Proceedings of the International Congress of Mathematicians

Rio de Janeiro 2018

VOLUME III

Invited Lectures

Boyan Sirakov
Paulo Ney de Souza
Marcelo Viana
Editors

SBM

World Scientific

Proceedings of the

International Congress of Mathematicians

Rio de Janeiro 2018

Editors
Boyan Sirakov, PUC – Rio de Janeiro
Paulo Ney de Souza, University of California, Berkeley
Marcelo Viana, IMPA – Rio de Janeiro
Technical Editors: Books in Bytes

Proceedings of the International Congress of Mathematicians
August 1 – 9, 2018, Rio de Janeiro, Brazil

Printed in Singapore

Published and Distributed by

World Scientific Publishing Co Pte Ltd
5 Toh Tuck Link
Singapore 596224

Tel: 65-6466-5775
Fax: 65-6467-7667
www.wspc.com
sales@wspc.com

ISBN: 978-981-3272-92-7 (volume print)
ISBN: 978-981-3272-87-3 (set print)
ISBN: 978-981-3272-88-0 (set ebook)

Proceedings of the **International Congress of Mathematicians**

Rio de Janeiro 2018

VOLUME III
Invited Lectures

Boyan Sirakov
Paulo Ney de Souza
Marcelo Viana
Editors

SBM

World Scientific

Contents

1 Logic and Foundations

2 Algebra

3 Number Theory

4 Algebraic and Complex Geometry

5 Geometry

6 Topology

8 Analysis and Operator Algebras

13 Combinatorics

14 Mathematical Aspects of Computer Science

16 Control Theory and Optimization

17 Mathematics in Science and Technology

18 Mathematics Education and Popularization of Mathematics

19 History of Mathematics

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1495–1528)

BOURGAIN–DELBAEN $\mathcal{L}_\infty$-SPACES, THE SCALAR-PLUS-COMPACT PROPERTY AND RELATED PROBLEMS

SPIROS A. ARGYROS AND RICHARD G. HAYDON

Abstract

We outline a general method of constructing $\mathcal{L}_\infty$-spaces, based on the ideas of Bourgain and Delbaen, showing how the solution to the Scalar-plus-Compact Problem, the embedding theorem of Freeman, Odell and Schlumprecht and other recent developments fit into this framework.

1 Introduction

Bourgain and Delbaen [1980] introduced a new class of separable Banach space that provided counterexamples to a number of open problems. On the one hand, each of these spaces $X_{\alpha,\beta}$ is an "$\mathcal{L}_\infty$-space", which means that its finite-dimensional structure resembles that of a $\mathcal{C}(K)$-space, and the dual $X^*_{\alpha,\beta}$ is isomorphic either to ℓ_1 or $\mathcal{C}[0,1]^*$; on the other hand, each infinite-dimensional subspace of $X_{\alpha,\beta}$ has a further subspace isomorphic to some ℓ_p $(1 \leq p < \infty)$, so that the global structure of $X_{\alpha,\beta}$ is very different from that of $\mathcal{C}(K)$-spaces, or complemented subspaces of such spaces.

Much more recently, it has become clear that the spaces $X_{\alpha,\beta}$ are a special case of a more general class, of what we now call Bourgain–Delbaen spaces (or BD-spaces for short). Taking a suitably general definition, it has been shown by Argyros, Gasparis, and Motakis [2016] that every separable $\mathcal{L}_\infty$-space is isomorphic to a BD-space, and BD-constructions have been used both in the solution of the Scalar-plus-Compact Problem by Argyros and Haydon [2011] and in the proof by Freeman, Odell, and Schlumprecht [2011] that every Banach space with separable dual embeds in an isomorphic predual of ℓ_1. The aim of this article is to sketch some general theory of BD constructions, trying to show how recent results fit together, and (hopefully) shedding light on some older theorems by presenting them in a BD framework. We do not have space for detailed proofs, especially

MSC2010: primary 46B03; secondary 46B26.

not of the scalar-plus-compact construction. Readers who are interested in an alternative account of this may wish to consult the Séminaire Bourbaki paper of Grivaux and Roginskaya [2014–2015], or the associated video available on YouTube. We include statements and sketched proofs of certain results that are as yet unpublished, including some from the long-promised paper of Argyros, Freeman, Haydon, Odell, Raikoftsalis, Schlumprecht, and Zisimopoulou [n.d.]. We also include a few easy proofs of known results, where we think that they may aid understanding of what is otherwise a rather abstract narrative.

The theory we sketch here builds on a rich body of literature starting with the paper of Tsirel'son [1974] that introduced the first example of a Banach space with no subspace isomorphic to c_0 or to any ℓ_p. Since Tsirelson norms are essential tools for us, we devote a section to sketching a fairly general theory of such norms, including an account of regular families of subsets of $\mathbb{N}$ and the crucial notion of asymptotic ℓ_1 structure.

We then move on to the mixed-Tsirelson spaces that have provided the raw material for the solution to the Distortion Problem by Odell and Schlumprecht [1994], the remarkable counterexamples of Gowers [1994a,b, 1996] and the whole theory of indecomposable and hereditarily indecomposable spaces as initiated by Gowers and Maurey [1993] and developed further by Argyros and Deliyanni [1997], Argyros and Felouzis [2000], and Argyros and Tolias [2004]. There are already some excellent surveys of this material in the literature and we shall try to avoid excessive duplication with papers such as Argyros and Tolias [2004], Maurey [1994] and Maurey [2003]. In particular, despite their importance, we shall say little about the notions of distortion and hereditary indecomposability.

Next we follow Argyros, Gasparis, and Motakis [2016] by introducing a notion of BD-space that is sufficiently general to embrace all separable $\mathscr{L}_\infty$-spaces, before restricting attention to "standard BD-spaces", a subclass amenable to detailed analysis and admitting norm estimates of Tsirelson type. We look at duality and subspace structure for such spaces and note that the spaces that are constructed in the Embedding Theorem of Freeman, Odell, and Schlumprecht [2011] may be taken to be of this type. We look at some natural notions of sub- and super-objects, and see how two known constructions, one due to Zippin [1977] and one to Cabello Sánchez, Castillo, Kalton, and Yost [2003] emerge naturally from BD methods. Finally, we move on to the scalar-plus-compact construction of Argyros and Haydon [2011] and more recent developments of this kind.

2 Notation

We write $\mathbb{N}$ for the set $\{1, 2, 3, \dots\}$ of natural numbers and ω for $\{0, 1, 2, \dots\}$, which we are usually considering as an ordinal. The cardinality of a set A is denoted $\#A$. When $f : X \to Y$ is a mapping and $A \subseteq X$, we write $f[A]$ for the image of A under f.

In a vector space, we write $\mathrm{sp}\langle x_i : i \in I\rangle$ for the linear span of a set $\{x_i : i \in I\}$; in a normed space $\overline{\mathrm{sp}}\langle x_i : i \in I\rangle$ denotes the corresponding closed linear span. The closed unit ball of a normed space X is denoted ball X. Two Banach spaces X and Y are said to be M-isomorphic if there exists a linear homeomorphism $T : X \to Y$ with $\|T\|\|T^{-1}\| \le M$. Sequences $(x_n)_{n\in\mathbb{N}}$ in X and $(y_n)_{n\in\mathbb{N}}$ in Y are said to be M-equivalent if there is an M-isomorphism $T : \overline{\mathrm{sp}}\langle x_n : n \in \mathbb{N}\rangle \to \overline{\mathrm{sp}}\langle y_n : n \in \mathbb{N}\rangle$ with $T(x_n) = y_n$ for all n. We say that a Banach space X is ℓ_p-saturated if there is a constant M such that every infinite-dimensional subspace of X has a further subspace M-isomorphic to ℓ_p.

A sequence $(M_n)_{n\in\mathbb{N}}$ of closed subspaces of a Banach space X is said to be a Schauder Decomposition if every $x \in X$ admits a unique representation as a norm-convergent sum $x = \sum_{n=1}^\infty x_n$ with $x_n \in M_n$ for all n. This is the case if and only if the linear direct sum $\bigoplus_{n\in\mathbb{N}} M_n$ is dense in X and there is a constant M such that for each N the usual projection $P_N : \bigoplus_{n\in\mathbb{N}} M_n \to \bigoplus_{n\le N} M_n$ extends to an operator on X with norm at most M. If each of the subspaces M_n is finite-dimensional we speak of a finite-dimensional decomposition, or f.d.d. Of course, if each of the subspaces is one-dimensional, with $M_n = \mathrm{sp}\langle x_n\rangle$, we get back to the well-known notion of a Schauder basis.

We use fairly standard notation for function spaces : $\mathbb{R}^\Gamma$ is the space of all scalar valued functions on a set Γ, and for $x \in \mathbb{R}^\Gamma$, the *support* of x is $\mathrm{supp}\, x = \{\gamma \in \Gamma : x(\gamma) \ne 0\}$; $\mathbb{R}^{(\Gamma)}$ is the space of functions of finite support, $\ell_\infty(\Gamma)$ the space of bounded functions, equipped with the supremum norm, $c_0(\Gamma)$ the norm closure of $\mathbb{R}^{(\Gamma)}$ in $\ell_\infty(\Gamma)$, and $\ell_p(\Gamma)$ the space of all functions x for which the norm $\|x\|_p$, defined by $\|x\|_p^p = \sum_{\gamma\in\Gamma} |x(\gamma)|^p$ is finite. For $1 \le p \le \infty$ we write ℓ_p (resp. ℓ_p^n) for $\ell_p(\Gamma)$ with $\Gamma = \mathbb{N}$ (resp. $\Gamma = \{1, 2, \dots, n\}$). When $\Gamma_1 \subset \Gamma_2$ we identify $\mathbb{R}^{\Gamma_1}$ with the subspace of $\mathbb{R}^{\Gamma_2}$ consisting of functions that vanish off Γ_1, and adopt the same convention for other function spaces.

According to the context, an element y of $\mathbb{R}^{(\Gamma)}$ may be regarded either as a vector, or as a functional acting on $\mathbb{R}^\Gamma$ via the duality $\langle y, x\rangle = \sum_\gamma x(\gamma)y(\gamma)$. If we are thinking of y as a functional, we shall generally employ a notation adorned with a star, writing for instance f^* instead of y. In particular, the element of $\mathbb{R}^\Gamma$ that takes the value 1 at a specific $\gamma \in \Gamma$ and is zero elsewhere may be denoted either e_γ if we are thinking of it as a vector, or as e_γ^* if we are thinking of it as the evaluation functional $x \mapsto \langle e_\gamma^*, x\rangle = x(\gamma)$.

We say that finite subsets $E_1, E_2, \dots$ of $\mathbb{N}$ are *successive*, and write $E_1 < E_2 < \cdots$ if $\max E_j < \min E_{j+1}$ for all j. If X is a space with a basis $(d_n)_{n\in\mathbb{N}}$ (resp. a finite-dimensional decomposition $(M_n)_{n\in\mathbb{N}}$) we define the *range* of a vector x, denoted $\mathrm{ran}\, x$, to be the minimal interval $I \subset \mathbb{N}$ such that $x \in \langle d_n : n \in I\rangle$ (resp. $x \in \bigoplus_{n\in I} M_n$). We say that vectors $x_1, x_2, \dots$ are successive, or that (x_n) is a *block sequence*, if $\mathrm{ran}\, x_1 < \mathrm{ran}\, x_2 < \cdots$, and that (x_n) is a *skipped* block sequence if $1 + \max \mathrm{ran}\, x_n < \min \mathrm{ran}\, x_{n+1}$ for all n. The closed linear span of a block sequence is called a *block subspace*. In the context of a normed space, we say that a sequence (x_n) is *normalized* if $\|x_n\| = 1$ for all n.

We say that an infinite-dimensional Banach space X is *indecomposable* if X cannot be expressed as the direct sum of two infinite-dimensional closed subspaces, and *hereditarily indecomposable* if every infinite-dimensional subspace of X is indecomposable. We recall that a bounded linear operator T on a Banach space X is *strictly singular* if there is no infinite dimensional subspace Y of X such that $T\restriction_Y$ is an isomorphism. We say that X has *few operators* if every bounded linear operator on X can be written $T = \lambda I + S$, where S is strictly singular. The Banach space is said to have *very* few operators, or to have the Scalar-plus-Compact Property if, in addition, every strictly singular operator on X is compact.

3 Regular families, Tsirelson norms and asymptotic ℓ_p spaces

Definition 3.1. We say that a collection $\mathcal{M}$ of finite subsets of $\mathbb{N}$ is *regular* if

1. $\mathcal{M}$ is *compact* (for the topology induced by the product topology on $\{0,1\}^{\mathbb{N}}$) and
2. $\mathcal{M}$ is *spreading*, i.e. if $M = \{m_1, m_2, \ldots, m_k\} \in \mathcal{M}$ and $n_j \geq m_j$ for all j then $N = \{n_1, n_2, \ldots, n_k\}$ is also in $\mathcal{M}$. (Such an N is called a *spread* of M.)

We note that a regular family is also *hereditary* in the sense that $N \subset M \in \mathcal{M}$ implies $N \in \mathcal{M}$, and that any compact family of subsets is contained in a regular family (namely the closure of the set of all its spreads).

Important examples of regular families include $\mathcal{A}_n = \{M \subset \mathbb{N} : \#M \leq n\}$, the Schreier family $\mathcal{S} = \{M \subset \mathbb{N} : \#M \leq \min M\}$ and the higher Schreier families $\mathcal{S}_\alpha$ introduced by Alspach and Argyros [1992] and defined for all countable ordinals α. There is an associative binary operation $*$ defined on the set of regular families by taking $\mathcal{M} * \mathcal{N}$ to be the set of all unions $\{M_1 \cup M_2 \cup \ldots M_n$ where $M_1, M_2, \ldots$ are successive members of $\mathcal{M}$ and $\{\min M_j : j \leq n\} \in \mathcal{N}$. As with any associative operation we can form powers $\mathcal{M}^{*n} = \mathcal{M} * \mathcal{M} * \cdots * \mathcal{M}$ (with n terms). For finite n the higher Schreier families are given by $\mathcal{S}_n = \mathcal{S}^{*n}$.

It follows from standard results (based on the Hahn–Banach Theorem or Ptak's combinatorial lemma) about weakly null sequences of continuous functions on a compact set that for every regular family and every $\epsilon > 0$ there exist n and a finite sequence $a_1, a_2, \ldots, a_n$ such that $a_j \geq 0$ for all j, $\sum_{j=1}^n a_j = 1$ while $\sum_{j \in M} a_j \leq \epsilon$ for all $M \in \mathcal{M}$. We call the vector $a = \sum_n a_n e_n$ a *basic convex combination* that is ϵ*-small* for $\mathcal{M}$. The papers of Alspach and Odell [1988] and Alspach and Argyros [1992] studied this phenomenon in greater detail, introducing ordinal indices that measure the speed of weak convergence of a sequence and the complexity of convex combinations obtained by the method of repeated averages. It was in this context that the special regular families $\mathcal{S}_\alpha$ ($\alpha < \omega_1$) were introduced in Alspach and Argyros [ibid.]. The same ideas yield a general result that is easy to state and sufficient for our present purposes.

Proposition 3.2. *Let $\mathcal{M}$ be a regular family. Then there is a regular family $\mathcal{M}^\#$ such that every maximal member N of $\mathcal{M}^\#$ is the support of a basic convex combination $a_{\mathcal{M},N,\epsilon}$ that is $2^{-\min N+1}$-small for $\mathcal{M}$.*

Let $\mathcal{M}$ be a regular family and let $E_1 < E_2 < \cdots < E_n$ be a sequence of successive subsets of $\mathbb{N}$. We say that the sequence $(E_j)_{j\le n}$ is $\mathcal{M}$-admissible if the set $\{\min E_j : j \le n\}$ is in $\mathcal{M}$ and that sequence $(f_j^*)_{j=1}^a$ in $\mathbb{R}^{(\mathbb{N})}$ is $\mathcal{M}$-admissible if the sequence of supports $\operatorname{supp} f_1^*, \ldots, \operatorname{supp} f_n^*$ is.

We are now ready to define Tsirelson norms. There are two ways to to do this: directly by an implicit functional equation for a norm $\|\cdot\|$ on $\mathbb{R}^{(\mathbb{N})}$, or by constructing a norming set $W \subset \mathbb{R}^{(\mathbb{N})}$ and defining $\|x\| = \sup_{f^*\in W} |\langle f^*, x\rangle|$. The first approach was introduced in Figiel and Johnson [1974] and often makes for elegant proofs; the second, closer in spirit to the original construction Tsirel'son [1974], is useful when a more delicate calculation based on an analysis of the functionals f^* is needed. In the context of Tsirilson spaces we shall follow the notation of Figiel and Johnson [1974], writing Ex for the vector given by

$$Ex(n) = \begin{cases} x(n) & \text{when } n \in E \\ 0 & \text{when } n \notin E. \end{cases}$$

Definition 3.3. Let $\mathcal{M}$ be a regular family and let $0 < \theta < 1$ be a real number. We define $W(\mathcal{M}, \theta)$ to be the minimal subset W of $\mathbb{R}^{(\mathbb{N})}$ with $\pm e_n^* \in W$ for all n and such that $\theta \sum_{j=1}^a f_j^* \in W$ whenever $f_1^*, \ldots, f_a^*$ is an $\mathcal{M}$-admissible sequence in W. We define the *Tsirelson* norm $\|\cdot\|_{T(\mathcal{M},\theta)}$ on $\mathbb{R}^{(\mathbb{N})}$ by

$$\|x\|_{T(\mathcal{M},\theta)} = \sup_{f^*\in W(\mathcal{M},\theta)} \langle f^*, x\rangle.$$

The space $T(\mathcal{M}, \theta)$ is then defined to be the completion of $\mathbb{R}^{(\mathbb{N})}$ with respect to this norm. An equivalent definition of the norm is to define $\|\cdot\|_{T(\mathcal{M},\theta)}$ directly as the smallest solution to the functional equation

$$\|x\| = \max\Big\{\|x\|_\infty\,,\ \theta \sup \sum_{j=1}^a \|E_j x\|\Big\},$$

where the supremum is taken over all $\mathcal{M}$-admissible sequences $E_1, \ldots, E_a$. *The* Tsirelson space T, as defined in Figiel and Johnson [ibid.] and studied extensively in Casazza and Shura [1989], is $T(\mathcal{S}, \frac{1}{2})$.

For general families $\mathcal{M}$, the above definition appears in the preprint ofArgyros and Deliyanni [1992], where results are obtained even in the case where $\mathcal{M}$ is only assumed to be compact (and not necessarily spreading). In the case of a regular family we have the following theorem, in which the second statement is due to Bellenot [1986].

Theorem 3.4. *Let $\mathcal{M}$ be a regular family and let $0 < \theta < 1$. If $\mathcal{M}$ has members of arbitrarily large finite cardinality then $T[\mathcal{M}, \theta]$ is reflexive with no subspace isomorphic to any space ℓ_p. If the members of $\mathcal{M}$ are of bounded cardinality then the unit vector basis of $T[\mathcal{M}, \theta]$ is equivalent to the usual basis of ℓ_p where $\theta = n^{-1/p'}$, $n = \max_{M \in \mathcal{M}} \#M$ and $1/p + 1/p' = 1$.*

We shall give a sketch proof of the first statement in this theorem, not because there is anything new in it (indeed it is very close to that given by Figiel and Johnson), but in order to give a very easy example of the use of special convex combinations, and to introduce the important notion of asymptotic ℓ_p structure.

Definition 3.5. Let X be a Banach space with a finite-dimensional Schauder decomposition $(M_n)_{n \in \mathbb{N}}$, and let $p \in [1, \infty]$. We say that (M_n) is *asymptotic ℓ_p* with constant $C > 1$, if for every n there exists N such that the sequence $(x_1, x_2, \dots, x_n)$ is C-equivalent to the unit vector basis of ℓ_p^n whenever $(x_j)_{j=1}^n$ is a normalized block sequence with $N \le \operatorname{ran} x_1 < \operatorname{ran} x_2 < \cdots \operatorname{ran} x_n$. We sometimes risk ambiguity by saying that the space X is asymptotic ℓ_p. There is a related, but weaker, notion close to what was called "asymptotic ℓ_p" by Maurey, Milman, and Tomczak-Jaegermann [1995]. To avoid ambiguity, we shall say that a finite-dimensional decomposition is *skipped-asymptotic ℓ_p* with constant C, if for every n there exists N such that the sequence $(x_1, x_2, \dots, x_n)$ is C-equivalent to the unit vector basis of ℓ_p^n whenever $(x_j)_{j=1}^n$ is a normalized skipped block sequence with $N \le \operatorname{ran} x_1$.

A space with a finite-dimensional decomposition that is *skipped-asymptotic ℓ_p* with $p < \infty$ cannot contain c_0 or ℓ_q for $q \neq p$. Theorem 3.4 thus follows from the next proposition.

Proposition 3.6. *Let $\mathcal{M}$ be a regular family with members of arbitrarily large finite cardinality and let $0 < \theta < 1$ be a real number. Then $T(\mathcal{M}, \theta)$ is reflexive and its usual unit-vector basis $(e_n)_{n \in \mathbb{N}}$ is asymptotic ℓ_1.*

Proof. For any n there is a set $M_0 \in \mathcal{M}$ of cardinality n; using the spreading property, there is a natural number N such that every set M with $\#M = n$ and $\min M \ge N$ is in $\mathcal{M}$. Now suppose that $x_1, \dots, x_n$ is a normalized block sequence with $\min \operatorname{ran} x_1 \ge N$ and set $x = \sum_{j=1}^n a_j x_j$; taking $E_j = \operatorname{ran} x_j$, we see that the sequence $E_1, E_2, \dots, E_n$

is $\mathcal{M}$-admissible, so that

$$\|\sum_{j=1}^{n} a_j x_j\| = \|x\| \geq \theta \sum_{j=1}^{n} \|E_j x\| = \theta \sum_{j=1}^{n} \|a_j x_j\| = \theta \sum_{j=1}^{n} |a_j|,$$

by the implicit formula for the norm. Thus $T(\mathcal{M}, \beta)$ is asymptotic ℓ_1.

Since $T(\mathcal{M}, \beta)$ has an unconditional basis, to show that it is reflexive it is enough to show that it has no subspace isomorphic to c_0 or ℓ_1, and c_0 is excluded by the asymptotic ℓ_1 property. By a result of James [1964], if a space has a subspace isomorphic to ℓ_1 it has a space nearly isometric to ℓ_1. So it will be enough for us to show that there is no normalized block sequence $(x_n)_{n\in\mathbb{N}}$ in $T(\mathcal{M}, \theta)$ satisfying the lower estimate

$$\|\sum_j a_j x_j\| \geq \theta' \sum_j |a_j|$$

for all scalars a_j, where $\theta' = \frac{1}{2}(1+\theta) < 1$.

To do this we shall employ an easy splitting lemma variants of which underlie many proofs about Tsirelson spaces, as well as the BD-spaces we shall be looking at later.

Lemma 3.7. *Let $\mathcal{M}$ be a regular family and let $E_1, \ldots, E_a$ be an $\mathcal{M}$-admissible sequence of subsets of $\mathbb{N}$. Let I be an interval in $\mathbb{N}$ and let $R_i = [p_i, q_i]$ ($i \in I$) be successive intervals in $\mathbb{N}$. Then we may write $I = I' \cup \bigcup_{k=0}^{a} I_k$ where*

$$\begin{aligned} I' &= \{i \in I : R_i \cap E_k = \varnothing \text{ for all } k\} \\ I_0 &= \{i \in I : R_i \cap E_k \neq \varnothing \text{ for more than one value of } k\} \\ I_k &= \{i \in I : R_i \cap E_l \neq \varnothing \text{ for } l = k \text{ but no other value of } l\}. \end{aligned}$$

The set $\{q_i : i \in I_0\}$ belongs to the family $\mathcal{M}$.

We now consider a normalized block sequence $(x_n)_{n\in\mathbb{N}}$ with $\operatorname{ran} x_j = R_j = [p_j, q_j]$ and form the sum $x = \sum_{i\in I} a_i x_i$ where the coefficients a_i are chosen so that $\sum_{i\in I} a_i e_{q_i}$ is a basic convex combination that is ϵ-small for $\mathcal{M} \cup \mathcal{A}_1$, where $\epsilon = \frac{1}{2}(1-\theta)$. What this means is that $\sum_{i\in J} a_i \leq \epsilon$ whenever $\{q_j : j \in J\} \in \mathcal{M}$.

By the implicit definition of the norm, either $\|x\| = \|x\|_\infty \leq \max_i |a_i| \leq \epsilon$ or there is an $\mathcal{M}$-admissible sequence $E_1, \ldots, E_a$ such that

$$\|x\| = \theta \sum_{k=1}^{a} \|E_k x\|.$$

Applying Lemma 3.7 we obtain

$$\begin{aligned}\|x\| &\le \sum_{i\in I_0}\|a_i x_i\| + \theta\sum_{k=1}^{a}\sum_{i\in I_k}\|a_i x_i\| \\ &\le \sum_{i\in I_0} a_i + \theta\sum_{i\in I} a_i \le \epsilon + \theta = \theta',\end{aligned}$$

because $\{q_j : j \in I_0\} \in \mathcal{M}$. □

The original Tsirelson space has a rich theory and, thanks to special properties of the Schreier family $\mathcal{S}$, we have a good understanding of its subspace structure; the reader is referred to Casazza and Shura [1989] for a comprehensive account of this material. Taken together with the generalized versions $T(\mathcal{S}_\alpha, \theta)$, the Tsirelson space has become much more than an isolated counterexample and now plays a key role in the general theory. For instance, the "subsequential $T(\mathcal{S}_\alpha, \theta)$-estimates" of Odell, Schlumprecht, and Zsák [2007] may be used to study the structure of general separable reflexive spaces.

4 Mixed Tsirelson spaces

The definition we are about to give of a "mixed Tsirelson space" might at first sight seem an idle generalization. But in fact Schlumprecht's space Schlumprecht [1991], which is of this type, opened the way to a new chapter in the theory, making possible Gowers's solutions to a group of previously intractable problems in Gowers [1994a,b, 1996], the theory of hereditarily indecomposable spaces introduced by Gowers and Maurey [1993] and the solution of the Distortion Problem in Hilbert space by Odell and Schlumprecht [1994].

Definition 4.1. Argyros and Deliyanni [1997] Let I be a countable set; for each $i \in I$ let $\mathcal{M}_i$ be a regular family and let $0 < \beta_i < 1$ be a real number. We define the norming set $W[(\mathcal{M}_i, \beta_i)_{i\in I}]$ to be the smallest subset of $\mathbb{R}^{(\mathbb{N})}$ that contains $\pm e_n^*$ for all $n \in \mathbb{N}$ and has the property that the functional f^* given by

$$f^* = \beta_i \sum_{r=1}^{a} f_r^*$$

is in W whenever $f_1^*, \dots, f_a^*$ is an $\mathcal{M}_i$-admissible sequence with $f_r^* \in W$ for all r. (Such a functional is said to have been created by an $(\mathcal{M}_i, \beta_i)$-*operation* or to be of *weight* β_i.) The space $T[(\mathcal{M}_i, \beta_i)_{i\in I}]$ is defined to be the completion of $\mathbb{R}^{(\mathbb{N})}$ with respect to the norm given by

$$\|x\| = \sup\{\langle f^*, x\rangle : f^* \in W[(\mathcal{M}_i, \beta_i)_{i\in I}]\}.$$

There is of course an alternative definition of the norm using an implicit formula.

It is possible to modify the definition of the norming set $W[(\mathcal{M}_i, \beta_i)_{i\in I}]$ by placing restrictions on the functionals f_r^* that are permitted in an $(\mathcal{M}_i, \beta_i)$-operation. One possibility is to insist that f_r^* must be a functional of a specific weight (determined by i and the preceding functionals $f_1^* \dots, f_{r-1}^*$); this is referred to as a *coding*, and is generally applied for i in some proper subset of I (typically, for i odd when $I = \mathbb{N}$). The idea of coding in this way can be traced back to the paper of Maurey and Rosenthal [1977]. Another possibility is to insist that f_r^* be an average $n_r^{-1} \sum_{k=1}^{n_r} g_{r,k}^*$ where the $g_{r,k}^*$ are successive elements of W; this approach, introduced by Odell and Schlumprecht [1995, 2000], is referred to as *saturation under constraints* and has found more recent applications in Argyros, Beanland, and Motakis [2013], Argyros and Motakis [2014], and Beanland, Freeman, and Motakis [2015].

The first example in the literature of what we now call a mixed Tsirelson space seems to be a space constructed by Tzafriri in Tzafriri [1979] to solve a delicate question about type and cotype. In our notation, this space is $T[(\mathcal{A}_n, \theta/\sqrt{n})_{n\in\mathbb{N}}]$. Schlumprecht's space Schlumprecht [1991] is $T[(\mathcal{A}_n, (\log_2(n+1))^{-1}_{n\in\mathbb{N}}]$. Most subsequent work on hereditary indecomposability and related topics has followed Argyros and Deliyanni [1997] by working not with the sequence $(\mathcal{A}_n)_{n\in\mathbb{N}}$ but with either a highly lacunary subsequence $(\mathcal{A}_{n_k})_{k\in\mathbb{N}}$ or a sequence $(\mathcal{S}_{\alpha_k})_{k\in\mathbb{N}}$ of higher Schreier families. For the applications we are considering here it is convenient to write $\beta_i = m_i^{-1}$ and make some standard assumptions about the natural numbers m_i and the families $\mathcal{M}_i$ that we work with.

Definition 4.2. Let $(m_i)_{i\in\mathbb{N}}$ be a sequence of natural numbers and let $(\mathcal{M}_i)_{i\in\mathbb{N}}$ be a sequence of regular families. We shall say that (m_i) and $(\mathcal{M}_i)$ satisfy the *Standard Assumptions* if

1. each m_i has the form 2^{l_i} with $l_i \in \mathbb{N}$, $l_1 \ge 2$ and $l_{i+1} \ge 2l_i$;
2. for each $i \in \mathbb{N}$, $\mathcal{M}_{i+1} \supseteq \mathcal{M}_i^{*l_{i+1}}$
3. for each $i \in \mathbb{N}$, and each maximal $N \in \mathcal{M}_{i+1}$ there is a basic convex combination supported on N that is m_{i+1}^{-2}-small for $(\mathcal{A}_4 * \mathcal{M}_i)^{*l_{i+1}}$.

The existence, given $\mathcal{M}_i$, of a family $\mathcal{M}_{i+1}$ satisfying (2) and (3) follows from Proposition 3.2. A basic convex combination having the property (3) will be called an $\mathcal{M}_{i+1}$-*special basic convex combination.*

The next proposition is central to the theory of mixed Tsirelson spaces and to many constructions that use coding to achieve hereditary indecomposability and few operators.

Proposition 4.3. *Let (m_i) and $(\mathcal{M}_i)$ satisfy the Standard Assumptions, let $h \geq 2$ be a natural number and let $a = \sum_{n\in N} a_n e_n$ be an $\mathcal{M}_h$-special basic convex combination. For a functional f^* of weight m_i^{-1} in $W[(\mathcal{M}_i, m_i^{-1})_{i\in I}]$ we have*

$$|\langle f^*, a\rangle| \leq \begin{cases} m_i^{-1} & \text{if } i \geq h \\ 2m_h^{-1}m_i^{-1} & \text{if } i < h \end{cases}$$

In particular, $\|a\| = m_h^{-1}$ and the equality $\langle f^*, a\rangle = \|a\|$ can be achieved only for a functional f^* of weight m_h^{-1}.

It is an important property of the Tsirelson space $T(\mathcal{S}, \beta)$ that every normalised block sequence is equivalent to a subsequence of the usual basis; in fact (x_n) is equivalent to (e_{q_n}), where $q_n = \max\operatorname{supp} x_n$ (see Casazza and Shura [1989]). No such result holds for arbitrary normalised block sequences in mixed Tsirelson spaces, and one could even say that it is this "failure" that underlies all their interesting properties. Indeed, provided the Standard Assumptions are satisfied, we can construct block sequences of vectors that do not behave like a subsequence of the usual basis. The following proposition is proved by a finite version of the "ℓ_1-improvement argument" of James [1964] that we used earlier in our discussion of Theorem 3.4.

Proposition 4.4. *Let (m_i) and $(\mathcal{M}_i)$ satisfy the Standard Assumptions, let $h \geq 2$ be a natural number and let $(w_n)_{n\in\mathbb{N}}$ be a block sequence in $T[(\mathcal{M}_i, \beta_i)_{i\in I}]$. Then there exists a maximal member M of $\mathcal{M}_h$, and a normalized finite block subsequence $(x_n)_{n\in I}$ of (w_n) such that $\{\max\operatorname{supp} x_n : n \in I\} = M$ and*

$$\Big\|\sum_{n\in I} \lambda_n x_n\Big\| \geq \tfrac{1}{4}\sum_{n\in I}|\lambda_n|,$$

for all scalars λ_n.

If $(x_n)_{n\in I}$ and M are as in the above proposition, so that $M = \{q_n : n \in I\}$ with $q_n = \max\operatorname{supp} x_n$, we may consider an $\mathcal{M}_h$-special basic convex combination $\sum_{n\in I} a_n e_{q_n}$, noting that

$$\Big\|\sum_{n\in I} a_n e_{q_n}\Big\| \leq m_h^{-1}$$

$$\Big\|\sum_{n\in I} a_n x_n\Big\| \geq \tfrac{1}{4},$$

by Propositions 4.3 and 4.4. Thus there is no constant C such that an inequality

$$\Big\|\sum_{n\in I} a_n x_n\Big\| \leq C\Big\|\sum_{n\in I} a_n e_{q_n}\Big\|$$

holds for an arbitrary normalized block sequence with $q_n = \max\operatorname{supp} x_n$.

However, there are special sequences for which we do have upper estimates of this form. These are the so-called "rapidly increasing sequences", or RIS. The idea is that from an arbitrary block sequence (w_n) we may construct, first of all, successive normalized finite block sequences $(x_n)_{n\in I_h}$ $(h = 2, 3, \dots)$ as in Proposition 4.4. We then form the $\mathcal{M}_h$-special convex combinations $y_h = \sum_{n\in I_h} a_n x_n$. A suitable subsequence ("rapidly increasing") of $(y_h)_{h\geq 2}$ is then a RIS and satisfies the desired norm estimates. We shall not go into detail either about rapidly increasing sequences or about the coding that can be introduced into a mixed Tsirelson construction to produce hereditary indecomposability and other exotic behaviours. We refer the reader to original papers such as Gowers and Maurey [1993] and Argyros and Deliyanni [1997], or the survey articles Maurey [1994] and Maurey [2003]. The idea in brief is that, with suitable coding, a construction of this kind results in a space X such that for every bounded linear operator T on X, there is a scalar λ such that $\|Tx_n - \lambda x_n\| \to 0$ for every RIS (x_n). Since every block subspace can be shown to contain a RIS, $T-\lambda I$ is strictly singular. The space X constructed in this way thus has few operators. In a later section we shall see what coding means in the context of a BD construction, and attempt to show why we are able to get from strict singularity to compactness of an operator $T - \lambda I$.

5 $\mathcal{L}_\infty$ spaces and Bourgain–Delbaen constructions

The $\mathcal{L}_p$-spaces were introduced some 50 years ago by Lindenstrauss and Pełczyński [1968] and were studied further in Lindenstrauss and Rosenthal [1969] (see also Nielsen and Wojtaszczyk [1973]). They provide an early, and striking, example of how conditions placed on finite-dimensional subspaces can have strong consequences for the isomorphic structure of an infinite-dimensional space, and remain one of the key areas of interest in Banach spaces.

Definition 5.1. Let X be a Banach space, let $1 \leq p \leq \infty$ and let $M > 1$ be a real number. We say that X is a $\mathcal{L}_{p,M}$-space if for every finite-dimensional subspace E of X there is a finite-dimensional subspace $F \supseteq E$ which is M-isomorphic to $\ell_p^{\dim F}$. If X is $\mathcal{L}_{p,M}$ for some M we say it is a $\mathcal{L}_p$-space.

Despite their definition in terms of finite-dimensional structure, the $\mathcal{L}_p$-spaces have many good infinite-dimensional properties, of which we now recall a few. For $1 \leq p \leq \infty$ every separable $\mathcal{L}_p$-space has a Schauder basis, and when $p < \infty$ every such space contains an isomorphic copy of ℓ_p. When $1 < p < \infty$ a separable Banach space is $\mathcal{L}_p$ if and only if it is isomorphic to a non-hilbertian complemented subspace of the Lebesgue space

$L_p(0,1)$. A useful characterization of $\mathscr{L}_\infty$-spaces in terms of extensions of compact operators was given in Lindenstrauss and Rosenthal [1969] (see also Theorem 4.2 of Zippin [2003]).

Proposition 5.2. *A Banach space X is an $\mathscr{L}_{\infty,M}$-space X if and only if, for every Banach space Z, every closed subspace $Y \subset Z$ and every compact operator $K : Y \to X$, there is a compact extension $L : Z \to X$ with $\|L\| \le M\|K\|$.*

In other respects, the structure of $\mathscr{L}_\infty$-spaces is more complicated than what we have for the case $p < \infty$. One of the main achievements of the original paper of Bourgain and Delbaen [1980] was to exhibit $\mathscr{L}_\infty$-spaces without subspaces isomorphic to c_0. We are now ready to look at the Bourgain–Delbaen construction, and its generalizations, in more detail.

It is immediate from the definition that a separable Banach space X is a $\mathscr{L}_\infty$-space if and only if there is a constant M and an increasing sequence of finite-dimensional subspaces $E_1 \subseteq E_2 \subseteq \cdots$ with $\overline{\bigcup_n E_n} = X$ and such that, for each n, E_n is M-isomorphic to $\ell_\infty(\Gamma_n)$ for some finite set Γ_n. In this set-up, the inclusion $E_n \hookrightarrow E_{n+1}$ corresponds to an isomorphic embedding $i_{n+1,n} : \ell_\infty(\Gamma_n) \to \ell_\infty(\Gamma_{n+1})$, and the structure of a separable $\mathscr{L}_\infty$-space is determined by this sequence of embeddings.

There is a particular class of isomorphic embeddings that is convenient to work with: we say that $i : \ell_\infty(\Gamma_1) \to \ell_\infty(\Gamma_2)$ is an *extension operator* if Γ_1 is a subset of Γ_2 and, for all $u \in \ell_\infty(\Gamma_1)$, $(iu)\restriction_{\Gamma_1} = u$. A quick way to describe Bourgain–Delbaen spaces is to say that they are $\mathscr{L}_\infty$-spaces constructed from a sequence of extension operators. Before giving a formal definition, we briefly forget about norms and boundedness, and consider the (very easy) linear algebra of such a sequence of extension operators.

Let $\Gamma_1 \subset \Gamma_2 \subset \cdots$ be an increasing sequence of finite sets, with $\Gamma = \bigcup_n \Gamma_n$; for each n, let r_n be the restriction mapping $\mathbb{R}^\Gamma \to \mathbb{R}^{\Gamma_n}$. We shall say that a sequence $(i_n)_{n\in\mathbb{N}}$ of linear mappings $i_n : \mathbb{R}^{\Gamma_n} \to \mathbb{R}^\Gamma$ is a *compatible sequence of extension mappings* if the following are true:

1. each i_n is an extension mapping, that is to say $r_n i_n r_n = r_n$;
2. the compatibility condition $i_n r_n i_m = i_m$ holds whenever $m < n$.

Given such a compatible sequence, each of the "one-step" mappings $i_{n+1,n} = r_{n+1} i_n$ is an extension mapping $\mathbb{R}^{\Gamma_n} \to \mathbb{R}^{\Gamma_{n+1}}$, and conversely, if we are given a sequence of one-step extension mappings $i_{n+1,n}$ as in the earlier discussion, there is a unique compatible family of extensions $i_n : \mathbb{R}^{\Gamma_n} \to \mathbb{R}^\Gamma$ satisfying $i_{n+1,n} = r_{n+1} i_n$.

We now introduce some notation that will be used consistently in the rest of this section. We consider a countably infinite set Γ, expressed as the union of an increasing sequence of finite subsets Γ_n, and write Δ_n for the difference set $\Gamma_n \setminus \Gamma_{n-1}$ when $n > 1$; for the case

$n = 1$ we set $\Delta_1 = \Gamma_1$. We say that an element γ is of *rank n* if $\gamma \in \Delta_n$. We suppose Γ to be equipped with a compatible sequence of linear extension mappings i_n. The images $E_n = i_n[\mathbb{R}^{\Gamma_n}]$ form an increasing sequence of subspaces of $\mathbb{R}^\Gamma$ and, for each n, we may define a projection $P_n = i_n r_n$ from $\mathbb{R}^\Gamma$ onto E_n. The dual projection P_n^* takes $\mathbb{R}^{(\Gamma)}$ onto $\mathbb{R}^{\Gamma_n}$, regarded as a subspace of $\mathbb{R}^{(\Gamma)}$. When $\gamma \in \Delta_{n+1}$ for some n we define $c_\gamma^* = P_n^* e_\gamma^*$, a functional supported on Γ_n, while if $\gamma \in \Gamma_1$ we set $c_\gamma^* = 0$. In either case we define $d_\gamma^* = e_\gamma^* - c_\gamma^*$. For any n and any $\gamma \in \Delta_n$ we define $d_\gamma = i_n e_\gamma \in \mathbb{R}^\Gamma$.

Proposition 5.3. *Let Γ be a countably infinite set, equipped with a compatible sequence of linear extension mappings as in the discussion above. The family $(d_\gamma^*)_{\gamma\in\Gamma}$ is an algebraic basis of $\mathbb{R}^{(\Gamma)}$ and d_γ ($\gamma \in \Gamma$) are the unique elements of $\mathbb{R}^\Gamma$ such that $(d_\gamma, d_\gamma^*)_{\gamma\in\Gamma}$ is a biorthogonal system. The extension mappings i_n, and the projections P_n and P_n^* are given by*

$$i_n u = \sum_{\gamma\in\Gamma_n} \langle d_\gamma^*, u\rangle d_\gamma \qquad (u \in \mathbb{R}^{\Gamma_n})$$

$$P_n x = \sum_{\gamma\in\Gamma_n} \langle d_\gamma^*, x\rangle d_\gamma \qquad (x \in \mathbb{R}^{\Gamma})$$

$$P_n^* f^* = \sum_{\gamma\in\Gamma_n} \langle f^*, d_\gamma\rangle d_\gamma^* \qquad (f^* \in \mathbb{R}^{(\Gamma)}).$$

If we think of Δ_m as being "to the left" of Δ_n when $m < n$, then $(d_\gamma^*)_{\gamma\in\Gamma}$ is a left-triangular basis of $\mathbb{R}^{(\Gamma)}$, while the family $(d_\gamma)_{\gamma\in\Gamma}$ is right-triangular.

Definition. Let Γ be a countably infinite set, expressed as the union of an increasing sequence of finite subsets Γ_n and equipped with a compatible family of linear extension mappings $i_n : \mathbb{R}^{\Gamma_n} \to \mathbb{R}^\Gamma$. We shall say that Γ is a *Bourgain–Delbaen set*, or more briefly a *BD-set*, if the mappings i_n take values in $\ell_\infty(\Gamma)$ and are uniformly bounded as operators from $\ell_\infty(\Gamma_n)$ to $\ell_\infty(\Gamma)$. We define $X(\Gamma)$ to be the closure in $\ell_\infty(\Gamma)$ of the union $\bigcup_{n\in\mathbb{N}} i_n[\ell_\infty(\Gamma_n)]$ and call $X(\Gamma)$ a *Bourgain–Delbaen space*.

When Γ is a BD-set then for any n and any $u \in \ell_\infty(\Gamma_n)$ we have $\|u\| \le \|i_n u\| \le M\|u\|$, where $M = \sup_n \|i_n\|$. Thus each of the increasing sequence of subspaces $E_n = i_n[\ell_\infty(\Gamma_n)]$ is M-isomorphic to $\ell_\infty(\Gamma_n)$, so that $X(\Gamma)$ is an $\mathscr{L}_{\infty,M}$-space. The subspaces of $X(\Gamma)$ defined by $M_n = i_n[\ell_\infty(\Delta_n)] = \mathrm{sp}\langle d_\gamma : \gamma \in \Delta_n\rangle$ form a Schauder decomposition of $X(\Gamma)$, the associated projection onto $M_1 \oplus M_2 \oplus \cdots \oplus M_n$ being P_n. In fact, suitably ordered, the vectors d_γ form a Schauder basis of $X(\Gamma)$ but we do not often need to use this finer structure. Similarly, the functionals d_γ^* form a ("left-triangular") Schauder basis of $\ell_1(\Gamma)$, but it is usually convenient to work with a coarser structure, the Schauder decomposition formed by the subspaces $M_n^* = \mathrm{sp}\langle d_\gamma^* : \gamma \in \Delta_n\rangle$; the corresponding projection onto the finite direct sum $M_1^* \oplus M_2^* \oplus \cdots \oplus M_n^* = \mathrm{sp}\langle d_\gamma^* : \gamma \in \Gamma_n\rangle = \ell_1(\Gamma_n)$

is P_n^*. It is convenient in this context to write $\Gamma_0 = \varnothing$ and P_0 (resp. P_0^*) for the zero operator on $\ell_\infty(\Gamma)$ (resp. $\ell_1(\Gamma)$). When $E = [m,n]$ is an interval in $\mathbb{N}$ we shall write $P_E = P_n - P_{m-1}$ and $P_E^* = P_n^* - P_{m-1}^*$. When we talk about "block-sequences" in $X(\Gamma)$ it will always be with respect to the above Schauder decomposition: thus the *range* $\operatorname{ran} x$ of of an vector x in $X(\Gamma)$ is the smallest interval E such that $P_E x = x$ and (x_n) is a block-sequence if $\operatorname{ran} x_n < \operatorname{ran} x_{n+1}$ for all n. We adopt similar notation for the range of a functional $f^* \in \ell_1$, but we give the word "support" its usual meaning $\operatorname{supp} f^* = \{\gamma \in \Gamma : f^*(\gamma) \neq 0\}$.

Since the extension mappings that are used to build BD-spaces form a rather special subclass of the class of all isomorphic embeddings of finite-dimensional ℓ_∞-spaces, it might be natural to guess that BD-spaces form a rather special sort of $\mathscr{L}_\infty$-space. So the following recent result of Argyros, Gasparis, and Motakis [2016] is perhaps surprising.

Theorem 5.4. *Every infinite-dimensional separable $\mathscr{L}_\infty$-space is isomorphic to a BD-space.*

We have already noted that a BD-structure on a set Γ is determined by specifying either the extension operators i_n or the functionals c_γ^*. When we use this method to carry out interesting constructions, it is usually most convenient to work with the c_γ^* and we need a criterion for norm-boundedness of the mappings i_n expressed in terms of these functionals.

Proposition 5.5. *Let $\Gamma = \bigcup_n \Gamma_n$ be a set equipped with a compatible sequence of extension mappings $(i_n)_{n\in\mathbb{N}}$ and let $M \geq 1$. Using our standard notation, the following are equivalent:*

(1) *Γ is a BD-set with constant M; this is to say $\|i_n\| \leq M$ for all n;*
(2) *For the norm of operators on $\ell_1(\Gamma)$, $\|P_n^*\| \leq M$ for all n;*
(3) *For every n, every $m < n$ and every $\gamma \in \Delta_n$, $\|P_m^* c_\gamma^*\|_1 \leq M$.*

As well as introducing the idea of building $\mathscr{L}_\infty$-spaces by successive extensions, the original paper of Bourgain and Delbaen gave a neat way to construct functionals c_γ^* satisfying condition (3) above. Modifying slightly the definitions that have appeared in earlier papers, and eliminating some special cases, we give a definition that captures the crucial idea.

Definition 5.6. Let Γ be a countably infinite set, equipped as usual with a compatible sequence of linear extension mappings (i_n). Let $\beta < \frac{1}{2}$ be a positive constant and let n be a natural number. We shall say that an element c^* of $\ell_1(\Gamma_n)$ is a *BD-functional*, with weight (at most) β, if c^* has one of the forms

$$c^* = \begin{cases} \beta P_{[s,n]}^* b^* & \text{or} \\ \alpha e_\xi^* + \beta P_{[s,n]}^* b^* \end{cases} \tag{5-1}$$

with $0 \le \alpha \le 1$, $s \le n+1$, $b^* \in \operatorname{ball} \ell_1(\Gamma_n \setminus \Gamma_{s-1})$ and (in the second case) $\xi \in \Delta_m$ for some $m < s$. We note that in the case where $s = n+1$, we have $b^* = 0$, so that c^* is either 0 or αe^*_ξ. A trivial generalization of the proof given by Bourgain and Delbaen [1980] gives us the following theorem.

Theorem 5.7. *Let Γ be a countably infinite set, equipped with a compatible sequence of linear extension mappings (i_n). Suppose that there is a constant $\beta < \frac{1}{2}$ such that, for each n and each $\gamma \in \Delta_{n+1}$, the functional c^*_γ is a BD-functional with weight at most β. Then Γ is a BD-set with $\sup_n \|i_n\| \le M = (1-2\beta)^{-1}$.*

Bourgain and Delbaen considered pairs of scalars α, β with $\beta < \frac{1}{2}$, $\alpha \le 1$ and $\alpha + \beta > 1$, constructing for each such pair an $\mathcal{L}_\infty$-space $X_{\alpha,\beta}$ with the Radon–Nikodým Property; in particular, $X_{\alpha,\beta}$ has no subspace isomorphic to c_0. When $\alpha = 1$ the space $X_{1,\beta}$ has the Schur Property and in particular is ℓ_1-saturated. For the case $\alpha < 1$, Haydon [2000] established ℓ_p-saturation, where $\alpha^q + \beta^q = 1$ and $1/p + 1/q = 1$. For a space $X(\Gamma)$ to have these properties (for a given pair α, β) it is sufficient that

1. for every $\gamma \in \Gamma$, c^*_γ is a BD functional with the given values of α and β;
2. for every $m < n < p$, every $\xi \in \Delta_m$, every $\eta \in \Delta_n$ and every choice of sign $\pm$ there exists $\gamma \in \Delta_p$ with $c^*_\gamma = \alpha e^*_\xi \pm \beta P^*_{(m,\infty)} e^*_\eta$,

It is easy to see that, for a general BD-space $X(\Gamma)$, the evaluation functionals e^*_γ form a system equivalent to the usual basis of $\ell_1(\Gamma)$. So $X(\Gamma)^*$ has a subspace $\overline{\operatorname{sp}}\langle e^*_\gamma : \gamma \in \Gamma\rangle$ naturally isomorphic to $\ell_1(\Gamma)$. It was Alspach [2000] who first observed that for the spaces $X_{\alpha,\beta}$ of Bourgain and Delbaen, when $\alpha < 1$, this subspace makes up the whole of the dual space. It seems hard to arrive at straightforward conditions on the functionals c^*_γ for this to be true in a general BD construction.

There remain open problems about the original BD spaces $X_{\alpha,\beta}$, for instance whether $X_{\alpha,\beta}$ and $X_{\alpha',\beta'}$ are non-isomorphic when α, β and α', β' are distinct pairs with $\alpha^q + \beta^q = 1 = \alpha'^q + \beta'^q$. But subsequent developments have concentrated on constructions involving BD-functionals with $\alpha = 1$. In the next section we sketch a fairly general framework within which a lot of constructions are possible, and where precise analysis can be carried out, including an exact description of the dual space.

6 Standard BD-spaces

We shall say that a BD-set Γ is a *standard* BD-set if there is a constant $\beta \in (0, \frac{1}{2})$, called the *weight* of Γ, such that, for each n and each $\gamma \in \Delta_{n+1}$, we have

$$c^*_\gamma = \begin{cases} \beta P^*_{[s,\infty)} b^* & \text{or} \\ e^*_\xi + \beta P^*_{[s,\infty)} b^* \end{cases}$$

with $s \le n+1$, $b^* \in \operatorname{ball} \ell_1(\Gamma_n \setminus \Gamma_{s-1})$ and (in the second case) $\xi \in \Gamma_{s-1}$. We call s the "cut", b^* the "top" and ξ (when it exists) the "base" of γ.

The key tool in the study of the structure of standard BD-spaces is what we call the *evaluation analysis* which expresses the evaluation e^*_γ as a sum of terms that are adapted to the finite-dimensional decomposition $(M^*_n)_{n\in\mathbb{N}}$ of $\ell_1(\Gamma)$. If base γ is undefined this is easy to write down:

$$e^*_\gamma = c^*_\gamma + d^*_\gamma = \qquad \beta P^*_{[s,\infty)} b^* + d^*_\gamma.$$

If $\xi = \operatorname{base}\gamma$ is defined then we have

$$e^*_\gamma = c^*_\gamma + d^*_\gamma = e^*_\xi + \beta P^*_{[s,\infty)} b^* + d^*_\gamma,$$

and we may continue by expressing e^*_ξ as $c^*_\xi + d^*_\xi$ and so on. What we end up with is the following

Proposition 6.1. *Let Γ be a standard BD-space of weight β and let γ be an element of Γ. Then there exist a natural number a and elements $\xi_1, \xi_2, \dots, \xi_a$ of Γ such that* base ξ_1 *is undefined, $\xi_a = \gamma$ and $\xi_k = \operatorname{base}\xi_{k+1}$ when $1 \le k < a$. We have the evaluation analysis*

$$e^*_\gamma = \sum_{k=1}^{a} (\beta P^*_{[s_k,\infty)} b^*_k + d^*_{\xi_k}),$$

where $s_k = \operatorname{cut}\xi_k$ and $b^*_k = \operatorname{top}\xi_k$.

A good way to investigate duality of standard BD spaces is to introduce a tree-order $\preccurlyeq$ on Γ. We can define this recursively by saying that $\xi \preccurlyeq \gamma$ if and only if either $\xi = \gamma$ or $\xi \preccurlyeq \operatorname{base}\gamma$. What this amounts to is that the elements ξ with $\xi \preccurlyeq \gamma$ are exactly the $\xi_1, \xi_2, \dots, \xi_a$ that occur in the evaluation analysis of Proposition 6.1. The elements γ with no base are minimal in this tree. In accordance with standard terminology, we shall say that a standard BD-set is *well-founded* if it has no infinite branch, that is to say, if there is no infinite sequence $(\xi_n)_{n\in\mathbb{N}}$ such that $\xi_n = \operatorname{base}\xi_{n+1}$ for all n.

Proposition 6.2. *Let Γ be a standard BD-set. Then $X(\Gamma)^* = \overline{\operatorname{sp}}\langle e^*_\gamma : \gamma \in \Gamma\rangle$ if and only if Γ is well-founded.*

One of the implications in the above proposition is easy to see: if $\beta = \{\xi_1 \prec \xi_2 \prec \xi_3 \prec \cdots\}$ is an infinite branch of the tree Γ then we can define a functional f^* by $\langle f^*, x\rangle = \lim_{n\to\infty} x(\xi_n)$. This functional, which it is natural to denote by e^*_β, is not in $\ell_1(\Gamma)$ since $\langle e^*_\beta, d_{\xi_n}\rangle = 1$ for all n, while $\langle g^*, d_{\xi_n}\rangle \to 0$ as $n \to \infty$ for any $g^* \in \ell_1(\Gamma)$. One way of proving the converse implication uses the fact that the set of extreme points of the unit ball of $X(\Gamma)^*$ is contained in the weak*-closure of the evaluations e^*_γ and applies Choquet's integral representation theorem.

Although in current applications we have need of only one very special sort of BD-set with infinite branches it may be interesting to note that there is a general duality result here too. We first note that the set B of infinite branches of Γ has a natural topology as a Polish space and that for any bounded Radon measure μ on B we may define a functional $R^*\mu \in X(\Gamma)^*$ by $\langle R^*\mu, x\rangle = \int_{\mathrm{B}} \langle e^*_\beta, x\rangle \mathrm{d}\mu(\beta)$. Subject to modest additional hypotheses on weights (of which the following proposition gives one example) we have a nice extension of Proposition 6.2, again provable by a Choquet argument.

Proposition 6.3. *Let Γ be a standard BD set of weight $\beta < \frac{1}{4}$. The dual space $X(\Gamma)^*$ is naturally isomorphic to $\ell_1(\Gamma) \oplus \mathcal{M}^{\mathrm{b}}(\mathrm{B})$.*

It is of course an elementary fact that every separable Banach space is isomorphic to a quotient of ℓ_1. The striking and unexpected result proved by Freeman, Odell, and Schlumprecht [2011] is that, when Y^* is a separable *dual* space, the quotient operator can be chosen to be the dual of an isomorphic embedding of Y into a space X with X^* isomorphic to ℓ_1. Unsurprisingly, they use a BD-construction, though one that does not quite fit with our definition of a standard BD-space. Nonetheless, their idea carries over to this framework yielding the following.

Theorem 6.4. *Let Y be a Banach space with separable dual and let $\beta < \frac{1}{2}$ be a positive real number. Then Y embeds isomorphically into a standard BD-space $X(\Gamma)$ of weight β with $X(\Gamma)^*$ naturally isomorphic to $\ell_1(\Gamma)$.*

7 Tsirelson-type estimates for standard BD spaces

Let Γ be a standard BD-set and let $\gamma \in \Gamma$ be an element with evaluation analysis

$$e^*_\gamma = \sum_{k=1}^{a} (\beta P^*_{[s_k,\infty)} b^*_k + d^*_{\xi_k})$$

as in Proposition 6.1. We shall say that a is the *age* of γ and that the set $\{s_1, s_2, \dots, s_a\}$ is the *history* of γ. If all branches of the tree structure on Γ are finite then the collection of all histories $\{\operatorname{hist}\gamma : \gamma \in \Gamma\}$ is a compact family of finite subsets of $\mathbb{N}$ and so is contained in some regular family $\mathcal{M}$.

It becomes clear that there may be a connection with Tsirelson norms if we consider γ with the above evaluation analysis and a sequence $(x_i)_{i=1}^a$ in $X(\Gamma)$ such that $\operatorname{ran} x_i \subseteq$

$[s_i, \operatorname{rank} \xi_i)$. We then have $\langle d^*_{\xi_k}, x_i\rangle = 0$ for all i and k, and

$$\langle \beta P^*_{[s_k,\infty)} b^*_k, x_i\rangle = \begin{cases} \langle b^*_k, x_k\rangle & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases}$$

This leads to

$$\langle e^*_\gamma, \sum_{i=1}^a x_i\rangle = \beta \sum_{i=1}^a \langle b^*_i, x_i\rangle \le \beta \sum_{i=1}^a \|x_i\|$$

while if the b^*_i can be chosen i such a way that $\langle b^*_i, x_i\rangle \ge \delta\|x_i\|$, we obtain

$$\langle e^*_\gamma, \sum_{i=1}^a x_i\rangle \ge \beta\delta \sum_{i=1}^a \|x_i\|,$$

a formula highly suggestive of Tsirelson norms.

Of course dealing with the general case where the sequence (x_i) does not fit so nicely with the evaluation analysis of γ requires some extra effort, but by making use of an appropriate version of Lemma 3.7 we obtain an upper Tsirelson estimate that is valid for all block sequences.

Theorem 7.1 (First Basic Inequality). *Let $\mathcal{M}$ be a regular family and let Γ be a standard BD-set of weight β such that $\operatorname{hist}\gamma \in \mathcal{M}$ for all $\gamma \in \Gamma$. Then for any normalized block sequence $(x_j)_{j=1}^n$ in $X(\Gamma)$ we have the upper Tsirelson estimate*

$$\|\sum_{j=1}^n a_j x_j\| \le \beta^{-1} \|\sum_{j=1}^n a_j e_{q_j}\|_{T(\mathcal{A}_3 * \mathcal{M}, \beta)},$$

where $q_j = \max \operatorname{ran} x_j$.

The above inequality gives an alternative way of proving Proposition 6.2, and also provides insight into the subspace structure of the "Bourgain–Tsirelson" space introduced in Haydon [2006].

Proposition 7.2. *Let Γ be a standard BD-set of weight $\beta < \frac{1}{2}$ with the following properties:*

1. *for every $\gamma \in \Gamma$ the history $\operatorname{hist}\gamma$ is in the Schreier family $\mathcal{S}$;*
2. *if $s < n$ are natural numbers, ξ, η are elements of Γ with $\operatorname{rank}\xi < s \le \operatorname{rank}\eta < n$ and $\{s\} \cup \operatorname{hist}\xi \in \mathcal{S}$ then for each choice of sign $\pm$ there exists an element γ with $\operatorname{rank}\gamma = n$, $\operatorname{base}\gamma = \xi$ and $\operatorname{top}\gamma = \pm e^*_\eta$.*

Then $X(\Gamma)$ is skipped-asymptotic ℓ_1 and every infinite-dimensional subspace of $X(\Gamma)$ contains a sequence equivalent to some subsequence of the unit-vector basis of the Tsirelson space $T(\mathcal{S},\beta)$.

A sketch of the proof proceeds as follows. Consider a normalized block sequence $(x_n)_{n\in\mathbb{N}}$ in $X(\Gamma)$, setting $\operatorname{ran} x_n = [p_n, q_n]$. Assumption (1), together with the First Basic Inequality, yields an upper estimate of the form

$$\|\sum_{j=1}^n a_j x_j\| \le C\|\sum_{j=1}^n a_j e_{q_j}\|_{T(\mathcal{A}_3 * \mathcal{S},\beta)}.$$

Provided (x_n) is a skipped-block sequence, Assumption (2) allows us to construct an element $\gamma \in \Gamma$ whose analysis does fit nicely with (x_n), leading to a lower estimate

$$\|\sum_{j=1}^n a_j x_j\| \ge \beta(1-2\beta)\sum_{j=1}^n |a_n|$$

for a skipped block sequence with $\min \operatorname{ran} x_1 \ge n$. This gives the skipped asymptotic ℓ_1 property. For "sufficiently skipped" block sequences we can get a better lower estimate

$$\|\sum_{j=1}^n a_j x_j\| \ge (1-\epsilon)\|\sum_{j=1}^n a_j e_{p_j}\|_{T(\mathcal{S},\beta)}.$$

To finish, we need two standard results from Casazza and Shura [1989] about the standard Tsirelson space: first that the $T(\mathcal{A}_3 * \mathcal{S},\beta)$-norm is equivalent to the $T(\mathcal{S},\beta)$-norm and secondly that the sequences $(e_{p_n})_{n\in\mathbb{N}}$ and $(e_{q_n})_{n\in\mathbb{N}}$ are equivalent in $T(\mathcal{S},\beta)$ whenever $p_1 \le q_1 < p_2 \le q_2 < \cdots$.

It may be helpful at this point to describe explicitly a BD set satisfying the conditions of Proposition 7.2. The recursive definition that follows is the simple prototype on which more complicated BD constructions, including the one in Argyros and Haydon [2011] are modeled.

Definition 7.3. The set Γ^{BT} is defined as $\bigcup_{n\in\mathbb{N}} \Delta_n$, where the sets Δ_n and the function hist are given recursively by setting $\Delta_1 = \{1\}$, $\Gamma_n = \bigcup_{m\le n} \Delta_m$ and

$$\begin{aligned}\Delta_{n+1} = \{(n+1, 0, s, \pm, \eta) &: 1 \le s \le n,\ \eta \in \Gamma_n \setminus \Gamma_{s-1}\}\\ &\cup \{(n+1, \xi, s, \pm, \eta) : 2 \le s \le n,\ \xi \in \Gamma_{s-1},\ \{s\} \cup \operatorname{hist}\xi \in \mathcal{S}, \eta \in \Gamma_n \setminus \Gamma_{s-1}\},\\ \operatorname{hist}(n+1, 0, s, \pm, \eta) = \{s\} &\qquad \operatorname{hist}(n+1, \xi, s, \pm, \eta) = \{s\} \cup \operatorname{hist}\xi.\end{aligned}$$

We then introduce a standard BD structure (of weight $\beta < \frac{1}{2}$) on Γ^{BT} by setting

$$\begin{aligned} &\operatorname{base}(n+1, 0, s, \pm, \eta) = \text{undefined} && \operatorname{top}(n+1, *, s, \pm, \eta) = \pm e_\eta^* \\ &\operatorname{base}(n+1, \xi, s, \pm, \eta) = \xi && \operatorname{cut}(n+1, *, s, \pm, \eta) = s\,. \end{aligned}$$

8 Two recent examples

We have already mentioned the original spaces $X_{\alpha,\beta}$ of Bourgain and Delbaen. In the case where $\alpha = 1$, these are standard BD-spaces in our terminology, and the branches of the tree structure of the corresponding BD-set are all infinite. When $\alpha < 1$, the ℓ_p-saturated space $X_{\alpha,\beta}$ is not a standard BD-space, and it is not immediately obvious how to construct a standard BD-space that is ℓ_p-saturated for $1 < p < \infty$. Such a construction has, however, been carried out by Gasparis, Papadiamantis, and Zisimopoulou [2010]. Rephrasing their result, we have the following.

Theorem 8.1. *For every real number p with $1 < p < \infty$ there is a well-founded standard BD-set Γ such that $X(\Gamma)$ is ℓ_p-saturated.*

It is worth noting that in this example there is an upper bound N on the ages of members of Γ (that is to say on the lengths of branches in the tree-structure of Γ). Another example with this property has been constructed by Argyros, Gasparis, and Motakis [2016]; this space lies at the opposite end of the spectrum from the space $X(\Gamma^{\mathrm{BT}})$, which is skipped-asymptotic ℓ_1.

Theorem 8.2. *There is a well-founded standard BD-set Γ such that the standard basis of $X(\Gamma)$ is asymptotic-ℓ_∞ but $X(\Gamma)$ does not contain c_0.*

The example above is of interest because it is relevant to problems about uniform homeomorphisms. It is not known whether a Banach space X that is uniformly homeomorphic to c_0 must be linearly homeomorphic (i.e. isomorphic) to c_0, but it is shown in Godefroy, Kalton, and Lancien [2001] that any such space X must be an isomorphic predual of ℓ_1 and have "summable Szlenk index"; Godefroy, Kalton and Lancien ask whether these properties already imply that X is isomorphic to c_0. The example of Argyros, Gasparis and Motakis shows that the answer is negative; it is not clear whether this space is uniformly homeomorphic to c_0. Very recently, this example has also found an application in descriptive set theory, in the proof by Kurka [2017] that the isomorphism class of c_0 is not Borel.

9 Self-determining subsets and BD augmentations

We have noted that the structure of a BD-set may be thought of in two (equivalent) ways, either in terms of the extension mappings $i_n : \ell_\infty(\Gamma_n) \to \ell_\infty(\Gamma)$ and the right-triangular basis $(d_\gamma)_{\gamma\in\Gamma}$ of $X(\Gamma)$ or in terms of the left-triangular basis functionals $d^*_\gamma = e^*_\gamma - c^*_\gamma$ in $\ell_1(\Gamma)$. This gives two different ways in which certain subsets of Γ may be naturally equipped with an induced BD structure. Introducing terminology that we shall use only temporarily, we shall say that a subset Γ' of Γ is *left-closed* if $\operatorname{supp} d^*_\gamma \subseteq \Gamma'$ whenever $\gamma \in \Gamma'$, and that a subset Γ'' is *right-closed* if $\operatorname{supp} d_\gamma \subseteq \Gamma''$ whenever $\gamma \in \Gamma''$.

When Γ' is left-closed, the functionals d^*_γ $(\gamma \in \Gamma')$ form a left triangular basis of $\ell_1(\Gamma')$ and so yield a BD-structure on Γ'; we then write $X(\Gamma')$ for the corresponding BD-space. On the other hand, when $\Gamma'' \subset \Gamma$ is right-closed, we have $i_n[\ell_\infty(\Gamma''_n)] \subset \ell_\infty(\Gamma'')$ for all n, so that Γ'' has its own BD-structure, defined by the extension operators $i''_n = i_n \restriction_{\ell_\infty(\Gamma''_n)}$. The connection between our two notions of closedness was established by Argyros and Motakis [2014].

Theorem 9.1. *Let Γ be a BD-set, let Γ' be a subset of Γ and let $\Gamma'' = \Gamma \setminus \Gamma'$. Then Γ' is left-closed if and only if Γ'' is right-closed. When this is the case, the restriction mapping $R' : x \mapsto x\restriction_{\Gamma'}$ is a quotient operator from $X(\Gamma)$ onto $X(\Gamma')$ and* $\ker R' = X(\Gamma'')$.

We note that in the set-up of Theorem 9.1, the BD-space $X(\Gamma'')$ is a subspace of $X(\Gamma)$, but that $X(\Gamma')$ typically is not. Indeed this happens only when Γ' is both left- and right-closed, and in this case $X(\Gamma)$ is just the direct sum of the disjointly supported spaces $X(\Gamma')$ and $X(\Gamma'')$. In general, $X(\Gamma)$ is a twisted sum of $X(\Gamma')$ and $X(\Gamma'')$, and the interesting cases are where this twisted sum is non-trivial.

In the terminology of Argyros and Motakis, left-closed subsets are called *self-determining* and we shall use that term in the rest of this paper. Once we have constructed a suitable "large" BD-set Γ then defining suitable self-determined subsets can be a useful and economical way of generating further examples with different properties. Thus a number of examples use self-determined subsets of a certain BD-set constructed by Argyros and Haydon [2011] and denoted $\Gamma^{\max}$ in that work: these include the scalar-plus-compact space of Argyros and Haydon [ibid.], the spaces constructed by Tarbard [2012, 2013], the main example of Argyros and Motakis [2016] and the recent construction due to Manoussakis, Pelczar-Barwacz, and Świętek [2017]. We look at the space of Argyros and Haydon [2011] in greater detail in a later section.

In order to prove refinements of Theorem 6.4, Freeman, Odell, and Schlumprecht [2011] introduced the tool of "augmenting" a BD-set, by adding extra elements to change chosen aspects of the structure of the associated BD-space. Let $\Gamma' = \bigcup_{n\in\mathbb{N}} \Delta'_n$ be a general BD-set; a BD-set $\Gamma = \Gamma' \cup \Gamma''$ that contains Γ' as a self-determining subset will be called an *augmentation* of Γ'. It is very easy to build augmentations as we have great freedom in choosing the functionals c^*_γ for the new elements $\gamma \in \Gamma''$. In particular, if Γ' is a standard BD-set of weight β, we can recursively build a standard augmentation, also of weight β, by adding new elements $\gamma \in \Delta''_{n+1}$ with complete freedom of choice of $\xi = \operatorname{base}\gamma \in \Gamma_n = \Gamma'_n \cup \Gamma''_n$, $s = \operatorname{cut}\gamma \le n+1$ and $\operatorname{top}\gamma \in \operatorname{ball}\ell_1(\Gamma_n \setminus \Gamma_{s-1})$. Augmentations will be used repeatedly in the rest of this article.

10 BD-sets with zero weight

A special case arises if we consider a BD-set such that, for each $\gamma \in \Gamma$, we have either $c^*_\gamma = 0$ or $c^*_\gamma = e^*_\xi$ for some ξ with $\operatorname{rank}\xi < \operatorname{rank}\gamma$. We may think of such a set as a standard BD-set with weight $\Gamma = \beta = 0$. It is not a surprise to find that the spaces $X(\Gamma)$ that arise in this way are actually objects with which we are already familiar.

Proposition 10.1. *Let Γ be a BD-set of zero weight. If Γ is well-founded, there is a locally compact topology on the countable set Γ such that $X(\Gamma) = \mathscr{C}_0(\Gamma)$. If Γ is ill-founded and B is the set of infinite branches of Γ then there is a locally compact topology on $\Gamma \cup \mathrm{B}$ such that $X(\Gamma)$ is naturally identifiable with $\mathscr{C}_0(\Gamma \cup \mathrm{B})$.*

The topologies in the above proposition are natural ones defined by the tree structure introduced in Section 6. In the well-founded case, the topology is the coarsest such that all the sets $U_\gamma = \{\delta \in \Gamma : \gamma \preccurlyeq \delta\}$ are open and closed; in the general case we take the sets V_γ ($\gamma \in \Gamma$) as basic clopen sets, where $V_\gamma = U_\gamma \cup \{\beta \in \mathrm{B} : \gamma \in \beta\}$. Two special cases are worth mentioning: if Γ is of weight zero and, moreover, no element of Γ has a base, then $c^*_\gamma = 0$ for all γ and $X(\Gamma)$ is just $c_0(\Gamma)$; if Γ is of weight 0 and $\operatorname{hist}\gamma \in \mathscr{S}$ for all $\gamma \in \Gamma$ then $X(\Gamma)$ is isometric to $\mathscr{C}_0(\alpha)$ for some ordinal $\alpha \le \omega^\omega$.

The technique of BD augmentations gives a neat way to construct certain examples of twisted sums, due originally to Cabello Sánchez, Castillo, Kalton, and Yost [2003]. The following is an example.

Theorem 10.2. *For every $\epsilon > 0$ there is a Banach space X and a subspace Y of X, isometric to $\mathscr{C}_0(\omega^\omega)$, such that X/Y is $(1+\epsilon)$-isomorphic to c_0 while the quotient mapping $X \to X/Y$ is strictly singular.*

To prove this, we start with the set $\Gamma' = \mathbb{N}$, equipped with the trivial BD-structure where $\operatorname{rank} n = n$ and $c^*_n = 0$ for all n; as we noted above, $X(\Gamma') = c_0$. We construct an augmentation of Γ' by adjoining new elements $\gamma \in \Gamma''$ as in the the construction of Γ^{BT} above, but subject to the condition that $\operatorname{supp}\operatorname{top}\gamma \subseteq \Gamma'$ for all γ (that is to say, $\eta \in \Gamma'$ in the notation of Definition 7.3). By Theorem 9.1 the restriction mapping $R' : x \mapsto x\restriction_{\Gamma'}$ is a quotient mapping from $X(\Gamma)$ onto $X(\Gamma') = c_0$ and the kernel of R' is the space $X(\Gamma'')$. But Γ'', considered as a BD-space in its own right, is of weight zero, because of the condition we placed on the tops. Because of the role of the Schreier family in the construction of Γ'', we see that $X(\Gamma'')$ is isometric to $\mathcal{C}_0(\omega^\omega)$. To prove the strict singularity of the quotient mapping R', we show that asymptotic ℓ_1 estimates of the kind used in Proposition 7.2 are valid for skipped block sequences (x_n) in $X(\Gamma)$ provided that $\inf_n \|R'x_n\| > 0$.

There is a natural way to associate with an arbitrary standard BD-set Γ, a BD-set $\mathring{\Gamma}$ of weight zero. We take $\mathring{\Gamma}$ to have the same underlying set as Γ, and retain the same definition of base γ, while redefining top γ to be 0 for all γ. While $X(\Gamma)$ and $X(\mathring{\Gamma})$ typically have very different Banach space structures, these two subspaces of $\ell_\infty(\Gamma)$ are quite close to each other in the Hausdorff metric.

Proposition 10.3. *Let Γ be a standard BD-space of weight β and let $\mathring{\Gamma}$ be the associated set of weight zero. For every $x \in X(\Gamma)$ (resp. $X(\mathring{\Gamma})$) and every $\epsilon > 0$, there exists $y \in X(\mathring{\Gamma})$ (resp. $X(\Gamma)$) with $\|x - y\| \le (2\beta M + \epsilon)\|x\|$, where $M = (1-2\beta)^{-1}$.*

Combining Proposition 10.3 with the embedding given by Theorem 6.4, we retrieve a result of Zippin [1977].

Theorem 10.4. *Zippin [ibid.] Let Y be a Banach space with separable dual and let ϵ be a positive real number. Then there is a countable locally compact space Γ and a subspace Z of $\ell_\infty(\Gamma)$ which is $(1+\epsilon)$-isomorphic to Y, such that for all $z \in Z$ there exists $h \in \mathcal{C}_0(\Gamma)$ with $\|z - h\| \le \epsilon\|z\|$.*

11 BD-sets with mixed weights

In order to introduce finer structure into our BD-sets and exploit the theory of mixed Tsirelson spaces, we make the following definition.

Definition 11.1. Let Γ be a countably infinite set and let $(m_i)_{i\in\mathbb{N}}$ be a sequence of natural numbers satisfying Definition 4.2(1). A weighted BD structure on Γ consists of the following mappings:

1. rank $: \Gamma \to \mathbb{N}$ such that each inverse image $\Delta_n = \operatorname{rank}^{-1}\{n\}$ is finite;
2. cut $: \Gamma \to \mathbb{N}$ satisfying $\operatorname{cut}\gamma \le \operatorname{rank}\gamma$;
3. top $: \Gamma \to \operatorname{ball}\ell_1(\Gamma)$ such that $\operatorname{supp}\operatorname{top}\gamma \subseteq \{\eta \in \Gamma : \operatorname{cut}\gamma \le \operatorname{rank}\eta < \operatorname{rank}\gamma\}$;
4. base $: \Gamma \to \Gamma \cup \{\text{undefined}\}$ such that $\operatorname{rank}\operatorname{base}\gamma < \operatorname{cut}\gamma$ whenever it is defined;
5. weight $: \Gamma \to \{m_i^{-1} : i \in \mathbb{N}\}$ such that $\operatorname{weight}\gamma = \operatorname{weight}\operatorname{base}\gamma$ when this is defined.

Given such a structure we set $\Gamma_n = \bigcup_{m\le n}\Delta_m = \{\gamma\in\Gamma : \operatorname{rank}\gamma \le n\}$ and define BD-functionals

$$c_\gamma^* = \qquad m_i^{-1}(I - P_{s-1}^*)b^*,$$

where $m_i^{-1} = \operatorname{weight}\gamma$, $s = \operatorname{cut}\gamma$, $b^* = \operatorname{top}\gamma$ and $\operatorname{base}\gamma$ is undefined, or

$$c_\gamma^* = e_\xi^* + m_i^{-1}(I - P_{s-1}^*)b^*,$$

with $\xi = \operatorname{base}\gamma$ when this is defined. We call Γ a *weighted BD-set*.

Provided Γ is well-founded, there are regular families $\mathcal{M}_i$ such that $\operatorname{hist}\gamma\in\mathcal{M}_i$ whenever γ is of weight m_i^{-1}. Our first task is to seek a class of sequences in $X(\Gamma)$ for which we can establish upper mixed-Tsirelson estimates. These will be the BD versions of the rapidly increasing sequences mentioned earlier in the context of mixed Tsirelson spaces. Although we define them in terms of evaluation estimates, rather than a particular method of construction, we shall shall continue to use the term RIS.

Definition 11.2. Let Γ be a well-founded BD-set with weights taking values in the sequence (m_i^{-1}). We shall say that a block sequence $(x_k)_{k\in\mathbb{N}}$ in $X(\Gamma)$ is a RIS if there exist a constant C and an increasing sequence (i_n) of natural numbers such that the following hold:

1. $\|x_k\| \le C$ for all n;
2. $|x_k(\gamma)| \le C m_h^{-1}$ whenever $\operatorname{weight}\gamma = m_h^{-1}$ and $h < i_k$;
3. $i_{k+1} > \max\{i : \exists\gamma \text{ with } \operatorname{rank}\gamma \le \max\operatorname{ran} x_k \text{ and } \operatorname{weight}\gamma = m_i^{-1}\}$.

We note that from a block sequence satisfying (1) and (2) we can always extract a subsequence satisfying (3) as well.

The next result is central to many subsequent developments. The proof is slightly intricate but is based on ideas that can be traced back to the splitting lemma, presented earlier as Lemma 3.7.

Theorem 11.3 (Second Basic Inequality)**.** *Let $(m_i)_{i\in\mathbb{N}}$ satisfy Definition 4.2(1), let $(\mathcal{M}_i)_{i\in\mathbb{N}}$ be sequence of regular families and let Γ be a weighted BD-set such that* $\operatorname{hist}\gamma\in\mathcal{M}_i$ *whenever* $\operatorname{weight}\gamma = m_i^{-1}$. *Let $(x_k)_{k\in\mathbb{N}}$ be a C-RIS in $X(\Gamma)$ with* $\max\operatorname{ran} x_k = q_k$. *If $I\subset\mathbb{N}$ is a finite interval, λ_k $(k\in I)$ are scalars and $\gamma\in\Gamma$ then there exist $k_0\in I$ and $g^*\in W[(\mathcal{A}_4 * \mathcal{M}_i, m_i^{-1})_{i\in\mathbb{N}}]$ such that either $g^* = 0$ or* $\operatorname{weight} g^* = \operatorname{weight}\gamma$ *and* $\operatorname{supp} g^* \subseteq \{q_k : k_0 < k \in I\}$, *and such that*

$$|\sum_{k\in I}\lambda_k x_k| \le 2C|\lambda_{k_0}| + 2C\langle g^*, \sum_{k_0<k\in I}\lambda_k e_{q_k}\rangle.$$

In particular, we have the upper estimate

$$\|\sum_k \lambda_k x_k\| \le 4C \|\sum_k \lambda_k e_{q_k}\|_{T[(\mathcal{M}_i * \mathcal{A}_4, m_i^{-1})_{i\in\mathbb{N}}]}.$$

The second crucial property of weighted BD-spaces, the one that will enable us to pass from "few operators" to "very few operators" in the next section, is the existence of two types of RIS that do not have analogues in the usual mixed Tsirelson framework. These are defined in terms of "local support". If $x \in X(\Gamma)$ has finite range $\operatorname{ran} x_k = [p, q]$ then we can write $x = i_q u$, where $u \in \ell_\infty(\Gamma_q)$ and $\operatorname{supp} u \subseteq \Gamma_q \setminus \Gamma_{p-1}$; we call $\operatorname{supp} u$ the *local support* of x. We say that a bounded block sequence (x_k) in $X(\Gamma)$ has *bounded local weight* if $\inf_k \min\{\operatorname{weight}\gamma : \gamma \in \operatorname{loc\,supp} x_k\} > 0$, and that (x_k) has rapidly decreasing local weight if $\max\{\operatorname{weight}\gamma : \gamma \in \operatorname{loc\,supp} x_k\}$ tends to 0 sufficiently fast. The interest of these definitions is in the following two propositions, the first of which is proved using the Evaluation Analysis 6.1.

Proposition 11.4. *Let Γ be a BD-set with weights in the sequence $(m_i^{-1})_{i\in\mathbb{N}}$ and let (x_k) be a bounded block sequence in $X(\Gamma)$. If (x_k) has either bounded local weight, or rapidly decreasing local weight, (x_k) is a RIS.*

Proposition 11.5. *Let Γ be a well-founded BD set, let Y be a Banach space and let $T : X(\Gamma) \to Y$ be a bounded linear operator. If $\|T(x_n)\| \to 0$ for every RIS in $X(\Gamma)$ then $\|T(x_n)\| \to 0$ for every bounded block sequence in $X(\Gamma)$; hence T is a compact operator.*

The proof of this second proposition is worth including, since it is easy and exploits the local ℓ_∞-structure of $X(\Gamma)$. We consider a bounded block sequence (x_k) with $\operatorname{ran} x_k = [p_k, q_k]$ and assume if possible that $\|Tx_k\| > \delta > 0$ for all k. We can write $x_k = i_{q_k} u_k$ where $\operatorname{supp} u_k \subseteq \Gamma_{q_k} \setminus \Gamma_{p_k - 1}$ and $\|u_k\| \le \|x_k\|$. For any $h \in \mathbb{N}$ we define $v_k^h \in \ell_\infty(\Gamma_{q_k})$ by setting

$$v_k^h(\gamma) = \begin{cases} u_k(\gamma) & \text{if weight}\,\gamma \ge m_h^{-1} \\ 0 & \text{otherwise} \end{cases}$$

For any h, the sequence (y_k^h) given by $y_k^h = i_{q_k} v_k^h$ is bounded ($\|y_k^h\| \le \|i_{q_k}\| \|v_k^h\| \le M\|x_k\|$) and has bounded local weight; hence it is a RIS and $\|Ty_k^h\| \to 0$ by hypothesis. This means that if we write $z_k^h = x_k^h - y_k^h = i_{q_k}(u_k - v_k^h)$, we must have $\|Tz_k^h\| > \frac{1}{2}\delta$ for all large enough k. It is now easy to construct sequences $h(r)$ and $k(r)$ tending to ∞ with r such that $\|Tz_{k(r)}^{h(r)}\| > \delta$ for all r. This contradicts our hypothesis, since a suitable subsequence of $(z_{k(r)}^{h(r)})$ has rapidly decreasing local weight and so is a RIS.

12 The scalar-plus-compact property and invariant subspaces

In Argyros and Haydon [2011], where $\mathcal{M}_i$ is taken to be $\mathcal{A}_{n_i}$, for an appropriately fast-growing sequence n_i, a large BD-set, denoted $\Gamma^{\max}$ was introduced, providing a framework in which other constructions can be made by taking self-determining subsets. The same can be done in general, using a recursive construction like that of Γ^{BT} given earlier.

Definition 12.1. We define $\Gamma^{\max}[(\mathcal{M}_i, m_i^{-1})_{i\in\mathbb{N}}]$ to be the union $\bigcup_{n\in\mathbb{N}} \Delta_n$, where the sets Δ_n and the functions hist and weight are given recursively by setting $\Delta_1 = \{1\}$, $\Gamma_n = \bigcup_{m\le n} \Delta_m$ and

$$
\begin{aligned}
\Delta_{n+1} = \{(n+1, i, 0, s, b^*) : i \le n+1,\ 1 \le s \le n,\ b^* \in B(s,n)\}\\
\cup \{(n+1, i, \xi, s, b^*) : 2 \le s \le n,\ \xi \in \Gamma_{s-1},\ \operatorname{weight}\xi = m_i^{-1},\ \{s\} \cup \operatorname{hist}\xi \in \mathcal{M}_i,\\
b^* \in B(s,n)\}, \operatorname{hist}(n+1, i, 0, s, b^*) = \{s\}, \quad \operatorname{hist}(n+1, i, \xi, s, b^*) = \{s\} \cup \operatorname{hist}\xi,\\
\operatorname{weight}(n+1, i, *, s, b^*) = m_i^{-1}.
\end{aligned}
$$

The differences with Definition 7.3 are first that we introduce i to allow for the mixed weights, and secondly that we allow the top b^* of an element to be a general element of ball $\ell_1(\Gamma_n \setminus \Gamma_{s-1})$, rather than restricting it to be an evaluation functional $\pm e_\eta^*$. For the sets Δ_{n+1} to be finite, we do need to place some restrictions on b^*, requiring it to lie in some finite ϵ-net $B(s,n)$ (which, of course, ought really to have been included in the recursive definition).

As we have said, the main role of $\Gamma^{\max}$ is to provide a framework for other constructions, but the $X(\Gamma^{\max})$, a mixed-Tsirelson version of the space $X(\Gamma^{\mathrm{BT}})$ is of some interest in its own right. It is natural to ask about its subspace structure, and whether there is an analogue of Proposition 7.2. It has been shown by Świętek that such an analogue does hold, at least for the version of $\Gamma^{\max}$ given in Argyros and Haydon [ibid.]; the reader is referred to Świętek [2017] for details.

In order to construct a space with the scalar-plus-compact property, we define a self determining subset Γ' of $\Gamma^{\max}$ by following the same recursive construction, while restricting the choice of the tops b^* of odd-weight elements. First we fix a *coding function* σ, an increasing injection from the set of odd-weight elements of $\Gamma^{\max}$ into $\mathbb{N}$. Then we insist that for an element $(n+1, 2j-1, 0, s, b^*)$ to be in Δ'_{n+1}, b^* must have the form e_η^* for some $\eta \in \Gamma'_n$ with weight of the form m_{4i-2}^{-1}, while for an element $(n+1, 2j-1, \xi, s, b^*)$ b^* must equal e_η^* for some $\eta \in \Gamma'$ of weight $m_{4\sigma(\xi)}$. We call the resulting BD-set $\Gamma^{\mathrm{K}}[[(\mathcal{M}_i, m_i^{-1})_{i\in\mathbb{N}}]$.

Theorem 12.2. *Provided the sequences (m_i) and $(\mathcal{M}_i)$ satisfy the Standard Assumptions, the space $X(\Gamma^{\mathrm{K}})$ has the scalar-plus-compact property.*

While the full strength of the Standard Assumptions was not needed in the previous section, we really need it here, in order to form special convex combinations and apply mixed-Tsirelson estimates such as Proposition 4.3. The proof, while complicated, is closely modeled on earlier proofs that certain spaces have few operators, eventually showing that for any bounded linear operator T on $X(\Gamma^K)$ there is a scalar λ such that $\|Tx_k - \lambda x_k\| \to 0$ for every RIS. The difference here is that we can now deduce compactness of $T - \lambda I$ by Proposition 11.5.

By the theorem of Aronszajn and Smith [1954], every bounded linear operator on $X(\Gamma^K)$ has a proper, non-trivial invariant subspace; $X(\Gamma^K)$ was the first infinite-dimensional space known to have this Invariant Subspace Property (though, of course, it is a famously open problem whether the more familiar space ℓ_2 does). A class of spaces with the Invariant Subspace Property but not the Scalar-plus-Compact Property was constructed by Tarbard [2012].

Theorem 12.3. *Tarbard [ibid.] For each natural number N there is a weighted BD-set Γ^N and a strictly singular operator S on $X(\Gamma^N)$ such that*

1. *S^N is non-compact*
2. *$S^{N+1} = 0$*
3. *every bounded linear operator T on $X(\Gamma^N)$ can be written uniquely as $T = \lambda I + \sum_{k=1}^{N} \lambda_k S^k + K$ with K compact.*

Tarbard's spaces have the Invariant Subspace Property, by Lomonosov's theorem, because any operator can be written $\lambda I + U$ with U^{N+1} compact.

No example is known of an infinite-dimensional reflexive space with the Scalar-plus-Compact Property (indeed, all known examples are $\mathcal{L}_\infty$-spaces). We must mention, however, the paper Argyros and Motakis [2014] which uses the method of saturation with constraints to construct an infinite-dimensional reflexive space all of whose subspaces have the Invariant Subspace Property.

The space $X(\Gamma^{\mathrm{K}})$ of Argyros and Haydon [2011] is hereditarily indecomposable and saturated with reflexive subspaces. Recent work has shown that the scalar-plus-compact property can hold in spaces with very different subspace structure: the space constructed by Manoussakis, Pelczar-Barwacz, and Świętek [2017] has the scalar-plus-compact property but is saturated with unconditional basic sequences; Argyros and Motakis [2016] combine the BD construction with the method of saturation with constraints to construct a space that has the scalar-plus-compact property, but no infinite-dimensional reflexive subspaces.

13 Calkin algebras

The Calkin algebra of a Banach space is the quotient $\mathcal{L}(X)/\mathcal{K}(X)$ where $\mathcal{L}(X)$ is the algebra of all bounded linear operators on X and $\mathcal{K}(X)$ the ideal of compact operators. Obviously the Calkin algebra of $X(\Gamma^K)$ is one-dimensional and that of Tarbard's space $X(\Gamma^N)$ is $N+1$-dimensional. It is natural to pose a question about the structure of Calkin algebras as Banach algebras:

Which unital Banach algebras arise as Calkin algebras of Banach spaces?

Tarbard [2013] gave a further example.

Theorem 13.1. *There is a well-founded BD-set Γ^∞ such that the Calkin algebra of $X(\Gamma^\infty)$ is isometrically isomorphic to the convolution algebra $\ell_1(\omega)$.*

Direct sums of versions of the example of Argyros and Haydon [2011] are used in Kania and Laustsen [2017] to show that every finite-dimensional semisimple algebra can be realised as a Calkin algebra. A major advance has come from Motakis, Puglisi, and Zisimopoulou [2016] who build on the interesting theory of BD-direct sums developed by Zisimopoulou [2014].

Theorem 13.2. *For every countable compact space K there is a $\mathcal{L}_\infty$ Banach space X with Calkin algebra isomorphic to $\mathcal{C}(K)$.*

Most recently, Motakis, Puglisi, and Tolias [2017] give a broad class of algebras of diagonal operators that may be realized as Calkin algebras, including examples that are hereditarily indecomposable and quasi-reflexive as Banach spaces. We are however still a long way from an answer to our question and we are still waiting for a first example of a unital Banach algebra that is *not* isomorphic to a Calkin algebra.

14 Indecomposable extensions of Banach spaces with separable duals

In this section and the next we shall be investigating the question of when a Banach space Y can be embedded as a subspace of a separable indecomposable space, or, more ambitiously, a separable space with the scalar-plus-compact property. The analogous question of when Y can be expressed as a quotient of a separable, hereditarily indecomposable space has been completely solved: it was shown by Argyros and Tolias [2004] that this is the case if and only if Y has no subspace isomorphic to ℓ_1. We conjecture that there is a similar best-possible for irreducible extensions.

Conjecture 14.1. For a Banach space Y with separable dual the following are equivalent:

1. Y embeds isomorphically into a separable space X with the scalar-plus-compact property;

2. Y embeds isomorphically into a separable indecomposable space X ;

3. Y has no subspace isomorphic to c_0.

It is clear that (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) because a c_0 subspace of a separable space is automatically complemented by Sobczyk's theorem. We note that the non-separable space $\mathcal{C}(K)$ constructed by Koszmider [2004] is indecomposable but contains $\mathcal{C}[0,1]$ (and hence copies of all separable spaces). Thus, to have a sensible conjecture we certainly need the word "separable" in (2). In passing we remark that we know of no non-separable Banach space with the scalar-plus-compact property, nor whether there is an upper bound on the size of such a space. While a hereditarily indecomposable space necessarily embeds in ℓ_∞ (Argyros and Tolias [2004]), it has recently been shown (subject to GCH) by Koszmider, Shelah, and Świętek [2018] that there exist arbitrarily large indecomposable Banach spaces. At the risk of drifting seriously off-topic, we wonder whether every Banach space not containing ℓ_∞ embeds in some indecomposable space.

While we cannot prove Conjecture 14.1, some partial results in this direction do exist. First, it was shown in Argyros, Freeman, Haydon, Odell, Raikoftsalis, Schlumprecht, and Zisimopoulou [2012] that every separable superreflexive Banach space Y embeds into a BD-space with the scalar-plus-compact property. The same authors can now do a little better, replacing the best-possible hypothesis that c_0 does not embed in Y with the same condition applied to the double dual Y^{**}.

Theorem 14.2. *Let Y be a Banach space such that Y^* is separable and c_0 does not embed into Y^{**}. Then Y embeds into a BD-space with the scalar-plus-compact property.*

The plan of the proof is to start by applying Theorem 6.4 to embed Y into $X(\Gamma')$ where Γ' is a well-founded standard BD-set of weight $m_1^{-1} = \frac{1}{4}$. Since Γ' is well-founded, there is a regular family $\mathcal{M}_1$ that contains all the histories $\operatorname{hist}\gamma$. We introduce $m_2, m_3, \ldots$ and $\mathcal{M}_2, \mathcal{M}_r, \ldots$ so that the Standard Assumptions are satisfied. The aim is then to build an augmentation $\Gamma = \Gamma' \cup \Gamma''$, such that $X(\Gamma)$ has the scalar-plus-compact property and contains the subspace Y. Even this last part is not as straightforward as it sounds, since, as we have noticed, $X(\Gamma')$ need not be a subspace of $X(\Gamma)$. For its subspace Y to be contained in $X(\Gamma)$ we shall need a rather special sort of augmentation.

Definition 14.3. Let Γ' be a standard BD-set, let Y be a subspace of $X(\Gamma')$ and let $\Gamma = \Gamma' \cup \Gamma''$ be a standard augmentation of Γ. Identify Y with the subspace

$$\{x \in \ell_\infty(\Gamma) : x\restriction_{\Gamma'} \in Y \text{ and } x\restriction_{\Gamma''} = 0\}.$$

of $\ell_\infty(\Gamma)$. We shall say that the augmentation *respects the subspace* Y if, for every $\gamma \in \Gamma''$, we have

$$P^*_{[s,\infty)} b^* \in Y^\perp \quad \text{and} \quad \xi \in \Gamma'' \text{ (if it exists)},$$

where, as usual $s = \operatorname{cut} \gamma$, $b^* = \operatorname{top} \gamma$ and $\xi = \operatorname{base} \gamma$.

Proposition 14.4. *If Γ is a standard augmentation of a standard BD-set Γ' that respects the subspace Y of $X(\Gamma)$ then Y is a subspace of $X(\Gamma)$.*

So our plan to prove Theorem 14.2 will be to construct an exotic augmentation of Γ' that respects the subspace Y. We shall need plenty of choice in selecting the tops of newly added $\gamma \in \Gamma''$ in order to achieve the scalar-plus-compact property, while ensuring that the condition $P^*_{[s,\infty)} b^* \in Y^\perp$ is satisfied. Theorem 6.4 does not necessarily have quite the property we want, so we may need to adjust it slightly, using a standard M-basis argument.

Proposition 14.5. *Let Γ' be a well-founded standard BD-set and let Y be a closed subspace of $X(\Gamma')$. Then there is small perturbation Z of Y such that $\mathbb{Q}\text{-sp}\langle d^*_\gamma : \gamma \in \Gamma'\rangle \cap Z^\perp$ is norm-dense in $\ell_1(\Gamma') \cap Z^\perp$.*

We assume that Y already has the property of Z in the above proposition and construct an augmentation that respects Y but is otherwise modeled on Theorem 12.2. We arrange that for every bounded linear operator $T : X(\Gamma) \to X(\Gamma)$ there is a scalar λ such that $Q(T - \lambda I)$ is compact, where $Q : X(\Gamma) \to X(\Gamma)/Y$ is the quotient operator. It is only now that we need the hypothesis on the original space Y in order to apply the following result, which we call the “Quotient-Compact Property”.

Proposition 14.6. *Let Z be a Banach space, let Y be a subspace of Z such that c_0 does not embed in Y^{**},let X be a Banach space with X^* isomorphic to ℓ_1 and let $S : X \to Z$ be a bounded linear operator. If QS is compact, where $Q : Z \to Z/Y$ is the quotient operator, then S is compact.*

Applying this proposition with $X = Z = X(\Gamma)$ and $S = T - \lambda I$ allows us to finish the proof of Theorem 14.2.

15 A space with the scalar-plus-compact property that contains ℓ_1

In this section, based on joint work of the authors with Th. Raikoftsalis, we sketch the construction of a Banach space that contains ℓ_1 and has the scalar-plus-compact property.

We start with a simple case of an ill-founded BD-set of zero weight, taking Γ^{d} to be the dyadic tree

$$\Gamma^{\mathrm{d}} = 2^{<\omega} = \bigcup_{n\in\omega} 2^n = \{(\,), (0), (1), (00), (01), \dots\}.$$

We use fairly standard notation for this tree. When $\gamma \in 2^n$ we write $n = \operatorname{length}\gamma$ (which is also the *domain* of γ if we are thinking of γ as a function) and for $\xi, \gamma \in 2^{<\omega}$ we write $\xi \prec \gamma$ if $\operatorname{length}\xi < \operatorname{length}\gamma$ and $\xi(i) = \gamma(i)$ for $0 \le i < \operatorname{length}\xi$, that is to say ξ is the restriction of γ to the domain of ξ. If $\gamma \in 2^{n+1}$ we write γ^- for the restriction of γ to n; so γ^- is the unique member of 2^n with $\gamma^- \prec \gamma$.

To endow Γ^{d} with a BD-structure, we define $\operatorname{rank}\gamma = 1 + \operatorname{length}\gamma$ and

$$\operatorname{base}\gamma = \gamma^-, \qquad \operatorname{top}\gamma = 0,$$

when $\operatorname{length}\gamma > 0$. Of course, the only reason for the term "+1" in the definition of rank is our earlier decision that the rank function on a BD-set should take values in $\mathbb{N}$, rather than ω. As in Section 10 we see that $X(\Gamma^{\mathrm{d}})$ may be identified naturally with the space $\mathcal{C}(2^{\le\omega})$, where $2^{\le\omega}$ is the compact metric space of all finite and infinite sequences in $\{0,1\}$. The space $\mathcal{C}(2^{\le\omega})$ has a subspace isomorphic to ℓ_1 and is a quotient of $X(\Gamma)$ whenever Γ is an augmentation of Γ^{d}. By the lifting property, $X(\Gamma)$ contains ℓ_1 for all such Γ.

Theorem 15.1. *There is an augmentation of Γ^{d} that has the scalar-plus-compact property.*

The idea is very simple: we take sequences (m_i) and $\mathcal{M}_i)$ satisfying the Standard Assumptions, and with $\mathcal{M}_1 = \mathcal{S}$, and then augment the BD-set $\Gamma' = \Gamma^{\mathrm{d}}$ by adding elements $\gamma \in \Gamma''$ as in the construction of Theorem 12.2. We do not have good upper norm estimates for sequences of vectors in $X(\Gamma)$, since Γ is not well-founded. But on the subspace $X(\Gamma'')$ the machinery of Theorem 12.2 can still be applied, leading to the following.

Proposition 15.2. *Let $\Gamma' = \Gamma^d$ and let $\Gamma = \Gamma' \cup \Gamma''$ be an augmentation as descibed above. For every bounded linear operator $U : X(\Gamma'') \to X(\Gamma)$ there is a scalar λ such that $U - \lambda J$ is compact, where $J : X(\Gamma'') \to X(\Gamma)$ is the inclusion operator.*

Once we have this proposition, we consider a bounded linear operator $T : X(\Gamma) \to X(\Gamma)$ and set $U = T\restriction_{X(\Gamma'')}$. There exists a scalar λ such that $U - \lambda J$ is compact; since $X(\Gamma)$ is a $\mathcal{L}_\infty$-space, there is a compact operator $S_1 : X(\Gamma) \to X(\Gamma)$ extending $U - \lambda J$ by Proposition 5.2. The operator $T - \lambda I - S_1$ vanishes on the subspace $X(\Gamma'')$, which is the kernel of the restriction quotient mapping $R' : X(\Gamma) \to X(\Gamma')$. So we can write $T = \lambda I + S_1 + S_2 R'$, where $S_2 : X(\Gamma') \to X(\Gamma)$ is some operator. Now $X(\Gamma')$ is a $\mathcal{C}(K)$-space, and $X(\Gamma)$ is skipped-asymptotic ℓ_1 (because of our choice $\mathcal{M}_1 = \mathcal{S}$). We finish the proof of Theorem 15.1 by using a general result about operators between such spaces.

Proposition 15.3. *If X is skipped-asymptotic ℓ_1 then every bounded linear operator from $\mathcal{C}(K)$ to X is compact.*

A more elaborate construction involving a Y-respecting augmentation of Γ^{d} leads to the analogue of Theorem 14.2 for spaces with nonseparable dual.

Theorem 15.4. *Let Y be a separable Banach space such that Y^{**} does not have a subspace isomorphic to c_0. Then Y may be embedded isomorphically into an indecomposable BD-$\mathscr{L}_\infty$ space $X(\Gamma)$. If every bounded linear operator from $\mathscr{C}[0,1]$ into Y is compact, then $X(\Gamma)$ may be chosen to have the scalar-plus-compact property.*

We have already noted in Proposition 10.9 of Argyros and Haydon [2011] that some extra condition is needed in Theorem 15.4 in order to get the scalar-plus-compact property, rather than just indecomposability. We can make this a little more precise with the following easy proposition, which seems to be a good note on which to end.

Proposition 15.5. *Let Y be a separable Banach space with non-separable dual. If Y embeds isomorphically in a separable $\mathscr{L}_\infty$-space with the scalar-plus-compact property then every bounded linear operator from $\mathscr{C}[0,1]$ to Y is compact.*

Proof. Assume that Y is a subspace of a $\mathscr{L}_\infty$-space X with the scalar-plus-compact property. Since Y^* is non-separable, so is X^*, which implies that X has a subspace isomorphic to ℓ_1 by a theorem of Lewis and Stegall [1973]; by a theorem of Pelczynski there is a quotient operator $Q : X \to \mathscr{C}[0,1]$. If $S : \mathscr{C}[0,1] \to Y$ is non-compact then so is $T = SQ : X \to Y \subset X$. On the other hand, Y cannot contain c_0 and so S is strictly singular. So T is strictly singular and non-compact contradicting the assumption about X. □

References

Dale E. Alspach (2000). "The dual of the Bourgain-Delbaen space". *Israel J. Math.* 117, pp. 239–259. MR: 1760594 (cit. on p. 1509).

Dale E. Alspach and Spiros A. Argyros (1992). "Complexity of weakly null sequences". *Dissertationes Math. (Rozprawy Mat.)* 321, p. 44. MR: 1191024 (cit. on pp. 1498, 1499).

Dale E. Alspach and Edward Odell (1988). "Averaging weakly null sequences". In: *Functional analysis (Austin, TX, 1986–87)*. Vol. 1332. Lecture Notes in Math. Springer, Berlin, pp. 126–144. MR: 967092 (cit. on p. 1498).

G. Androulakis and Thomas Schlumprecht (2001). "Strictly singular, non-compact operators exist on the space of Gowers and Maurey". *J. London Math. Soc. (2)* 64.3, pp. 655–674. MR: 1843416.

Spiros A. Argyros, Kevin Beanland, and Pavlos Motakis (2013). "Strictly singular operators in Tsirelson like spaces". *Illinois J. Math.* 57.4, pp. 1173–1217. MR: 3285871 (cit. on p. 1503).

Spiros A. Argyros and Irene Deliyanni (1992). "Banach spaces of the type of Tsirelson". arXiv: math/9207206 (cit. on p. 1500).

– (1997). "Examples of asymptotic l_1 Banach spaces". *Trans. Amer. Math. Soc.* 349.3, pp. 973–995. MR: 1390965 (cit. on pp. 1496, 1502, 1503, 1505).

Spiros A. Argyros and V. Felouzis (2000). "Interpolating hereditarily indecomposable Banach spaces". *J. Amer. Math. Soc.* 13.2, pp. 243–294. MR: 1750954 (cit. on p. 1496).

Spiros A. Argyros, Daniel Freeman, Richard G. Haydon, Edward Odell, Theocharis Raikoftsalis, Thomas Schlumprecht, and Despoina Z. Zisimopoulou (n.d.). *The Bourgain–Delbaen construction and indecomposable extensions of separable Banach spaces*. In preparation (cit. on p. 1496).

– (2012). "Embedding uniformly convex spaces into spaces with very few operators". *J. Funct. Anal.* 262.3, pp. 825–849. MR: 2863850 (cit. on p. 1523).

Spiros A. Argyros, Ioannis Gasparis, and Pavlos Motakis (2016). "On the structure of separable $\mathcal{L}_\infty$-spaces". *Mathematika* 62.3, pp. 685–700. MR: 3521348 (cit. on pp. 1495, 1496, 1508, 1514).

Spiros A. Argyros, Gilles Godefroy, and Haskell P. Rosenthal (2003). "Descriptive set theory and Banach spaces". In: *Handbook of the geometry of Banach spaces*. Ed. by W. B. Johnson and J. Lindenstrauss. Vol. 2. North-Holland, Amsterdam, pp. 1007–1069. MR: 1999190.

Spiros A. Argyros and Richard G. Haydon (2011). "A hereditarily indecomposable $\mathcal{L}_\infty$-space that solves the scalar-plus-compact problem". *Acta Math.* 206.1, pp. 1–54. MR: 2784662 (cit. on pp. 1495, 1496, 1513, 1515, 1520–1522, 1526).

Spiros A. Argyros and Pavlos Motakis (2014). "A reflexive hereditarily indecomposable space with the hereditary invariant subspace property". *Proc. Lond. Math. Soc. (3)* 108.6, pp. 1381–1416. MR: 3218313 (cit. on pp. 1503, 1515, 1521).

– (2016). "The scalar-plus-compact property in spaces without reflexive subspaces". arXiv: 1608.01962 (cit. on pp. 1515, 1521).

Spiros A. Argyros and Theocharis Raikoftsalis (2012). "The cofinal property of the reflexive indecomposable Banach spaces". *Ann. Inst. Fourier (Grenoble)* 62.1, pp. 1–45. MR: 2986263.

Spiros A. Argyros and Andreas Tolias (2004). "Methods in the theory of hereditarily indecomposable Banach spaces". *Mem. Amer. Math. Soc.* 170.806, pp. vi+114. MR: 2053392 (cit. on pp. 1496, 1522, 1523).

N. Aronszajn and K. T. Smith (1954). "Invariant subspaces of completely continuous operators". *Ann. of Math. (2)* 60, pp. 345–350. MR: 0065807 (cit. on p. 1521).

Kevin Beanland, Daniel Freeman, and Pavlos Motakis (2015). "The stabilized set of p's in Krivine's theorem can be disconnected". *Adv. Math.* 281, pp. 553–577. MR: 3366846 (cit. on p. 1503).

Steven F. Bellenot (1986). "Tsirelson superspaces and l_p". *J. Funct. Anal.* 69.2, pp. 207–228. MR: 865221 (cit. on p. 1500).

Jean Bourgain (1981). *New classes of* $\mathfrak{L}$sp$\langle p \rangle$*-spaces*. Vol. 889. Lecture Notes in Mathematics. Springer-Verlag, Berlin-New York, pp. v+143. MR: 639014.

Jean Bourgain and F. Delbaen (1980). "A class of special $\mathfrak{L}_\infty$ spaces". *Acta Math.* 145.3-4, pp. 155–176. MR: 590288 (cit. on pp. 1495, 1506, 1509).

F. Cabello Sánchez, J. M. F. Castillo, N. J. Kalton, and D. T. Yost (2003). "Twisted sums with $C(K)$ spaces". *Trans. Amer. Math. Soc.* 355.11, pp. 4523–4541. MR: 1990760 (cit. on pp. 1496, 1516).

Peter G. Casazza and Thaddeus J. Shura (1989). *Tsirelson's space*. Vol. 1363. Lecture Notes in Mathematics. With an appendix by J. Baker, O. Slotterbeck and R. Aron. Springer-Verlag, Berlin, pp. viii + 204. MR: 981801 (cit. on pp. 1499, 1502, 1504, 1513).

Dongyang Chen, William B. Johnson, and Bentuo Zheng (2011). "Commutators on $(\sum \ell_q)_p$". *Studia Math.* 206.2, pp. 175–190. MR: 2860306.

Per Enflo (1976). "On the invariant subspace problem in Banach spaces", p. 7. MR: 0473871.

– (1987). "On the invariant subspace problem for Banach spaces". *Acta Math.* 158.3-4, pp. 213–313. MR: 892591.

T. Figiel and W. B. Johnson (1974). "A uniformly convex Banach space which contains no l_p". *Compositio Math.* 29, pp. 179–190. MR: 0355537 (cit. on p. 1499).

Daniel Freeman, Edward Odell, and Thomas Schlumprecht (2011). "The universality of ℓ_1 as a dual space". *Math. Ann.* 351.1, pp. 149–186. MR: 2824850 (cit. on pp. 1495, 1496, 1511, 1515).

I. Gasparis (2003). "Strictly singular non-compact operators on hereditarily indecomposable Banach spaces". *Proc. Amer. Math. Soc.* 131.4, pp. 1181–1189. MR: 1948110.

I. Gasparis, M. K. Papadiamantis, and Despoina Z. Zisimopoulou (2010). "More ℓ_r saturated $\mathcal{L}^\infty$ spaces". *Serdica Math. J.* 36.2, pp. 149–170. MR: 2732550 (cit. on p. 1514).

G. Godefroy, N. J. Kalton, and G. Lancien (2001). "Szlenk indices and uniform homeomorphisms". *Trans. Amer. Math. Soc.* 353.10, pp. 3895–3918. MR: 1837213 (cit. on p. 1514).

William T. Gowers (1994a). "A Banach space not containing $c_0,\ l_1$ or a reflexive subspace". *Trans. Amer. Math. Soc.* 344.1, pp. 407–420. MR: 1250820 (cit. on pp. 1496, 1502).

– (1994b). "A solution to Banach's hyperplane problem". *Bull. London Math. Soc.* 26.6, pp. 523–530. MR: 1315601 (cit. on pp. 1496, 1502).

– (1996). "A solution to the Schroeder-Bernstein problem for Banach spaces". *Bulletin of the London Mathematical Society* 28.3, pp. 297–304 (cit. on pp. 1496, 1502).

William T. Gowers and Bernard Maurey (1993). "The unconditional basic sequence problem". *J. Amer. Math. Soc.* 6.4, pp. 851–874. MR: 1201238 (cit. on pp. 1496, 1502, 1505).

– (1997). "Banach spaces with small spaces of operators". *Math. Ann.* 307.4, pp. 543–568. MR: 1464131.

Sophie Grivaux and Maria Roginskaya (2014–2015). "Espaces de Banach possédant très peu d'operateurs, Séminaire Bourbaki 67ème année". *Séminaire Bourbaki*, p. 40 (cit. on p. 1496).

Richard G. Haydon (2000). "Subspaces of the Bourgain–Delbaen space". *Studia Math.* 139.3, pp. 275–293. MR: 1762585 (cit. on p. 1509).

– (2006). "Variants of the Bourgain–Delbaen construction". In: *Unpublished conference talk, Caceres* (cit. on p. 1512).

Robert C. James (1964). "Uniformly non-square Banach spaces". *Ann. of Math. (2)* 80, pp. 542–550. MR: 0173932 (cit. on pp. 1501, 1504).

Tomasz Kania and Niels Jakob Laustsen (2017). "Ideal structure of the algebra of bounded operators acting on a Banach space". *Indiana Univ. Math. J.* 66.3, pp. 1019–1043. MR: 3663335 (cit. on p. 1522).

Piotr Koszmider (2004). "Banach spaces of continuous functions with few operators". *Math. Ann.* 330.1, pp. 151–183. MR: 2091683 (cit. on p. 1523).

Piotr Koszmider, Saharon Shelah, and Michał Świętek (2018). "There is no bound on sizes of indecomposable Banach spaces". *Adv. Math.* 323, pp. 745–783. MR: 3725890 (cit. on p. 1523).

Ondřej Kurka (Nov. 2017). "The isomorphism class of c_0 is not Borel". arXiv: 1711.03328 (cit. on p. 1514).

D. R. Lewis and C. Stegall (1973). "Banach spaces whose duals are isomorphic to $l_1(\Gamma)$". *J. Functional Analysis* 12, pp. 177–187. MR: 0342987 (cit. on p. 1526).

J. Lindenstrauss and A. Pełczyński (1968). "Absolutely summing operators in L_p-spaces and their applications". *Studia Math.* 29, pp. 275–326. MR: 0231188 (cit. on p. 1505).

J. Lindenstrauss and H. P. Rosenthal (1969). "The $\mathcal{L}_p$ spaces". *Israel J. Math.* 7, pp. 325–349. MR: 0270119 (cit. on pp. 1505, 1506).

V. I. Lomonosov (1973). "Invariant subspaces of the family of operators that commute with a completely continuous operator". *Funkcional. Anal. i Priložen.* 7.3, pp. 55–56. MR: 0420305.

Antonis Manoussakis, Anna Pelczar-Barwacz, and Michał Świętek (2017). "An unconditionally saturated Banach space with the scalar-plus-compact property". *J. Funct. Anal.* 272.12, pp. 4944–4983. MR: 3639519 (cit. on pp. 1515, 1521).

Bernard Maurey (1994). "Quelques progrès dans la compréhension de la dimension infinie". In: *Espaces de Banach classiques et quantiques*. Vol. 1994. SMF Journ. Annu. Soc. Math. France, Paris, p. 29. MR: 1465453 (cit. on pp. 1496, 1505).

Bernard Maurey (2003). "Banach spaces with few operators". In: *Handbook of the geometry of Banach spaces, Vol. 2*. North-Holland, Amsterdam, pp. 1247–1297. MR: 1999196 (cit. on pp. 1496, 1505).

Bernard Maurey, V. D. Milman, and N. Tomczak-Jaegermann (1995). "Asymptotic infinite-dimensional theory of Banach spaces". In: *Geometric aspects of functional analysis (Israel, 1992–1994)*. Vol. 77. Oper. Theory Adv. Appl. Birkhäuser, Basel, pp. 149–175. MR: 1353458 (cit. on p. 1500).

Bernard Maurey and H. P. Rosenthal (1977). "Normalized weakly null sequence with no unconditional subsequence". *Studia Math.* 61.1, pp. 77–98. MR: 0438091 (cit. on p. 1503).

Pavlos Motakis, Daniele Puglisi, and Andreas Tolias (Nov. 2017). "Algebras of diagonal operators of the form scalar-plus-compact are Calkin algebras". arXiv: 1711.01340 (cit. on p. 1522).

Pavlos Motakis, Daniele Puglisi, and Despoina Z. Zisimopoulou (2016). "A hierarchy of Banach spaces with $C(K)$ Calkin algebras". *Indiana Univ. Math. J.* 65.1, pp. 39–67. MR: 3466455 (cit. on p. 1522).

N. J. Nielsen and P. Wojtaszczyk (1973). "A remark on bases in $\mathcal{L}_p$-spaces with an application to complementably universal $\mathcal{L}_\infty$-spaces". *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* 21, pp. 249–254. MR: 0322484 (cit. on p. 1505).

Edward Odell and Thomas Schlumprecht (1994). "The distortion problem". *Acta Math.* 173.2, pp. 259–281. MR: 1301394 (cit. on pp. 1496, 1502).

– (1995). "On the richness of the set of p's in Krivine's theorem". In: *Geometric aspects of functional analysis (Israel, 1992–1994)*. Vol. 77. Oper. Theory Adv. Appl. Birkhäuser, Basel, pp. 177–198. MR: 1353459 (cit. on p. 1503).

– (2000). "A Banach space block finitely universal for monotone bases". *Trans. Amer. Math. Soc.* 352.4, pp. 1859–1888. MR: 1637094 (cit. on p. 1503).

Edward Odell, Thomas Schlumprecht, and A. Zsák (2007). "Banach spaces of bounded Szlenk index". *Studia Math.* 183.1, pp. 63–97. MR: 2360257 (cit. on p. 1502).

A. Pełczyński (1968). "On Banach spaces containing $L_1(\mu)$". *Studia Math.* 30, pp. 231–246. MR: 0232195.

Thomas Schlumprecht (1991). "An arbitrarily distortable Banach space". *Israel J. Math.* 76.1-2, pp. 81–95. MR: 1177333 (cit. on pp. 1502, 1503).

Séminaire Choquet. 17e année: 1977/78. Initiation à l'analyse. Fasc. 1: Exposés 1 à 15 (1978). Secrétariat Mathématique, Paris, 109 pp. (not consecutively paged). MR: 522976.

Michał Świętek (Sept. 2017). "Regular subspaces of a Bourgain-Delbaen space $\mathfrak{B}_{mT}$". arXiv: 1709.06481 (cit. on p. 1520).

Matthew Tarbard (2012). "Hereditarily indecomposable, separable $\mathcal{L}_\infty$ Banach spaces with ℓ_1 dual having few but not very few operators". *J. Lond. Math. Soc. (2)* 85.3, pp. 737–764. MR: 2927806 (cit. on pp. 1515, 1521).

– (2013). *Operators on banach spaces of Bourgain-Delbaen type*. Thesis (D.Phil.)–University of Oxford (United Kingdom). ProQuest LLC, Ann Arbor, MI, (no paging). arXiv: 1309.7469. MR: 3271750 (cit. on pp. 1515, 1522).

Boris S. Tsirel'son (1974). "Not every Banach space contains ℓ_p or c_0". *Functional Analysis and its Applications* 8.2, pp. 138–141 (cit. on pp. 1496, 1499).

L. Tzafriri (1979). "On the type and cotype of Banach spaces". *Israel J. Math.* 32.1, pp. 32–38. MR: 531598 (cit. on p. 1503).

M. Zippin (1977). "The separable extension problem". *Israel J. Math.* 26.3-4, pp. 372–387. MR: 0442649 (cit. on pp. 1496, 1517).

– (2003). "Extension of bounded linear operators". In: *Handbook of the geometry of Banach spaces, Vol. 2*. North-Holland, Amsterdam, pp. 1703–1741. MR: 1999607 (cit. on p. 1506).

Despoina Z. Zisimopoulou (Feb. 2014). "Bourgain–Delbaen $\mathcal{L}^\infty$-sums of Banach spaces". arXiv: 1402.6564 (cit. on p. 1522).

Received 2017-12-18.

SPIROS A. ARGYROS
DEPARTMENT OF MATHEMATICS
NATIONAL TECHNICAL UNIVERSITY OF ATHENS
ATHENS 15780
GREECE
sargyros@math.ntua.gr

RICHARD G. HAYDON
MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
ANDREW WILES BUILDING
WOODSTOCK ROAD
OXFORD OX2 6GG
UK
richard.haydon@bnc.ox.ac.uk

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1529–1556)

HARMONIC MEASURE: ALGORITHMS AND APPLICATIONS

Christopher J. Bishop

Abstract

This is a brief survey of results related to planar harmonic measure, roughly from Makarov's results of the 1980's to recent applications involving 4-manifolds, dessins d'enfants and transcendental dynamics. It is non-chronological and rather selective, but I hope that it still illustrates various areas in analysis, topology and algebra that are influenced by harmonic measure, the computational questions that arise, the many open problems that remain, and how these questions bridge the gaps between pure/applied and discrete/continuous mathematics.

1 Conformal complexity and computational consequences

• **Three definitions:** First, the most intuitive definition of harmonic measure is as the boundary hitting distribution of Brownian motion. More precisely, suppose $\Omega \subset \mathbb{R}^n$ is a domain (open and connected) and $z \in \Omega$. We start a random particle at z and let it run until the first time it hits $\partial\Omega$. We will assume this happens almost surely; this is true for all bounded domains in $\mathbb{R}^n$ and many, but not all, unbounded domains. Then the first hit defines a probability measure on $\partial\Omega$. The measure of $E \subset \partial\Omega$ is usually denoted $\omega(z, E, \Omega)$ or $\omega_z(E)$. For E fixed, $\omega(z, E, \Omega)$ is a harmonic function of z on Ω, hence the name "harmonic measure".

Next, if Ω is regular for the Dirichlet problem, then, by definition, every $f \in C(\partial\Omega)$ has an extension $u_f \in C(\overline{\Omega})$ that is harmonic in Ω, and the map $z \to u_f(z)$, $z \in \Omega$ is a bounded linear functional on $C(\partial\Omega)$. By the Riesz representation theorem, $u_f(z) = \int_{\partial\Omega} f d\mu_z$, for some measure μ_z, and $\mu_z = \omega_z$. For domains with sufficient smoothness, Green's theorem implies harmonic measure is given by the normal derivative of Green's function times surface measure on the boundary. Thus the key to many results are estimates related to the gradient of Green's function.

The author is partially supported by NSF Grant DMS 16-08577.
MSC2010: primary 30C85; secondary 65D99, 30C30, 37F30.
Keywords: Harmonic measure, Brownian motion, conformal mapping, Hausdorff dimension, optimal meshing, conformal dynamics, dessins d'enfants, hyperbolic manifolds.

Finally, in the plane (but not in higher dimensions) Brownian motion is conformally invariant, so ω_z for a simply connected domain Ω is the image of normalized Lebesgue measure on the unit circle $\mathbb{T} = \{w : |w| = 1\}$ under a conformal map $f : \mathbb{D} = \{w : |w| < 1\} \to \Omega$ with $f(0) = z$. Because of the many tools from complex analysis, we generally have the best theorems and computational methods in this case.

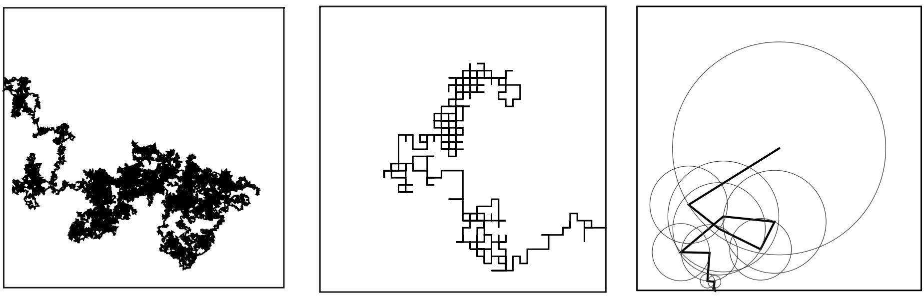


Figure 1: Continuous Brownian motion and two discrete approximations. In the center is a random walk on a grid; this is slow to use. On the right is the "walk-on-spheres" or "Kakutani's walk"; this is much faster to simulate.

• **The walk on spheres:** Suppose we want to compute the harmonic measure of one edge of a planar polygon. The most obvious approach is to approximate a Brownian motion by a random walk on a $\frac{1}{n} \times \frac{1}{n}$ grid. See Figure 1. However, it takes about n^2 steps for this walk to move distance 1, so for n large, it takes a long time for each particle to get near the boundary. A faster alternative is to note that Brownian motion is rotationally invariant, so it first hits a sphere centered on its starting point z in normalized Lebesgue measure. Fix $0 < \lambda < 1$ and randomly choose a point on

$$S_\lambda(z) = \{w : |w - z| = \lambda \cdot \mathrm{dist}(z, \partial\Omega)\}.$$

Now repeat. This random "walk-on-spheres" almost surely converges to a boundary point exponentially quickly, so only $O(\log n)$ steps are needed to get within $1/n$ of the boundary; see Binder and Braverman [2012]. I learned this process from a lecture of Shizuo Kakutani in 1986 and refer to it as Kakutani's walk.

However, even Kakutani's walk is only practical on small examples. Long corridors can make some edges very hard to reach, so we need a huge number of samples to estimate their harmonic measure. This is called the "crowding phenomena" (because the conformal pre-images of these edges are tiny; see below). For example, in a $1 \times r$ rectangle a Brownian path started at the center has only probability $\approx \exp(-\pi r/2)$ of hitting one of the short ends; for $r = 10$, the probability[1] is $\omega \approx 3.837587979 \times 10^{-7}$. See Figure 2. Thus random walks are not a time efficient method of computing harmonic measure (but they are memory efficient; see the work of Binder, Braverman, and Yampolsky [2007]).

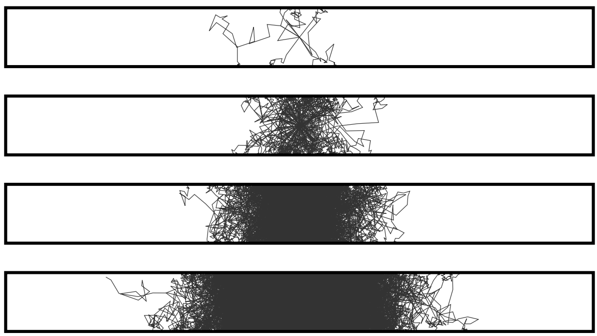


Figure 2: 10, 100, 1000 and 10000 samples of the Kakutani walk inside a 1×10 polygon. This illustrates the exponential difficulty of traversing narrow corridors.

• **The Schwarz-Christoffel formula:** Conformal mapping gives the best way of computing harmonic measure in a planar domain. See Figure 3. Many practical methods exist; surveys of various techniques include DeLillo [1994], Papamichael and Saff [1993], Trefethen [1986], Wegmann [2005]. Some fast and flexible current software includes `SCToolbox` by Toby Driscoll, `Zipper` by Don Marshall, and `CirclePack` by Ken Stephenson. To quote an anonymous referee of Bishop [2010a]: "Algorithmic conformal mapping is a small topic – one cannot pretend that thousands of people pay attention to it. What it does have going for it is durability. These problems have been around since 1869 and they have proved of lasting interest and importance."

When $\partial\Omega$ is a simple polygon, the conformal map $f : \mathbb{D} \to \Omega$ is given by the Schwarz-Christoffel formula (e.g., Christoffle [1867], Schwarz [1869], Schwarz [1890]):

$$f(z) = A + C \int_0^z \prod_{k=1}^{n} (1 - \frac{w}{z_k})^{\alpha_k - 1} dw,$$

where $\{\alpha_1\pi, \ldots, \alpha_n\pi\}$, are the interior angles of the polygon and $\mathbf{z} = \{z_1, \ldots, z_n\} \subset \mathbb{T} = \{z : |z| = 1\}$ are the preimages of the vertices (we call these the SC-parameters or the pre-vertices). For references, variations, and history of this formula, see Driscoll and Trefethen [2002].

[1] In fact, $\omega = \frac{2}{\pi} \arcsin((3 - 2\sqrt{2})^2(2 + \sqrt{5})^2(\sqrt{10} - 3)^2(5^{1/4} - \sqrt{2}^4)$; see page 262 of Bornemann, Laurie, Wagon, and Waldvogel [2004].

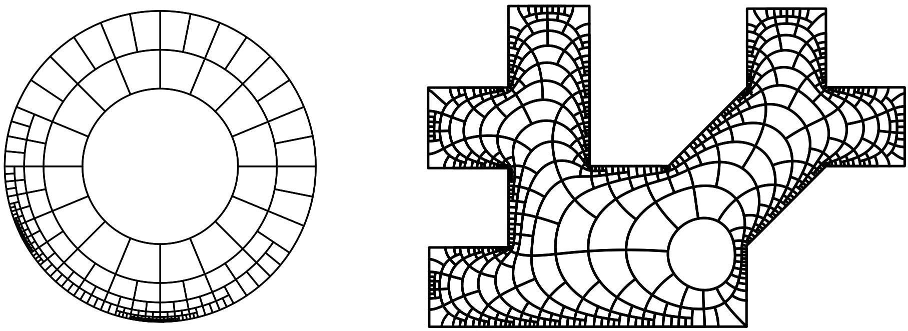


Figure 3: A conformal map to a polygon. The disk is meshed by boxes to a scale where vertex preimages are well separated. Counting boxes, we can estimate that the horizontal edge at top left has harmonic measure $\approx 2^{-16}$, another illustration of crowding.

The Schwarz-Christoffel formula does not really give us the conformal map; one must still solve for the n unknown SC-parameters, and this is a difficult problem. There are various heuristic methods that work as follows: make a parameter guess, compute the corresponding map, compare the image with the desired domain and modify the guess accordingly. Davis [1979] uses a simple side-length comparison: if a side is too long (or short), one simply decreases (or increases) the gap between the corresponding parameters proportionally. The more sophisticated CRDT algorithm of Driscoll and Vavasis [1998] uses cross ratios of adjacent Delaunay triangles to make the updated guess. However, neither Davis' method nor CRDT comes with a proof of convergence, nor a bound on how many steps are needed to achieve a desired accuracy.

• **The fast mapping theorem:** However, such bounds are possible (see Bishop [2010a]):

Theorem 1. *Given $\epsilon > 0$ and an n-gon P, there is $\mathbf{w} = \{w_1, \dots, w_n\} \subset \mathbb{T}$ so that*

1. *$\mathbf{w}$ can be computed in at most Cn steps, where $C = O(1 + \log \frac{1}{\epsilon} \log\log \frac{1}{\epsilon})$,*

2. *$d_{QC}(\mathbf{w}, \mathbf{z}) < \epsilon$ where $\mathbf{z}$ are the true SC-parameters.*

Here a step means an infinite precision arithmetic operation or function evaluation. The error in Theorem 1 is measured using a distance between n-tuples defined by

$$d_{QC}(\mathbf{w}, \mathbf{z}) = \inf\{\log K : \exists\ K\text{-quasiconformal } h : \overline{\mathbb{D}} \to \overline{\mathbb{D}} \text{ such that } h(\mathbf{z}) = \mathbf{w}\}.$$

A homeomorphism $h : \mathbb{D} \to \mathbb{D}$ is K-quasiconformal (K-QC) if it is absolutely continuous on almost all lines (so partial derivatives make sense a.e.) and $|\mu_h| \le k < 1$, where $\mu_h = h_{\overline{z}}/h_z$ is the complex dilatation of h (e.g., see Ahlfors [2006]). Geometrically, this says that infinitesimal circles are mapped to infinitesimal ellipses with eccentricity bounded by $K = (k+1)/(k-1) \geq 1$. In general, QC maps are non-smooth and can even map a line segment to fractal arc; see Bishop, Hakobyan, and Williams [2016] and its references.

The possible boundary values of a QC map $h : \mathbb{D} \to \mathbb{D}$ are exactly the quasisymmetric (QS) circle homeomorphisms. We say $h : \mathbb{T} \to \mathbb{T}$ is M-QS if $|h(I)| \leq M|h(J)|$ whenever I and J are disjoint, adjacent intervals of the same length on $\mathbb{T}$.

A map $f : \mathbb{D} \to \mathbb{D}$ is called a quasi-isometry (QI) for the hyperbolic metric ρ if there is an $A < \infty$ so that $A^{-1} \leq \rho(f(z), f(w))/\rho(z,w) \leq A$ whenever $\rho(z,w) \geq 1$; thus f is bi-Lipschitz at large scales, but we make no assumptions at small scales, not even continuity. Nevertheless, such an f does extend to a homeomorphism of the boundary circle, and the class of these extensions is again the QS-homeomorphisms. Thus QC and QI self-maps of $\mathbb{D}$ have the same set of boundary values.

Using the QC-metric on n-tuples has several advantages: it implies approximation in the Hausdorff metric and ensures points occur in the correct order on $\mathbb{T}$. When K is close to 1, the QS formulation holds with $M \approx 1$ and implies that the relative gaps between points are correct in a scale invariant way. We also have $d_{QC}(\mathbf{w}, \mathbf{z}) = 0$ iff the n-tuples are Möbius images of each other; this occurs iff the corresponding polygons are similar, which makes d_{QC} a natural metric for comparing shapes (to be precise, d_{QC} is only a metric if we consider n-tuples modulo Möbius transformations). Finally, this metric is easy to bound by computing any vertex-preserving QC map between the corresponding polygons, e.g., the obvious piecewise linear map coming from two compatible triangulations. See Figure 4. Using this, we can bound the QC-distance to the true SC-parameters without knowing what those parameters are. Computing the exact QC-distance between n-tuples is much harder, e.g. see Goswami, Gu, Pingali, and Telang [2015].

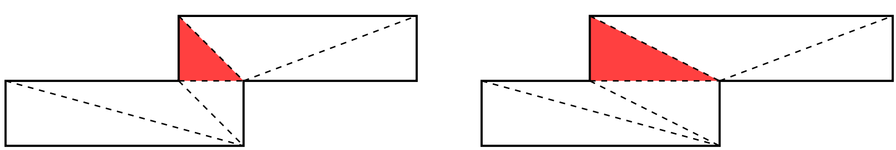


Figure 4: Equivalent triangulations of two polygons define a piecewise linear QC map and give an upper bound for the QC distance. Here $K = 2$ and the most distorted triangle is shaded.

• **Applications to computational geometry:** We will first discuss some applications of the fast mapping theorem (FMT), and then discuss its proof. As explained below, the proof of the FMT depends on ideas from computational geometry (CG), and it returns the favor by solving certain problems in CG. Optimal meshing is the problem of efficiently decomposing a domain Ω into nice pieces. Assume $\partial\Omega$ is an n-gon. "Efficient" means we want the number of mesh elements to be bounded by a polynomial in n (independent of Ω). "Nice" means the pieces are triangles or quadrilaterals that have angles strictly bounded between 0° and 180°, whenever possible. Some results that use the FMT (or ideas from its proof) include:

► **Thick/thin decomposition:** Every polygon can be written as a union of disjoint thick and thin pieces that are analogous to the thick/thin pieces of a hyperbolic manifold (regions where the injectivity radius is larger/smaller than some ϵ). See Figure 5. For an n-gon, each thin piece is either a neighborhood of a vertex (parabolic thin parts), or corresponds to a pair of sides that have small extremal distance within Ω (hyperbolic thin parts); the thin parts are in 1-to-1 correspondence with the thin parts of the n-punctured Riemann sphere formed by gluing two copies of the polygon along its (open) edges. Despite there being $\simeq n^2$ pairs of edges, there are $O(n)$ thin parts, and they can be found in time $O(n)$ using the FMT with $\epsilon \simeq 1$; see Bishop [2010a].

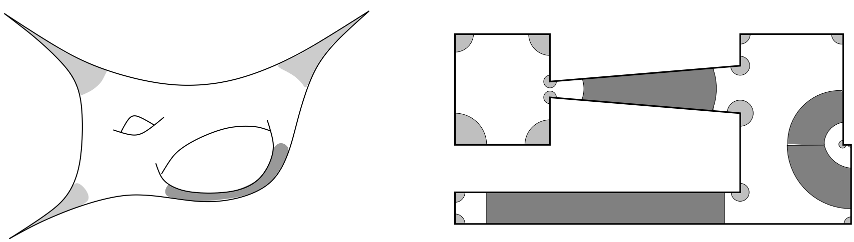


Figure 5: Thin parts of a surface and a polygon are shaded (light = parabolic, dark = hyperbolic), and the thick pieces are white.

► **Optimal quad-meshing:** Any n-gon has an $O(n)$ quadrilateral mesh where every angle is less than 120° and all the new angles are at least 60°; see Bishop [2010b], Bishop [2016b] Here "new" means that existing angles $< 60°$ remain, but are not subdivided. Both the complexity and angle bounds are sharp. The thick/thin decomposition plays a major role here: the thin parts are meshed with an ad hoc Euclidean construction and the thick parts are meshed by transferring a hyperbolic mesh from $\mathbb{D}$ by a nearly conformal map. Is there a similar approach in 3 dimensions, perhaps using decompositions into pieces that are meshed using some of the eight natural 3-dimensional geometries?

► **The NOT theorem:** Every planar triangulation with n elements can be refined to a **no**n**o**btuse **t**riangulation (all angles $\leq \theta = 90°$, called a NOT for brevity) with $O(n^{2.5})$ triangles; see Bishop [2016a]. No polynomial bound is possible if $\theta < 90°$ and the previous best result was with $\theta = 132°$, due to Tan [1996]. See also Mitchell [1993]. A gap remains between the $O(n^{2.5})$ algorithm and the n^2 worst known example. The proof of the NOT theorem involves perturbing a natural C^1 flow associated to the triangulation, in order to cause collisions between certain flow lines. Is there any connection to closing lemmas in dynamics, e.g., Pugh [1967]? Perhaps the gap could be reduced using dynamical ideas, or ideas from the NOT theorem applied to flows on surfaces.

The NOT theorem has an amusing consequence: suppose several adjoining countries have polygonal boundaries (with n edges in total) and the governments all want to place cell towers so that a cell phone always connects to a tower (the closest one) in the same country as the phone. Is this possible using a polynomial number of towers? More mathematically, we are asking for a finite point set S whose Voronoi cells conform to the given boundaries (the Voronoi cells of S are the points closest to each element of S). If the countries are nonobtuse triangles this is easy to do, so the NOT theorem implies this is possible in general using $O(n^{2.5})$ points, the first polynomial bound for this problem stated in Salzberg, Delcher, Heath, and Kasif [1995].

• **Proof of the FMT:** Like the other methods mentioned earlier, the fast mapping algorithm iteratively improves an initial guess for the conformal map. However, whereas Davis' method and CRDT use conformal maps onto an approximate domain, and try to improve the domain, the fast mapping algorithm uses approximately conformal maps onto the correct target domain and improves the degree of conformality. More precisely, each iteration computes the dilatation μ_f of a QC map $f : \mathbb{D} \to \Omega$, and attempts to solve the Beltrami equation $g_{\overline{z}} = \mu_f g_z$ with a homeomorphism $g : \mathbb{D} \to \mathbb{D}$. If g was an exact solution, then $F = f \circ g^{-1}$ would be the desired conformal map. The exact solution is given by a infinite series involving the Beurling transform (see e.g., Ahlfors [2006]) but the FMT uses only the leading term of this series and approximately solves the resulting linear equation (thus it is a higher dimensional version of Newton's method). Iterating gives a sequence of QC maps that converge quadratically to a conformal map, assuming the initial dilatation μ is small enough. A variation of this method was implemented by Green [2011].

To bound the total time, we have to estimate the time needed for each iteration, and the time needed to find a starting guess for which we can uniformly bound the number of iterations needed to reach accuracy ϵ (it is not obvious that such a point even exists). The first step involves representing the map as a collection of series expansions on the disk, and applying discretized integral operators using the fast multipole method and structured linear algebra. The second part is less standard: we use computational geometry to make a "rough-but-fast" QC approximation to the Riemann map and use 3-dimensional hyperbolic geometry to prove that this guess is close to the correct answer, with a dilatation bound independent of the domain. It is (fairly) easy to reduce from "bounded dilatation" to "small dilatation" by a continuation argument, so we will only discuss proving the uniform bound.

2 Disks, domes, dogbones, dimension and dendrites

• **The medial axis flow:** The medial axis (MA) of a planar domain Ω is the set of all interior points that have $\geq$ 2 distinct closest points on $\partial\Omega$. For polygons, these are the centers of maximal disks in Ω, but the latter set can be strictly larger in general; see Bishop and Hakobyan [2008]. If $\partial\Omega$ is a polygon, then the medial axis is a finite tree. See Figure 6.

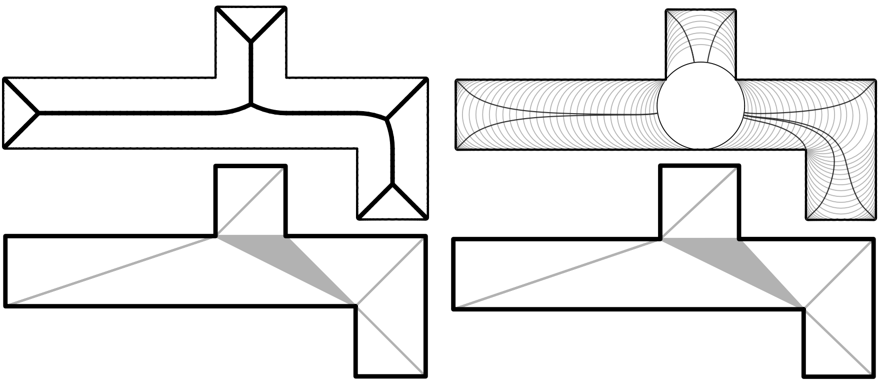


Figure 6: The top shows the medial axis of a domain (left) and the medial axis foliation and flow (right). The bottom show triangulations of the target polygon and initial guess using the MA-flow parameters. Here $K = 1.24$, (the most distorted triangle is shaded), but the polygons appear almost identical.

If we fix one medial axis disk D_0 as the "root" of this tree, then arcs of the remaining disks foliate $\Omega \setminus D_0$. Each boundary point can be connected to D_0 by a path orthogonal to this foliation; see Figure 6. The medial axis flow defines Möbius transformations between medial axis disks, hence preserves certain cross ratios, and given the medial axis, we can use this to compute the images of all n boundary vertices in $O(n)$ time. The medial axis itself can be computed in linear time by a result of Chin, Snoeyink, and Wang [1999], so the MA-flow gives a linear time (i.e., "fast") initial guess for the SC-parameters.

• **The convex hull theorem:** Why is our "fast guess" an accurate guess? The answer is best understood by moving from 2 to 3 dimensions. The "dome" of a planar domain Ω is the surface $S = S(\Omega) \subset \mathbb{H}^3 = \mathbb{R}^3_+ = \{(x, y, t) : t > 0\}$ that is the boundary of the union of all hemispheres whose base disk is contained in Ω. In fact, it suffices to consider only medial axis base disks. See Figure 7.

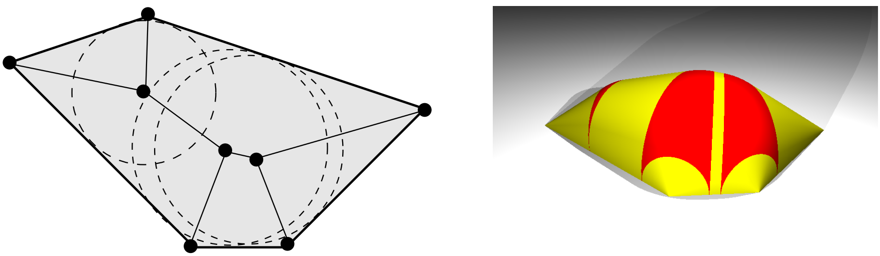


Figure 7: A polygonal domain and its dome. The red patches on the dome each correspond to the dome of a vertex disk of the medial axis; the yellow regions correspond to domes of edge disks.

Recall that $\mathbb{H}^3$ has a hyperbolic metric $d\rho = ds/t$. Each hemisphere below the dome S is a hyperbolic half-space, and the region above S is the intersection of their complements, hence is hyperbolically convex. Thus the dome of Ω is also the boundary of the hyperbolic convex hull in $\mathbb{H}^3$ of $\Omega^c = \mathbb{C} \setminus \Omega$. We define the "nearest point retraction" $R : \Omega \to S(\Omega)$ by expanding a horo-sphere in $\mathbb{R}^3_+$ tangent to $\mathbb{R}^2$ at $z \in \Omega$ until it first hits S at a point $R(z)$. See Figure 8. Dennis Sullivan's convex hull theorem (CHT) states that R is a quasi-isometry from the hyperbolic metric on Ω to the hyperbolic path metric on the dome. Sullivan [1981b] originally proved the CHT in the context of hyperbolic 3-manifolds (see below) and the version above is due to Epstein and Marden [1987]. See also Bishop [2001], Bishop [2002], Bridgeman and Canary [2010].

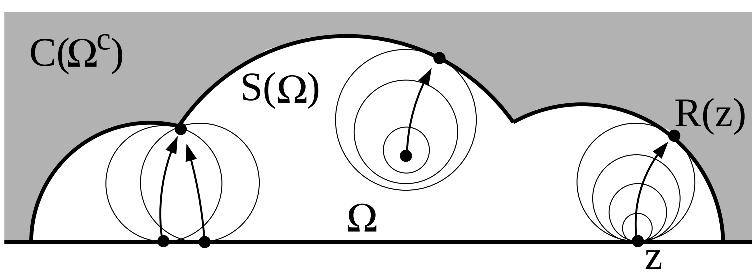


Figure 8: The dome S of Ω is the boundary of the hyperbolic convex hull of Ω^c (shaded). The retraction map $R : \Omega \to S$ defined by expanding horoballs need not be 1-to-1, but is a quasi-isometry.

The dome S with its hyperbolic path metric is isometric to the hyperbolic disk. The isometry $\iota : S \to \mathbb{D}$ can be visualized by thinking of S as bent along a disjoint collection of geodesics, and "flattening" the bends until we get a hyperbolic plane (the hemisphere above D_0; this is clearly isomorphic to $\mathbb{D}$). Remarkably, the restriction of this map to

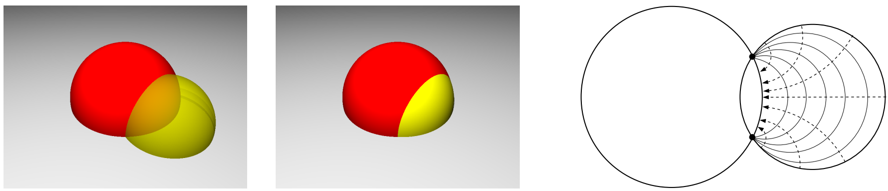


Figure 9: The dome of two overlapping disks consists of two hyperbolic half-planes joined along a geodesic (left). Flattening this bend means rotating one half-plane around the geodesic until it is flush with the other (center). On $\mathbb{R}^2$, this rotation corresponds to the medial axis flow in the base domain. The same observation applies to all finite unions of disks, and the general case follows by a limiting argument.

$\partial S = \partial\Omega$ equals the MA-flow map $\partial\Omega \to \partial D_0$. Figure 9 gives the idea of the proof. Since $\iota \circ R : \Omega \to \mathbb{D}$ is a quasi-isometry (and because QI and QC maps of $\mathbb{D}$ have the same boundary values), the MA-flow map $\partial\Omega \to \partial D_0$ has a uniformly QC extension $\sigma : \Omega \to D_0$. Thus our "fast guess" is indeed a "good guess" with uniform bounds.

• **Convex hulls and 3-manifolds:** As mentioned above, Sullivan's CHT was first discovered in the context of hyperbolic 3-manifolds. By definition, such a manifold M is the quotient of $\mathbb{H}^3$ by a Kleinian group, i.e., a discrete group G of orientation preserving hyperbolic isometries. This is completely analogous to a Riemann surface being the quotient of the hyperbolic disk by a Fuchsian group. The accumulation set of any G-orbit on $\partial\mathbb{H}^3 = \mathbb{R}^2 \cup \{\infty\}$ is called the limit set Λ of G; this is often a fractal set. The complement of Λ is called the ordinary set Ω. In this paper we will always assume $\Omega \neq \varnothing$. We let $C(\Lambda) \subset \mathbb{H}^3$ denote the hyperbolic convex hull of Λ. It is G-invariant, so its quotient defines a region $C(M) \subset M$ called the convex core of M; this is also the convex hull of all the closed geodesics in M. We define the "boundary at infinity" of M as $\partial_\infty M = \Omega/G$; this is a union of Riemann surfaces, one for each connected component of Ω. The dome of each component of Ω is a boundary component of $C(\Lambda)$, and corresponds to a boundary component of $C(M)$. The original formulation of Sullivan's CHT (which he attributes to Thurston) is that $\partial_\infty M$ is uniformly QC-equivalent to $\partial C(M)$.

A case of particular interest is when M is homeomorphic to $\Sigma \times \mathbb{R}$ for some compact surface Σ and $C(M)$ is compact (this is called a co-compact quasi-Fuchsian manifold). See Figure 10. Then Λ is a Jordan curve, so $\partial C(M)$ has two components, Ω_1

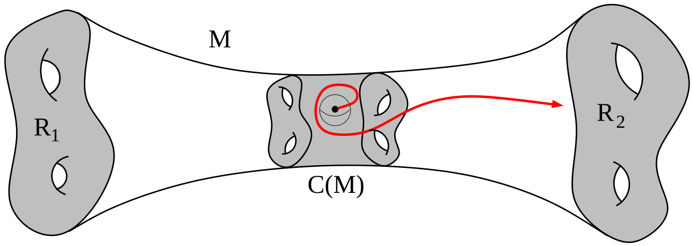


Figure 10: A co-compact quasi-Fuchsian manifold. The tunnel vision function is the harmonic measure of one component of $\partial_\infty M$.

and Ω_2. Since $u = \omega(z, \Omega_2, \mathbb{H}^3)$ is invariant under G, it defines a harmonic function $u(z) = \omega(z, R_2, M)$ on M. (Here u is harmonic for the hyperbolic metric on $\mathbb{H}^3$, not the Euclidean metric; the two concepts agree in 2 dimensions, but not in 3.) This is the "tunnel vision" function: for $z \in M$, $u(z)$ is the normalized area measure (on the tangent 2-sphere) of the geodesic rays starting at z that tend towards $R_2 \subset \partial_\infty M$. Thus u is the "brightness" at z if R_2 is illuminated but R_1 is dark. It is easy to check that $u \geq 1/2$ on the component of $\partial C(M)$ facing R_2 and is $\leq 1/2$ on the other component. Thus the level set $\{z : u(z) = \frac{1}{2}\}$ is contained inside $C(M)$.

• **Dogbones and 4-manifolds:** The topology of the tunnel vision level sets has an interesting connection to 4-dimensional geometry. If Λ is a circle, then the level sets $\{u(z) = \lambda\}$, $0 < \lambda < 1$, are topological disks, but if Λ approximates $\partial\Omega$, where

$$\Omega = \{z : |z-1| < 1/2\} \cup \{z : |z+1| < 1/2\} \cup \{z = x + iy : |x| < 1, |y| < \epsilon\},$$

and ϵ is small, then they can be non-trivial and u has a critical point. See Figure 11.

This critical point has a surprising consequence. Claude LeBrun has shown how to turn the hyperbolic 3-manifold M into a closed anti-self-dual 4-manifold N, so that N has an almost-Kähler structure if and only if u has no critical points. For definitions and details, see Bishop and LeBrun [2017]. The simplest case is to take $M \times \mathbb{T}$ and collapse $\partial_\infty M$ to two points; this gives a conformally flat N, but a hierarchy of topologically distinct non-flat examples also exists. In Bishop and LeBrun [ibid.] we construct a co-compact Fuchsian group that can be deformed to a quasi-Fuchsian group with limit set approximating the dogbone curve. Thus the almost-Kähler metrics sweep out an open, non-empty, but proper subset of the moduli space of anti-self-dual metrics on the corresponding 4-manifold N, giving the first example of this phenomena. Thus harmonic measure solves a problem about 4–manifolds, and 4–manifolds raise new questions about harmonic measure: for which planar domains Ω does $\omega(z, \Omega, \mathbb{H}^3)$ have a critical point? The group in Bishop and LeBrun [ibid.] has a huge number of generators; how many are really needed

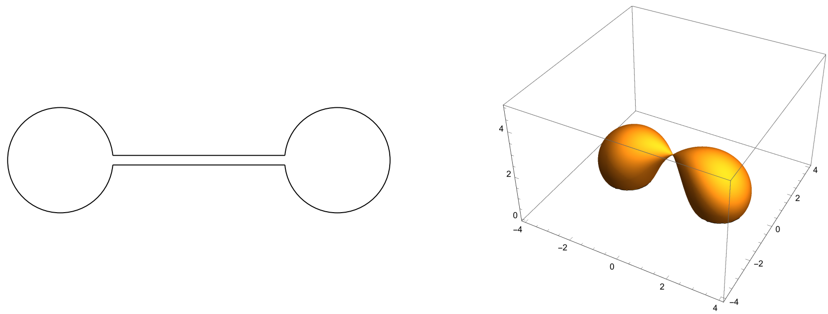


Figure 11: The dogbone domain (left) approximates two disjoint disks if the corridor is very thin. For two disks, the level surfaces $\{u(z) = \lambda\}$ evolve from two separate surfaces into a connected surface, so u must have a critical point; the critical surface is shown at right.

to get an example with a critical point? Are critical points common near the boundary of Teichmüller space for any large G?

• **Heat kernels and Hausdorff dimension:** As above, suppose $M \simeq \Sigma \times \mathbb{R}$ is hyperbolic and $C(M)$ is compact. By compactness, a Brownian motion inside $C(M)$ hits $\partial C(M)$ almost surely; as noted earlier, it then has probability $\geq 1/2$ of hitting the corresponding component of $\partial_\infty M$. This implies Brownian motion on M leaves $C(M)$ almost surely, which implies Brownian motion on $\mathbb{H}^3$ leaves $C(\Lambda)$ almost surely, which is equivalent to $\text{area}(\Lambda) = 0$. This observation can be made much more precise.

The heat kernel, $k_M(x, y, t)$, on a manifold M gives the probability that a Brownian motion starting at x at time 0 will be at y at time t. Thus the probability of being in $C(M)$ at time t is $p(x,t) = \int_{C(M)} k_M(x, y, t) dy$. The heat kernel can be written in terms of the eigenvalues and eigenfunctions of the Laplacian on M, $k_M(x, y, t) = \sum_{n=0}^{\infty} e^{-\lambda_n t} \varphi_n(x) \varphi_n(y)$, so it seems reasonable that $p(x,t) = O(\exp(-\lambda_0 t))$. See Davies [1988], Grigor'yan [1995], which make this precise. The lift of k_M to $\mathbb{H}^3$ is a sum over G-orbits of

$$k_{\mathbb{H}^3}(w, z, t) = (4\pi t)^{-3/2} \frac{\rho(z, w)}{\sinh(\rho(z, w))} \exp(-t - \frac{\rho(z, w)^2}{4t}).$$

Let $G_n = \{g \in G : n < \rho(0, g(0)) \leq n + 1\}$ and $N_n = \# G_n$. The critical exponent $\delta = \limsup_k (\log N_k)/k$, is always a lower bound for $\dim(\Lambda)$, and equality holds in many cases, e.g., when G is finitely generated. See Bishop and P. W. Jones [1997a], Sullivan [1984].

Putting these estimates together (and dropping the non-exponential terms) gives

$$e^{-\lambda_0 t} \simeq k_M(x,x,t) \simeq \sum_n \sum_{g \in G_n} k^3_{\mathbb{H}}(0, g(0), t) \simeq e^{-t} \sum_n e^{-(1-\delta)n - n^2/4t}.$$

The final sum is dominated by the term $n = -2t(1-\delta)$, and comparing the exponents gives $\lambda_0 = \delta(2-\delta)$, a well known formula relating the geometry of Λ to Brownian motion on M. Are other relations possible? If $C(M)$ is non-compact, but has finite volume, Sullivan [ibid.] showed the limit set has finite, positive packing measure (instead of Hausdorff measure, as happens when $C(M)$ is compact). Is this reflected by some property of Brownian motion or harmonic measure on M?

When $\mathrm{vol}(C(M)) = \infty$, Peter Jones and I proved that either (1) $\lambda_0 = 0$ and $\dim(\Lambda) = \delta = 2$ or (2) $\lambda_0 > 0$ and $\mathrm{area}(\Lambda) > 0$. See Bishop and P. W. Jones [1997a]. Again, this reduces to harmonic measure estimates: bounding $\omega(z, \partial C(M), M)$ at points deep inside $C(M)$. Both cases above can occur in general, but the second case ($\mathrm{area}(\Lambda) > 0$) is impossible for finitely generated groups with $\Omega \neq \varnothing$; this is the Ahlfors measure conjecture and was proven independently by Agol [2004] and by Calegari and Gabai [2006].

• **Dimension of dendrites:** We can strengthen the Ahlfors conjecture in some cases. Consider a singly degenerate manifold $M \simeq \Sigma \times \mathbb{R}$ where $C(M)$ contains one end of M, and also assume that M has positive injectivity radius (i.e., non-trivial loops have length bounded away from zero). See Figure 12. Then the limit set Λ is a dendrite (connected and does not separate the plane) of dimension 2 and area zero. Such limit sets are notoriously difficult to understand and compute.

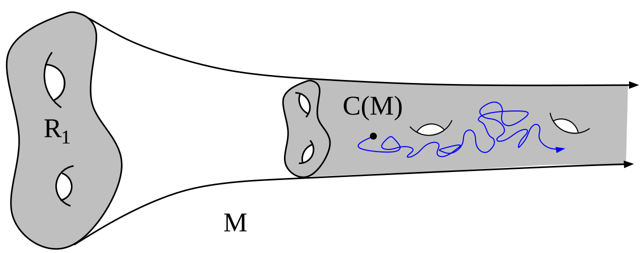


Figure 12: Co-compact quasi-Fuchsian manifolds can limit on a singly degenerate M: $C(M)$ contains a geometrically infinite end of M, and its complement is a geometrically finite end.

In this case, the tunnel vision function is constant, but there is an interesting alternative. By pushing the pole of Green's function G to ∞ through the geometrically infinite end, normalizing at a fixed point, and using estimates of $|\nabla G|$ in terms of the injectivity radius, one can show there is a positive harmonic function u on M that is zero on $R_1 \subset \partial_\infty M$, and grows linearly in the geometrically infinite end, i.e., $u(z) \simeq 1 + \mathrm{dist}(z, \partial C(M))$ for $z \in C(M)$. See Bishop and P. W. Jones [1997b]. Note that u lifts to a positive harmonic U on $\mathbb{H}^3$, and U must be the Poisson integral of a measure μ supported on Λ.

We expect Brownian motion, B_t, on the geometrically infinite end of M to behave like a Brownian path in $[0, \infty)$. By the law of the iterated logarithm (LIL), we then expect $u(B_n)$ to be as large as $\sqrt{n \log\log n}$ infinitely often (i.o.). Since a Brownian path on $\mathbb{H}^3$ tends to the boundary at linear speed in the hyperbolic metric, this means that at μ-a.e. $z \in \Lambda$, i.o. we have $U((1-e^{-n}) \cdot z) \simeq \sqrt{n \log\log n}$. Estimates for the Poisson kernel then imply that μ–a.e. point of Λ is covered by disks such that

$$\mu(D(z,t)) \simeq \varphi(t) = t^2 \sqrt{\log \frac{1}{t} \log\log\log \frac{1}{t}}.$$

In fact, this optimistic calculation is actually correct; the paper Bishop and P. W. Jones [ibid.] shows that Λ has finite, positive Hausdorff φ-measure, verifying a conjecture of Sullivan [1981a]. The optimal gauge φ for the general case, where injectivity radius approaches zero, remains unknown. What about subsets of Λ defined using geodesic rate of escape as in Gönye [2008], or Lundh [2003]?

3 Logarithms, length and Liouville

• **Makarov's theorems:** The LIL above for dendritic limit sets was much easier to discover because the connection between harmonic measure, random walks and Hausdorff dimension had already been uncovered by a celebrated result of Nick Makarov a decade earlier; see Makarov [1985]. Suppose Ω is planar and simply connected. He showed that if

$$\varphi_C(t) = t \exp\left(C \sqrt{\log \frac{1}{1-t} \log\log\log \frac{1}{1-t}} \right),$$

then there is a $C = C_1$ so that $\omega(E) = 0$ whenever E has zero φ_C-measure. However, there is also a $C = C_2$, and a fractal domain Ω, so that $\omega(E) = 1$ for some set $E \subset \partial\Omega$ of φ_C measure zero. In fact, we can take Ω to be the interior of the von Koch snowflake, or any sufficiently "wiggly" fractal (some cases were known earlier, e.g., Carleson's paper Carleson [1985]). Makarov discovered that if $f : \mathbb{D} \to \Omega$ is conformal, then the harmonic function $g = \log|f'|$ behaves precisely like the dyadic martingale $\{u_n\}$ on $\mathbb{T}$ defined on each nth generation dyadic interval $I \subset \mathbb{T}$ by

$$u_n = \lim_{r \nearrow 1} \frac{1}{|I|} \int_I g(re^{i\theta}) d\theta. \tag{1}$$

Distortion estimates for f' imply this limit exists and $|g(z) - u_n(I)| = O(1)$, for any z in the Whitney square corresponding to I. See Figure 13.

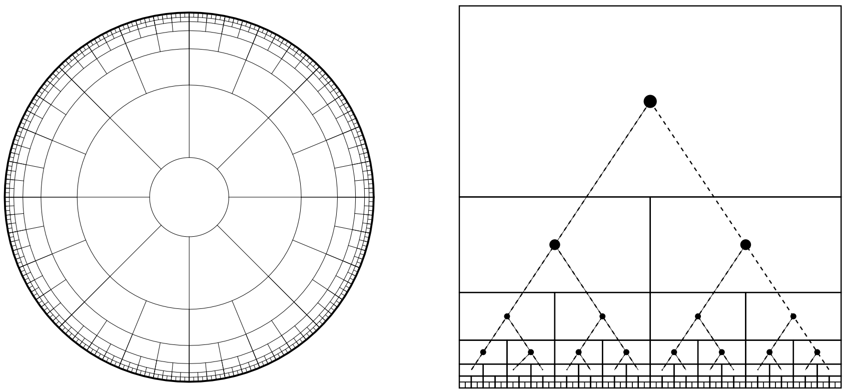


Figure 13: A Whitney decomposition of the disk and an enlargement near the boundary. Each box corresponds to a dyadic interval on the boundary. Although $g = \log|f'|$ is non-constant on each box, it is within $O(1)$ of the associated martingale value.

The $\{u_n\}$ have bounded differences, and the LIL for such martingales implies $|u_n(x)| = O(\sqrt{n \log\log n})$, for a.e. $x \in \mathbb{T}$. This, in turn, gives

$$|g(r \cdot x)| = O\left(\sqrt{\log \frac{1}{1-r} \log\log\log \frac{1}{1-r}}\right),$$

as $r \nearrow 1$ for a.e. $x \in \mathbb{T}$, and this implies Makarov's LIL. Makarov's discovery has since been refined and exploited in many interesting ways, e.g., it makes sense to talk about the asymptotic variance of $g = \log|f'|$ near the boundary and precise estimates for this have led to exciting developments in the theory of conformal and quasiconformal mappings, e.g., see the papers Astala, Ivrii, Perälä, and Prause [2015], Hedenmalm [2017], Ivrii [2016].

Makarov's LIL is just half of a remarkable theorem: $\dim(\omega) = 1$ for any simply connected planar domain, where $\dim(\omega) = \inf_E\{\dim(E) : \omega(E) = 1\}$. Since $\varphi_C(t) = o(t^\alpha)$ for any $\alpha < 1$, the LIL shows $\dim(E) < 1$ implies $\omega(E) = 0$. Hence $\dim(\omega) \geq 1$. On the other hand, since $g = \log|f'|$ behaves like a martingale, along a.e. radius it is either bounded or oscillates between $-\infty$ and ∞. The boundary set where the former happens maps to σ-finite length (since this set is a countable union of sets where $|f'|$ is radially bounded) and the latter set maps to zero length (since $|f'| \to 0$ along some radial sequence). Thus $\dim(\omega) \le 1$. See Pommerenke [1986]. For extensions to general planar domains, see P. W. Jones and Wolff [1988], Wolff [1993].

The obvious generalization to higher dimensions is that $\dim(\omega) = n$ for domains in $\mathbb{R}^{n+1}$. Bourgain [1987], proved $\dim(\omega) \le n + 1 - \epsilon(n)$, Wolff [1995] constructed ingenious fractal "snowballs" in $\mathbb{R}^3$ where $\dim(\omega)$ can be strictly larger than or strictly smaller than 2, so the generalization above is false. In the plane, $\log|\nabla u|$ is subharmonic if u is harmonic, and the failure of this key fact in $\mathbb{R}^3$ is the basis of Wolff's examples. However, in $\mathbb{R}^{n+1}$, $|\nabla u|^p$ is subharmonic if $p > 1 - 1/n$, and this suggests $\dim(\omega) \le n + 1 - 1/n$ for all $\Omega \subset \mathbb{R}^{n+1}$, but this remains completely open.

• **Harmonic measure and rectifiability:** The F. and M. Riesz theorem (F. Riesz and M. Riesz [1920]) states that for a simply connected planar domain with a finite length boundary, harmonic measure and 1-measure are mutually absolutely continuous. Extending this has been a major goal in the study of harmonic measure for the last century.

For example, McMillan [1969] proved that for a general simply connected domain in $\mathbb{R}^2$, ω gives full measure to the union of two special subsets of the boundary: the cone points and the twist points. Cone points are simply vertices of cones inside Ω, and on these points ω and Hausdorff 1-measure are mutually absolutely continuous. McMillan's theorem generalizes the F. and M. Riesz theorem since almost every point of a rectifiable curve is a tangent point, and hence is a cone point for each side.

A point $w \in \partial\Omega$ is a twist point if $\arg(z - w)$ on Ω is unbounded above and below in any neighborhood of w. More geometrically, any curve in Ω terminating at w must twist arbitrarily far in both directions around w. On the twist points we have

$$\limsup_{r\to 0} \frac{\omega(D(x,r))}{r} = \infty, \qquad \liminf_{r\to 0} \frac{\omega(D(x,r))}{r} = 0. \tag{2}$$

The left side is due to Makarov [1985]; it implies that on the twist points, ω is supported on a set of zero length. The right side is due to Choi [2004].

Choi's theorem has an interesting consequence. Suppose E consists of twist points, fix $\epsilon > 0$, and cover ω-a.e. point of E using disjoint disks such that $\omega(D(x_j, r_j)) < \epsilon r_j$ (use the Vitali covering lemma). Then any curve γ containing E has length at least

$$\ell(\gamma) \ge \sum_j r_j \ge \frac{1}{\epsilon} \sum_j \omega(D_j \cap E) \ge \frac{\omega(E)}{\epsilon},$$

i.e., $\ell(\gamma) = \infty$ if $\omega(E) > 0$. This implies the "local" F. and M. Riesz Theorem: if E is a zero length subset of a rectifiable curve, then $\omega(E) = 0$ for any simply connected domain. A quantitative version of this, proven in Bishop and P. W. Jones [1990], Bishop and P. W. Jones [1994], was one of the first applications of Jones' β-numbers and his traveling salesman theorem characterizing planar rectifiable sets in terms of β-numbers P. W.

Jones [1990]. There has been steady progress since this result on the relationship between harmonic measure and rectifiability, and even a sketch of this area would fill a survey longer than this one. A recent landmark, giving a converse to the local Riesz theorem in all dimensions, is due to Azzam, Hofmann, Martell, Mayboroda, Mourgoglou, Tolsa, and Volberg [2016]: if $\omega|_E \ll \mathcal{H}_n|_E$ then $\omega|_E$ is rectifiable (it's support can be covered by countably many Lipschitz graphs). Since ω is the normal derivative of Green's function G, $\omega \ll \mathcal{H}_n$ roughly means that $|\nabla G|$ is bounded near a subset of E, and this implies that the Riesz transforms (which relate the components of ∇G) are bounded operators with respect to ω on a suitable subset. Several recent deep results on singular integral operators and geometric measure theory then imply rectifiability; e.g., see Léger [1999], Nazarov, Tolsa, and Volberg [2014], Nazarov, Treil, and Volberg [2003].

The left side of Equation (2) has an amusing corollary. If $x \in \partial\Omega_1 \cap \partial\Omega_2$, where $\Omega_1, \Omega_2 \subset \mathbb{R}^{n+1}$ are disjoint with harmonic measures ω_1, ω_2 (fix some base point in each), then by Bishop [1991]

$$\omega_1(D(x,r)) \cdot \omega_2(D(x,r)) = O(r^{2n}). \tag{3}$$

Now assume $n = 1$ and $\gamma = \partial\Omega_1 = \partial\Omega_2$ is a closed Jordan curve. By the left side of Equation (2), ω-a.e. twist point of Ω_1 can be covered by disks where $\omega_1(D) \gg r$, so by Equation (3), these disks must also satisfy $\omega_2(D) \ll r \ll \omega_1(D)$. This implies $\omega_1 \perp \omega_2$ on the twist points of γ. On the tangent points of γ, ω_1 and ω_2 are mutually continuous to each other and to 1-measure, so $\omega_1 \perp \omega_2$ on γ if and only if the set of tangent points of γ has zero length; see Bishop, Carleson, Garnett, and P. W. Jones [1989]. This happens for the von Koch snowflake, as well as many other fractal curves. See Figure 14.

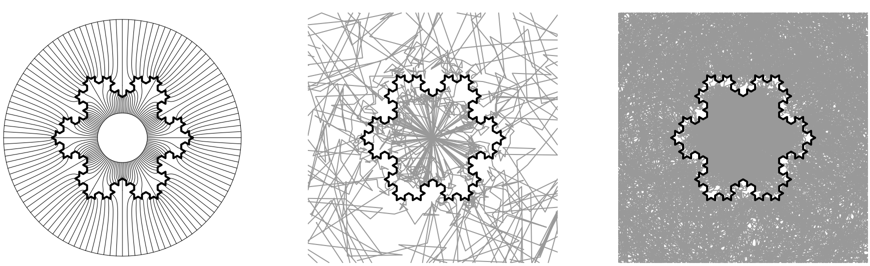


Figure 14: Conformal images of 120 equally spaced radial lines, illustrating the singularity of the inner and outer harmonic measures. On the right are 100 and 1000 Kakutani walks on each side; white shows points that are hard to hit from either side.

One way to think about Equation (2) is to consider a castle whose outer wall is a snowflake. If the fractal fortress is attacked by randomly moving warriors, then only a zero length subset of the wall needs to be defended, whereas if the fortress wall was finite length then it must all be defended. Thus a fractal fortress would be easier to defend (at least against a drunken army). However, because of the local Riesz theorem, it would take an officer infinite time to inspect all the defended positions.

• **Conformal welding:** We would like to compare ω_1, ω_2 for the two sides of a curve γ, but ω_1/ω_2 does not make sense in general. Instead, we consider the orientation preserving (o.p.) circle homeomorphism $h = g^{-1} \circ f$, where f and g are conformal maps from the two sides of the unit circle to the two sides of γ. Such an h is called a "conformal welding" (CW). Not every circle homeomorphism is a conformal welding (see Figure 15), and a useful characterization is likely to be very difficult to find.

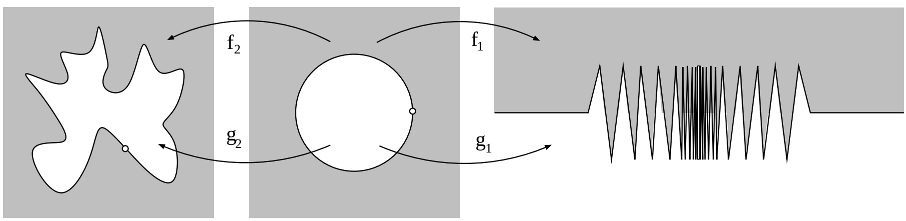


Figure 15: If f_1, g_1 map the two sides of $\mathbb{T}$ to the two sides of a $\sin(1/x)$ curve γ, then $h = g_1^{-1} \circ f_1$ is a homeomorphism, but is not a CW. Otherwise, $h = g_2^{-1} \circ f_2$ with maps corresponding to a Jordan curve, and then (by Morera's theorem) $f_2 \circ f_1^{-1}$ and $g_2 \circ g_1^{-1}$ would define a conformal map from the complement of a segment to the complement of a point, contradicting Liouville's theorem.

If $h(z) = z$, then the maps f, g are equal on $\mathbb{T}$, so by Morera's theorem they define a 1-1 entire function, and this must be linear by Liouville's theorem. Thus only circles can have equal harmonic measures on both sides. If h is bi-Lipschitz with constant near 1, David [1982] showed the corresponding curve is rectifiable, but for large constants the curve can have infinite length (see citeMR852832), or even dimension close to 2 (see Bishop [1988]). Nothing is known about where this transition occurs.

Every "nice" o.p. circle homeomorphism is a conformal welding, where "‘nice" means quasisymmetric; this includes every diffeomorphism but also many singular maps. These send full Lebesgue measure on $\mathbb{T}$ to zero measure; this happens exactly when $\omega_1 \perp \omega_2$, as for the snowflake. Surprisingly, all sufficiently "wild" homeomorphisms are also conformal weldings, where "wild" means log-singular: there is a set E of zero logarithmic capacity on the circle so that $\mathbb{T} \setminus h(E)$ also has zero logarithmic capacity. Zero logarithmic capacity sets are very small, e.g., Hausdorff dimension zero, so log-singular homeomorphisms are very, very singular. See Lundberg [2005]. Moreover, each log-singular map h corresponds to a dense set of all closed curves in the Hausdorff metric, so the association $h \leftrightarrow \gamma$ is far from 1-to-1. See Bishop [2007].

To illustrate the gap between these two cases, consider the space of circle homeomorphisms with the metric $d(f, g) = |\{x : f(x) \neq g(x)\}|$. This space has diameter 2π and the set of QS-homeomorphisms and log-singular homeomorphisms are distance 2π apart. However, conformal weldings are known to be dense in this space; see Bishop [ibid.]. Are they a connected set in this metric? Residual? Borel? For some generalizations and applications of conformal welding, see the papers of Feiszli [2008], Hamilton [1991], and Rohde [2017].

4 True trees and transcendental tracts

• **Dessins d'enfants:** As noted above, a curve γ with $\omega_1 = \omega_2$ must be a circle. Thus in terms of harmonic measure, a circle is the most "natural" way to draw a closed Jordan curve. What happens for other topologies? Can we draw any finite planar tree T so harmonic measure is equal on "both sides"? More precisely, with respect to the point at infinity, can we draw T so that

(1) every edge has equal harmonic measure,

(2) any subset of any edge has equal harmonic measure from both sides?

Perhaps surprisingly, the answer is yes, every finite planar tree T has such drawing, called the "true form of the tree" (or a "true tree" for short). To prove this, consider Figure 16. Let τ be a quasiconformal map of the exterior Ω of T to $\mathbb{D}^* = \{z : |z| > 1\}$, with each side of T mapping to an arc of length π/n, and arclength on each edge mapping to a multiple of arclength in the image. Let $J(z) = \frac{1}{2}(z + \frac{1}{z})$ be the Joukowsky map; this is conformal from $\mathbb{D}^*$ to $U = \mathbb{C} \setminus [-1, 1]$. Then $q(x) = J(\tau(z)^n)$ is quasiregular off T and continuous across T, so is quasiregular on $\mathbb{C}$.

By the measurable Riemann mapping theorem there is a QC "correction" map $\varphi : \mathbb{C} \to \mathbb{C}$ so that $p = q \circ \varphi$ is holomorphic. Since p is also n–to–1, it must be a polynomial of degree n. Its only critical values are ± 1, so it is a generalized Chebyshev, or Shabat, polynomial and $T' = \varphi(T) = p^{-1}([-1, 1])$ is a true tree.

It is easy to see that the polynomial p can be normalized to have its coefficients in some algebraic number field. This connection is part of Grothendieck's' theory of *dessins d'enfants* and is closely connected to the spherical case of Belyi's theorem: a Riemann surface is algebraic iff it supports a meromorphic function ramified over three values. There are many fascinating related problems, e.g., Grothendieck proved that the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts faithfully on the set of planar trees, but the orbits are unknown (some

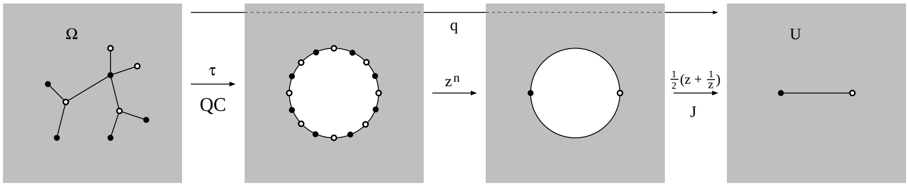


Figure 16: For a true tree, the conformal map $\tau : \mathbb{C} \setminus T \to \mathbb{D}^*$ sends sides of T to arcs of equal length arcs. In general, we choose a QC map τ that sends normalized arclength on sides of T to arclength on $\mathbb{T}$; then $q(z) = J(\tau(z)^n)$ is continuous across T and quasiregular on $\mathbb{C}$.

things are known, e.g., equivalent trees have the same set of vertex degrees). For more background see G. A. Jones and Wolfart [2016], Schneps [1994], Shabat and Zvonkin [1994].

It is a difficult problem to compute the correspondence between trees and polynomials, but this has been done by hand for trees with 10 or fewer edges, Kochetkov [2007], Kochetkov [2014]. It is possible to go much farther using harmonic measure. Don Marshall and Steffen Rohde have adapted Marshall's conformal mapping program ZIPPER; to compute the true form of a given planar tree (even with thousands of edges). See Marshall and Rohde [2007]. For small trees (less than 50 edges or so) they can obtain the vertices (and hence the polynomial) to thousands of digits of accuracy. Given enough digits of an algebraic integer $\alpha \in \mathbb{R}$ one can search for an integer relation among $1, \alpha, \alpha^2, \dots,$ that determines the field, e.g., using Helaman Ferguson's PSLQ algorithm; see Ferguson, Bailey, and Arno [1999].

Alex Eremenko asked if Shabat polynomials have special geometry. In Bishop [2014], I showed the answer is no in the sense that given any compact, connected set K there are polynomials with critical values ± 1 whose critical sets approach K in the Hausdorff metric. In particular, the true tree $T = p^{-1}([-1, 1])$ can be ϵ-close to any connected shape, i.e., "true trees are dense". See Figure 17.

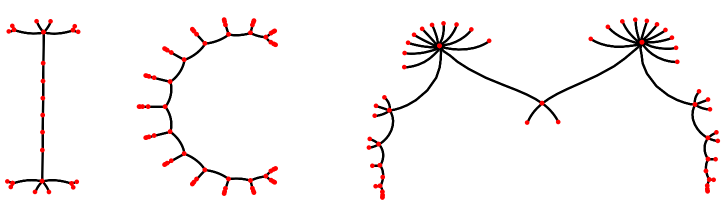


Figure 17: True trees approximating some random letters of the alphabet.

Is there a higher dimensional analog of true trees? In what other settings does "equal harmonic measure from both sides" makes sense and lead to interesting problems? If we drop (1) from the definition of a true tree, then we get trees that connect their vertices using minimum logarithmic capacity. See Stahl [2012].

• **Dessins d'adolescents:** Given the connection between true trees and polynomials, it is natural to ask about a correspondence between infinite planar trees and entire functions, e.g., is every unbounded planar tree T equivalent to $f^{-1}([-1,1])$ for some entire function f with critical values ± 1? The answer is no: one can show the infinite 3-regular tree is not of this form. However, a version of the "true trees are dense" construction does hold. Consider how to adapt the construction in Figure 16 to unbounded trees, as in Figure 18. Now, $\Omega = \mathbb{C} \setminus T$ is a union of unbounded, simply connected domains, called tracts, and each of these tracts can be mapped to $\mathbb{H}_r = \{x + iy : x > 0\}$, by a conformal map τ. The power function z^n is replaced by $\exp : \mathbb{H}_r \to \mathbb{D}^*$, but is still followed by the Joukowsky map, giving a holomorphic function $F(z) = J(\exp(\tau(z)))$ on each tract, but F need not be continuous across T. Fixing this requires some assumptions (some regularity of T that replaces finiteness). Via τ, the vertices of T define a partition of $i\mathbb{R} = \partial\mathbb{H}_r$ and we assume that this partition satisfies

(1) adjacent intervals have comparable length,

(2) interval lengths are all $\geq \pi$.

Under these hypotheses, the QC-folding theorem from Bishop [2015] gives a quasi-regular g that agrees with F outside $T(r) = \cup_{e \in T}\{z : \mathrm{dist}(z,e) < r \cdot \mathrm{diam}(e)\}$, where the union is over all edges in T. The tree $T' = g^{-1}([-1,1])$ satisfies $T \subset T' \subset T(r)$. The measurable Riemann mapping theorem gives a quasiconformal φ so that $f = g \circ \varphi^{-1}$ is an entire function with critical values ± 1 and no other singular values (the singular set $S(f)$ is the closure of the critical values and finite asymptotic values, i.e., limits of f along curves to ∞).

Since g is holomorphic off $T(r)$, μ_φ is supported in $T(r)$ and is uniformly bounded in terms of the assumptions on T. In many applications $T(r)$ has finite, even small, area, and φ is close to the identity. Thus the QC-folding theorem converts an infinite planar tree T satisfying some mild restrictions into an entire function f with $S(f) = \{\pm 1\}$, and such that $T' = f^{-1}([-1,1])$ is "close to" T in a precise sense.

Let $\mathcal{T}$ denote the transcendental entire functions (non-polynomials). The Speiser class is $\mathcal{S} = \{f \in \mathcal{T} : S(f) \text{ is finite}\}$, and the Eremenko-Lyubich class is $\mathcal{B} = \{f \in \mathcal{T} : S(f) \text{ is bounded}\}$. The QC-folding theorem (or simple modifications) gives a flexible way to construct examples in $\mathcal{S}$ and $\mathcal{B}$ with specified singular sets, including:

► A $f \in \mathcal{B}$ with a wandering domain. Wandering domains do not exist for rational functions by Sullivan's non-wandering theorem Sullivan [1985], nor in $\mathcal{S}$ by work of Eremenko and Lyubich [1992] and Goldberg and Keen [1986]. See Figure 19. See also the paper of Lazebnik [2017].

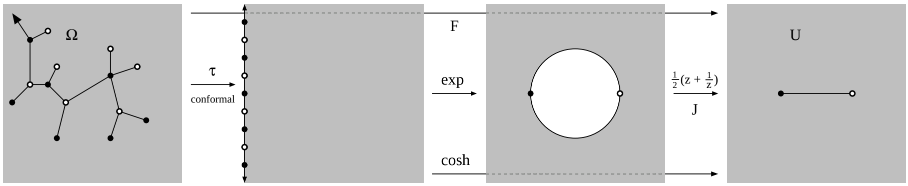


Figure 18: The transcendental version of Figure 16. F is holomorphic off T but not necessarily across T. QC-folding defines a quasiregular g so that $g = F$ outside a "small" neighborhood of T.

► A $f \in \mathcal{S}$ so that area$(\{z : |f(z)| > \epsilon\}) < \infty$ for all ϵ. This is a strong counterexample to the area conjecture of Eremenko and Lyubich [1992].

► A $f \in \mathcal{S}$ whose escaping set has no non-trivial path components; this improves the counterexample to the strong Eremenko conjecture in $\mathcal{B}$ due to Rottenfusser, Rückert, Rempe, and Schleicher [2011].

► A $f \in \mathcal{S}$ so that $\limsup_{r\to\infty} \log m(r, f)/\log M(r, f) = -\infty$ where m, M denote the min, max of $|f|$ on $\{|z| = r\}$. In 1916 Wiman had conjectured $\limsup \geq -1$, as occurs for $\exp(z)$. Beurling [1949] gave a partial positive result, but Hayman [1952] found a counterexample in general, and QC-folding now improves this to $\mathcal{S}$.

► $f \in \mathcal{S}$ with Julia sets so that $\dim(\mathcal{J}) < 1 + \epsilon$ Albrecht and Bishop [2017]. Examples in $\mathcal{B}$ are due to Stallard [1997], Stallard [2000], who also showed $\dim(\mathcal{J}) > 1$ for $f \in \mathcal{B}$. Baker [1975] proved $\dim(\mathcal{J}) \geq 1$ for all $f \in \mathcal{T}$, and examples with $\dim(\mathcal{J}) = 1$ (even with finite spherical linear measure) exist Bishop [2018], but it is unknown whether they can lie on a rectifiable curve on the sphere.

References

I. Agol (2004). "Tameness of hyperbolic 3-manifolds". arXiv: math/0405568 (cit. on p. 1541).

Lars Valerian Ahlfors (2006). *Lectures on quasiconformal mappings*. Second. Vol. 38. University Lecture Series. With supplemental chapters by C. J. Earle, I. Kra, M. Shishikura and J. H. Hubbard. American Mathematical Society, Providence, RI, pp. viii+162. MR: 2241787 (cit. on pp. 1532, 1535).

S. Albrecht and C. J. Bishop (2017). "Speiser class Julia sets with dimension near one". preprint (cit. on p. 1550).

K. Astala, O. Ivrii, A. Perälä, and I. Prause (2015). "Asymptotic variance of the Beurling transform". *Geom. Funct. Anal.* 25.6, pp. 1647–1687. MR: 3432154 (cit. on p. 1543).

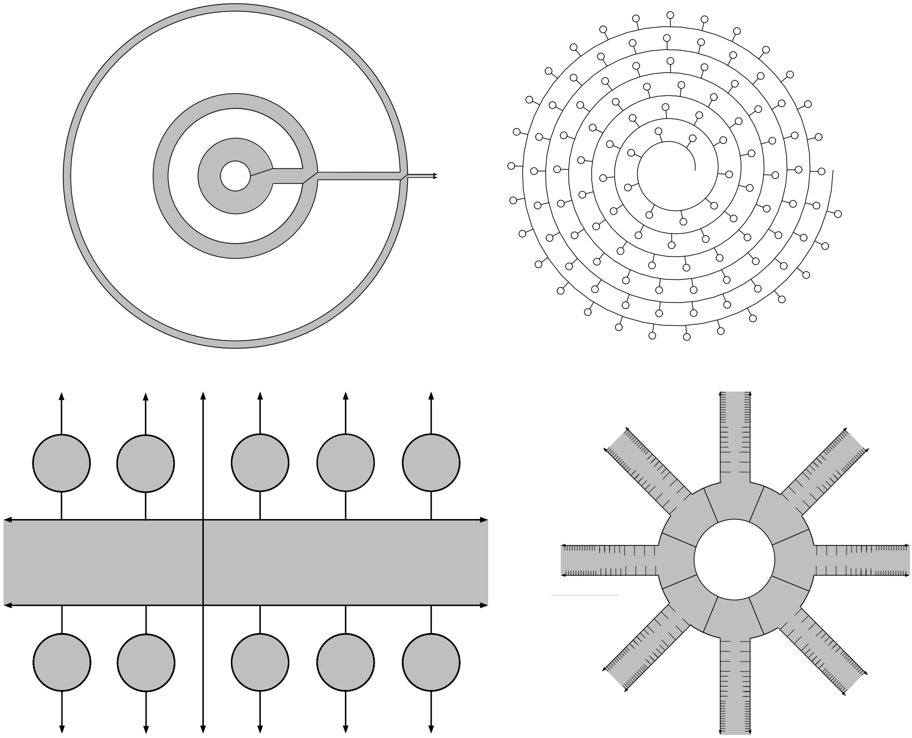


Figure 19: The folding theorem reduces constructing certain entire functions to drawing a picture. Here are the pictures associated to counterexamples for the area conjecture (upper left), Wiman's conjecture (upper right), an Eremenko-Lyubich wandering domain (lower left) and a Speiser class Julia set of dimension near 1.

J. Azzam, S. Hofmann, J. M. Martell, S. Mayboroda, M. Mourgoglou, X. Tolsa, and A. Volberg (2016). "Harmonic measure is rectifiable if it is absolutely continuous with respect to the co-dimension-one Hausdorff measure". *C. R. Math. Acad. Sci. Paris* 354.4, pp. 351–355. MR: 3473548 (cit. on p. 1545).

I. N. Baker (1975). "The domains of normality of an entire function". *Ann. Acad. Sci. Fenn. Ser. A I Math.* 1.2, pp. 277–283. MR: 0402044 (cit. on p. 1550).

Arne Beurling (1949). "Some theorems on boundedness of analytic functions". *Duke Math. J.* 16, pp. 355–359. MR: 0029980 (cit. on p. 1550).

I. Binder and M. Braverman (2012). "The rate of convergence of the walk on spheres algorithm". *Geom. Funct. Anal.* 22.3, pp. 558–587. MR: 2972600 (cit. on p. 1530).

I. Binder, M. Braverman, and M. Yampolsky (2007). "On the computational complexity of the Riemann mapping". *Ark. Mat.* 45.2, pp. 221–239. MR: 2342601 (cit. on p. 1531).

C. J. Bishop (1988). "A counterexample in conformal welding concerning Hausdorff dimension". *Michigan Math. J.* 35.1, pp. 151–159. MR: 931947 (cit. on p. 1546).

– (1991). "A characterization of Poissonian domains". *Ark. Mat.* 29.1, pp. 1–24. MR: 1115072 (cit. on p. 1545).

– (2001). "Divergence groups have the Bowen property". *Ann. of Math. (2)* 154.1, pp. 205–217. MR: 1847593 (cit. on p. 1537).

– (2002). "Quasiconformal Lipschitz maps, Sullivan's convex hull theorem and Brennan's conjecture". *Ark. Mat.* 40.1, pp. 1–26. MR: 1948883 (cit. on p. 1537).

– (2007). "Conformal welding and Koebe's theorem". *Ann. of Math. (2)* 166.3, pp. 613–656. MR: 2373370 (cit. on p. 1547).

– (2010a). "Conformal mapping in linear time". *Discrete Comput. Geom.* 44.2, pp. 330–428. MR: 2671015 (cit. on pp. 1531, 1532, 1534).

– (2010b). "Optimal angle bounds for quadrilateral meshes". *Discrete Comput. Geom.* 44.2, pp. 308–329. MR: 2671014 (cit. on p. 1534).

– (2014). "True trees are dense". *Invent. Math.* 197.2, pp. 433–452. MR: 3232011 (cit. on p. 1548).

– (2015). "Constructing entire functions by quasiconformal folding". *Acta Math.* 214.1, pp. 1–60. MR: 3316755 (cit. on p. 1549).

– (2016a). "Nonobtuse triangulations of PSLGs". *Discrete Comput. Geom.* 56.1, pp. 43–92. MR: 3509031 (cit. on p. 1534).

– (2016b). "Quadrilateral meshes for PSLGs". *Discrete Comput. Geom.* 56.1, pp. 1–42. MR: 3509030 (cit. on p. 1534).

– (2018). "A transcendental Julia set of dimension 1". to appear in Invent. Math. (cit. on p. 1550).

C. J. Bishop, L. Carleson, J. B. Garnett, and P. W. Jones (1989). "Harmonic measures supported on curves". *Pacific J. Math.* 138.2, pp. 233–236. MR: 996199 (cit. on p. 1545).

C. J. Bishop and H. Hakobyan (2008). "A central set of dimension 2". *Proc. Amer. Math. Soc.* 136.7, pp. 2453–2461. MR: 2390513 (cit. on p. 1536).

C. J. Bishop, H. Hakobyan, and M. Williams (2016). "Quasisymmetric dimension distortion of Ahlfors regular subsets of a metric space". *Geom. Funct. Anal.* 26.2, pp. 379–421. MR: 3513876 (cit. on p. 1533).

C. J. Bishop and P. W. Jones (1990). "Harmonic measure and arclength". *Ann. of Math. (2)* 132.3, pp. 511–547. MR: 1078268 (cit. on p. 1544).

– (1994). "Harmonic measure, L^2 estimates and the Schwarzian derivative". *J. Anal. Math.* 62, pp. 77–113. MR: 1269200 (cit. on p. 1544).

– (1997a). "Hausdorff dimension and Kleinian groups". *Acta Math.* 179.1, pp. 1–39. MR: 1484767 (cit. on pp. 1540, 1541).

– (1997b). "The law of the iterated logarithm for Kleinian groups". In: *Lipa's legacy (New York, 1995)*. Vol. 211. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 17–50. MR: 1476980 (cit. on p. 1542).

C. J. Bishop and C. LeBrun (2017). "Anti-Self-Dual 4-Manifolds, Quasi-Fuchsian Groups, and Almost-Kähler Geometry". arXiv: 1708.03824 (cit. on p. 1539).

F. Bornemann, D. Laurie, S. Wagon, and J. Waldvogel (2004). *The SIAM 100-digit challenge*. SIAM, Philadelphia, PA, pp. xii+306. MR: 2076374 (cit. on p. 1531).

J. Bourgain (1987). "On the Hausdorff dimension of harmonic measure in higher dimension". *Invent. Math.* 87.3, pp. 477–483. MR: 874032 (cit. on p. 1544).

M. Bridgeman and R. D. Canary (2010). "The Thurston metric on hyperbolic domains and boundaries of convex hulls". *Geom. Funct. Anal.* 20.6, pp. 1317–1353. MR: 2738995 (cit. on p. 1537).

D. Calegari and D. Gabai (2006). "Shrinkwrapping and the taming of hyperbolic 3-manifolds". *J. Amer. Math. Soc.* 19.2, pp. 385–446. MR: 2188131 (cit. on p. 1541).

Lennart Carleson (1985). "On the support of harmonic measure for sets of Cantor type". *Ann. Acad. Sci. Fenn. Ser. A I Math.* 10, pp. 113–123. MR: 802473 (cit. on p. 1542).

F. Chin, J. Snoeyink, and C. A. Wang (1999). "Finding the medial axis of a simple polygon in linear time". *Discrete Comput. Geom.* 21.3, pp. 405–420. MR: 1672988 (cit. on p. 1536).

S. Choi (2004). "The lower density conjecture for harmonic measure". *J. Anal. Math.* 93, pp. 237–269. MR: 2110330 (cit. on p. 1544).

E. B. Christoffle (1867). "Sul problema della tempurature stazonaire e la rappresetazione di una data superficie". *Ann. Mat. Pura Appl. Serie II*, pp. 89–103 (cit. on p. 1531).

G. David (1982). "Courbes corde-arc et espaces de Hardy généralisés". *Ann. Inst. Fourier (Grenoble)* 32.3, pp. xi, 227–239. MR: 688027 (cit. on p. 1546).

E. B. Davies (1988). "Gaussian upper bounds for the heat kernels of some second-order operators on Riemannian manifolds". *J. Funct. Anal.* 80.1, pp. 16–32. MR: 960220 (cit. on p. 1540).

R. T. Davis (1979). "Numerical methods for coordinate generation based on Schwarz-Christoffel transformations". In: *4th AIAA Comput. Fluid Dynamics Conf., Williamsburg VA*, pp. 1–15 (cit. on p. 1532).

T. K. DeLillo (1994). "The accuracy of numerical conformal mapping methods: a survey of examples and results". *SIAM J. Numer. Anal.* 31.3, pp. 788–812. MR: 1275114 (cit. on p. 1531).

T. A. Driscoll and L. N. Trefethen (2002). *Schwarz-Christoffel mapping*. Vol. 8. Cambridge Monographs on Applied and Computational Mathematics. Cambridge: Cambridge University Press, pp. xvi+132. MR: 2003e:30012 (cit. on p. 1531).

T. A. Driscoll and S. A. Vavasis (1998). "Numerical conformal mapping using cross-ratios and Delaunay triangulation". *SIAM J. Sci. Comput.* 19.6, 1783–1803 (electronic). MR: 99d:65086 (cit. on p. 1532).

D. B. A. Epstein and A. Marden (1987). "Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces". In: *Analytical and geometric aspects of hyperbolic space (Coventry/Durham, 1984)*. Vol. 111. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, pp. 113–253. MR: 903852 (cit. on p. 1537).

A. E. Eremenko and M. Yu. Lyubich (1992). "Dynamical properties of some classes of entire functions". *Ann. Inst. Fourier (Grenoble)* 42.4, pp. 989–1020. MR: 1196102 (cit. on pp. 1549, 1550).

M. Feiszli (2008). *Conformal shape representation*. Thesis (Ph.D.)–Brown University. ProQuest LLC, Ann Arbor, MI, p. 161. MR: 2717596 (cit. on p. 1547).

H. R. P. Ferguson, D. H. Bailey, and S. Arno (1999). "Analysis of PSLQ, an integer relation finding algorithm". *Math. Comp.* 68.225, pp. 351–369. MR: 1489971 (cit. on p. 1548).

L. R. Goldberg and L. Keen (1986). "A finiteness theorem for a dynamical class of entire functions". *Ergodic Theory Dynam. Systems* 6.2, pp. 183–192. MR: 857196 (cit. on p. 1549).

Z. Gönye (2008). "Dimension of escaping geodesics". *Trans. Amer. Math. Soc.* 360.10, pp. 5589–5602. MR: 2415087 (cit. on p. 1542).

M. Goswami, X. Gu, V. P. Pingali, and G. Telang (2015). "Computing Teichmüller maps between polygons". In: *31st International Symposium on Computational Geometry*. Vol. 34. LIPIcs. Leibniz Int. Proc. Inform. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, pp. 615–629. MR: 3392809 (cit. on p. 1533).

C. M. Green (2011). *The Ahlfors Iteration for Conformal Mapping*. Thesis (Ph. D.)–State University of New York at Stony Brook. ProQuest LLC, Ann Arbor, MI, p. 72. MR: 2995922 (cit. on p. 1535).

A. Grigor'yan (1995). "Heat kernel of a noncompact Riemannian manifold". In: *Stochastic analysis (Ithaca, NY, 1993)*. Vol. 57. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 239–263. MR: 1335475 (cit. on p. 1540).

D. H. Hamilton (1991). "Generalized conformal welding". *Ann. Acad. Sci. Fenn. Ser. A I Math.* 16.2, pp. 333–343. MR: 1139801 (cit. on p. 1547).

W. K. Hayman (1952). "The minimum modulus of large integral functions". *Proc. London Math. Soc. (3)* 2, pp. 469–512. MR: 0056083 (cit. on p. 1550).

H. Hedenmalm (2017). "Bloch functions and asymptotic tail variance". *Adv. Math.* 313, pp. 947–990. MR: 3649242 (cit. on p. 1543).

O. Ivrii (2016). "Quasicircles of dimension $1 + k^2$ do not exist". arXiv: 1511.07240 (cit. on p. 1543).

G. A. Jones and J. Wolfart (2016). *Dessins d'enfants on Riemann surfaces*. Springer Monographs in Mathematics. Springer, Cham, pp. xiv+259. MR: 3467692 (cit. on p. 1548).

P. W. Jones (1990). "Rectifiable sets and the traveling salesman problem". *Invent. Math.* 102.1, pp. 1–15. MR: 1069238 (cit. on p. 1544).

P. W. Jones and T. H. Wolff (1988). "Hausdorff dimension of harmonic measures in the plane". *Acta Math.* 161.1-2, pp. 131–144. MR: 962097 (cit. on p. 1544).

Yu. Yu. Kochetkov (2007). "Planar trees with nine edges: a catalogue". *Fundam. Prikl. Mat.* 13.6. translation in J. Math. Sci. (N. Y.) 158 (2009), no. 1, 114–140, pp. 159–195. MR: 2476034 (cit. on p. 1548).

– (2014). "Short catalog of plane ten-edge trees". arXiv: 1412.2472 (cit. on p. 1548).

K. Lazebnik (2017). "Several constructions in the Eremenko-Lyubich class". *J. Math. Anal. Appl.* 448.1, pp. 611–632. MR: 3579902 (cit. on p. 1549).

J. C. Léger (1999). "Menger curvature and rectifiability". *Ann. of Math. (2)* 149.3, pp. 831–869. MR: 1709304 (cit. on p. 1545).

K. A. Lundberg (2005). *Boundary behavior of uniformly convergent conformal maps*. Thesis (Ph.D.)–State University of New York at Stony Brook. ProQuest LLC, Ann Arbor, MI, p. 81. MR: 2707849 (cit. on p. 1547).

T. Lundh (2003). "Geodesics on quotient manifolds and their corresponding limit points". *Michigan Math. J.* 51.2, pp. 279–304. MR: 1992947 (cit. on p. 1542).

N. G. Makarov (1985). "On the distortion of boundary sets under conformal mappings". *Proc. London Math. Soc. (3)* 51.2, pp. 369–384. MR: 794117 (cit. on pp. 1542, 1544).

D. E. Marshall and S. Rohde (2007). "Convergence of a variant of the zipper algorithm for conformal mapping". *SIAM J. Numer. Anal.* 45.6, 2577–2609 (electronic). MR: 2361903 (cit. on p. 1548).

J. E. McMillan (1969). "Boundary behavior of a conformal mapping". *Acta Math.* 123, pp. 43–67. MR: 0257330 (cit. on p. 1544).

S. A. Mitchell (1993). "Refining a triangulation of a planar straight-line graph to eliminate large angles". In: *34th Annual Symposium on Foundations of Computer Science (Palo Alto, CA, 1993)*. Los Alamitos, CA: IEEE Comput. Soc. Press, pp. 583–591. MR: MR1328453 (cit. on p. 1534).

F. Nazarov, X. Tolsa, and A. Volberg (2014). "On the uniform rectifiability of AD-regular measures with bounded Riesz transform operator: the case of codimension 1". *Acta Math.* 213.2, pp. 237–321. MR: 3286036 (cit. on p. 1545).

F. Nazarov, S. Treil, and A. Volberg (2003). "The Tb-theorem on non-homogeneous spaces". *Acta Math.* 190.2, pp. 151–239. MR: 1998349 (cit. on p. 1545).

N. Papamichael and E. B. Saff, eds. (1993). *Computational complex analysis*. J. Comput. Appl. Math. **46** (1993), no. 1-2. Amsterdam: North-Holland Publishing Co., i–iv and 1–326. MR: 93m:00020 (cit. on p. 1531).

Ch. Pommerenke (1986). "On conformal mapping and linear measure". *J. Analyse Math.* 46, pp. 231–238. MR: 861701 (cit. on p. 1544).

C. C. Pugh (1967). "The closing lemma". *Amer. J. Math.* 89, pp. 956–1009. MR: 0226669 (cit. on p. 1535).

F. Riesz and M. Riesz (1920). "Über die randwerte einer analytischen funktion". In: Almqvists and Wilkets, Uppsala, pp. 27–44 (cit. on p. 1544).

S. Rohde (2017). "Conformal laminations and Gehring trees" (cit. on p. 1547).

G. Rottenfusser, J. Rückert, L. Rempe, and D. Schleicher (2011). "Dynamic rays of bounded-type entire functions". *Ann. of Math. (2)* 173.1, pp. 77–125. MR: 2753600 (cit. on p. 1550).

S. Salzberg, A. L. Delcher, D Heath, and S. Kasif (1995). "Best-case results for nearest-neighbor learning". *IEEE Trans. Pattern Anal. MAch. Intell.* 17.6, pp. 599–608 (cit. on p. 1535).

L. Schneps, ed. (1994). *The Grothendieck theory of dessins d'enfants*. Vol. 200. London Mathematical Society Lecture Note Series. Papers from the Conference on Dessins d'Enfant held in Luminy, April 19–24, 1993. Cambridge University Press, Cambridge, pp. iv+368. MR: 1305390 (cit. on p. 1548).

H. A. Schwarz (1869). "Conforme Abbildung der Oberfläche eines Tetraeders auf die Oberfläche einer Kugel". *J. Reine Ange. Math.* Also in collected works, Schwarz [1890], pp. 84-101, pp. 121–136 (cit. on p. 1531).

– (1890). *Gesammelte Mathematische Abhandlungen*. Springer, Berlin (cit. on pp. 1531, 1556).

G. B. Shabat and A. Zvonkin (1994). "Plane trees and algebraic numbers". In: *Jerusalem combinatorics '93*. Vol. 178. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 233–275. MR: 1310587 (cit. on p. 1548).

H. R. Stahl (2012). "Sets of minimal capacity and extremal domains". arXiv: 1205.3811 (cit. on p. 1549).

G. M. Stallard (1997). "The Hausdorff dimension of Julia sets of entire functions. III". *Math. Proc. Cambridge Philos. Soc.* 122.2, pp. 223–244. MR: 1458228 (cit. on p. 1550).

– (2000). "The Hausdorff dimension of Julia sets of entire functions. IV". *J. London Math. Soc. (2)* 61.2, pp. 471–488. MR: 1760674 (cit. on p. 1550).

D. Sullivan (1981a). "Growth of positive harmonic functions and Kleinian group limit sets of zero planar measure and Hausdorff dimension two". In: *Geometry Symposium, Utrecht 1980 (Utrecht, 1980)*. Vol. 894. Lecture Notes in Math. Springer, Berlin-New York, pp. 127–144. MR: 655423 (cit. on p. 1542).

– (1981b). "Travaux de Thurston sur les groupes quasi-Fuchsiens et les variétés hyperboliques de dimension 3 fibrées sur S^1". In: *Bourbaki Seminar, Vol. 1979/80*. Vol. 842. Lecture Notes in Math. Springer, Berlin-New York, pp. 196–214. MR: 636524 (cit. on p. 1537).

– (1984). "Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups". *Acta Math.* 153.3-4, pp. 259–277. MR: 766265 (cit. on pp. 1540, 1541).

– (1985). "Quasiconformal homeomorphisms and dynamics. I. Solution of the Fatou-Julia problem on wandering domains". *Ann. of Math. (2)* 122.3, pp. 401–418. MR: 819553 (cit. on p. 1549).

T.-S. Tan (1996). "An optimal bound for high-quality conforming triangulations". *Discrete Comput. Geom.* 15.2, pp. 169–193. MR: 1368273 (cit. on p. 1534).

L. N. Trefethen, ed. (1986). *Numerical conformal mapping*. Reprint of J. Comput. Appl. Math. **14** (1986), no. 1-2. Amsterdam: North-Holland Publishing Co., pp. iv+269. MR: 87m:30013 (cit. on p. 1531).

R. Wegmann (2005). "Methods for numerical conformal mapping". In: *Handbook of complex analysis: geometric function theory. Vol. 2*. Amsterdam: Elsevier, pp. 351–477. MR: MR2121864 (cit. on p. 1531).

Thomas H. Wolff (1993). "Plane harmonic measures live on sets of σ-finite length". *Ark. Mat.* 31.1, pp. 137–172. MR: 1230270 (cit. on p. 1544).

– (1995). "Counterexamples with harmonic gradients in $\mathbf{R}^3$". In: *Essays on Fourier analysis in honor of Elias M. Stein (Princeton, NJ, 1991)*. Vol. 42. Princeton Math. Ser. Princeton Univ. Press, Princeton, NJ, pp. 321–384. MR: 1315554 (cit. on p. 1544).

Received 2017-11-29.

Christopher J. Bishop
Mathematics Department
Stony Brook University
Stony Brook, NY 11794-3651
bishop@math.stonybrook.edu

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1557–1578)

DECOUPLINGS AND APPLICATIONS

CIPRIAN DEMETER

Abstract

We describe a Fourier analytic tool that has found a large number of applications in Number Theory, Harmonic Analysis and PDEs.

1 Introduction

The circle of ideas described in this note have grown inside the framework of *restriction theory*. This area of harmonic analysis was born in the late 60s, when Elias Stein has considered the problem of restricting the Fourier transform of an L^p function $F : \mathbb{R}^n \to \mathbb{C}$ to the sphere $\mathbb{S}^{n-1}$. When $p = 1$, the Fourier transform is always a continuous function, its value is well defined at each point. At the other extreme, when $p = 2$, $\widehat{F}$ is merely measurable, so restricting it to a set of Lebesgue measure zero such as $\mathbb{S}^{n-1}$ is meaningless. It turns out that the range $1 < p < 2$ hosts a completely new phenomenon. A plethora of restriction-type estimates exist in this range, for a wide variety of curved manifolds other than the sphere. These are quantified by various operator bounds on the so-called *extension operator*, to be introduced momentarily.

A major breakthrough in the analysis of the extension operator came with the discovery of its relation to the quantitative forms of the *Kakeya set conjecture*. In one of its simplest forms, this conjecture asserts that each subset of $\mathbb{R}^n$ containing a unit line segment in every direction must have full Hausdorff dimension n. This is trivial when $n = 1$, relatively easy to prove when $n = 2$, and wide open for $n \geq 3$. The quantitative formulations of the conjecture involve estimating the overlap of collections of congruent tubes of arbitrary orientations. The afore-mentioned extension operator has an oscillatory nature, but it can be decomposed into pieces which have roughly constant magnitude on appropriate tubes. Then one can gain valuable information by understanding the worst conspiracies that tubes can use to maximize their overlap. Using the intuition from the case when tubes

The author is partially supported by the NSF Grant DMS-1161752.
MSC2010: primary 42B20; secondary 11L05, 42B37.

are replaced with lines can sometimes be helpful, though it has been well documented that the thickness of the tubes creates significant additional complications.

It turns out that the overlap question is easier to understand if one intersects families of tubes with very separated directions. This property will be called *transversality*. A critical role in the arguments presented in the following is played by the multilinear Kakeya estimate of Bennett, Carbery, and Tao [2006], which proves a sharp bound on the intersection of n transverse families of tubes. The way to harness the power of multilinear estimates in order to prove linear ones was explained by Bourgain and Guth in the fundamental paper Bourgain and Guth [2011].

A built-in feature of any restriction estimate is that of *scale*. Scales arise by localizing the operator to spatial balls of finite radius. The operator norms at various scales are typically compared using a process called *induction on scales*. A bootstrapping argument forces these operator norms to only grow mildly with the scale. Sometimes an ϵ removal argument is available to completely eliminate this dependence. In other cases, such as with decouplings, finding an ϵ removal mechanism continues to remain a challenge. *Parabolic rescaling* and its variants is a key tool that allows moving back and forth between different scales. This exploits the invariance of the manifold under certain affine transformations which interact well with the Fourier transform.

We will start by analyzing a few classical exponential sum estimates and will continue by showing how decouplings lead to new ones. We close by presenting the proof of the simplest decoupling at critical exponent, an L^6 result for the parabola.

2 Stein–Tomas–Strichartz and exponential sum estimates on small balls

We will denote by $e(z)$ the quantity $e^{2\pi i z}$, $z \in \mathbb{R}$. For $F \in L^1(\mathbb{R}^n)$ we recall its Fourier transform

$$\widehat{F}(\xi) = \int_{\mathbb{R}^n} F(x) e(-x \cdot \xi) dx, \quad \xi \in \mathbb{R}^n.$$

Let D be an open cube, ball or annulus in $\mathbb{R}^m$, $1 \le m \le n-1$. Given a smooth function $\psi : D \to \mathbb{R}^{n-m}$ we define the manifold

$$\mathfrak{M} = \mathfrak{M}^{\psi} = \{(\xi, \psi(\xi)) : \ \xi \in D\} \tag{1}$$

and its associated extension operator for $f : D \to \mathbb{C}$

$$Ef(x) = E^{\mathfrak{M}} f(x) = \int_D f(\xi) e(\bar{x} \cdot \xi + x^* \cdot \psi(\xi)) d\xi, \quad x = (\bar{x}, x^*) \in \mathbb{R}^m \times \mathbb{R}^{n-m}.$$

For a subset $S \subset D$ we will denote $E(f1_S)$ by $E_S f$. The defining formula shows that $E_S f$ is the Fourier transform of the pullback of the measure $f d\xi$ from $\mathbb{R}^m$ to the manifold.

Examples of interesting manifolds arising this way include the truncated paraboloid

$$\mathbb{P}^{n-1} = \{(\xi_1, \dots, \xi_{n-1}, \xi_1^2 + \dots + \xi_{n-1}^2) : |\xi_i| < 1\},$$

the hemispheres

$$\mathbb{S}_{\pm}^{n-1} = \{(\xi, \pm\sqrt{1 - |\xi|^2}),\ |\xi| < 1\},$$

the truncated cone

$$\mathbb{C}o^{n-1} = \{(\xi, |\xi|) :\ 1 < |\xi| < 2\}$$

and the moment curve

$$\Gamma_n = \{(\xi, \xi^2, \dots, \xi^n) :\ \xi \in (0,1)\}.$$

To provide the reader with some motivation for considering the extension operator, let $\Psi(\bar{x}, t)$ be the solution of the free Schrödinger equation with initial data g

$$\begin{cases} 2\pi i \Psi_t = \Delta_{\bar{x}} \Psi, & (\bar{x}, t) \in \mathbb{R}^{n-1} \times \mathbb{R} \\ \Psi(\bar{x}, 0) = g(\bar{x}) \end{cases}.$$

A simple computation reveals that $\Psi(\bar{x}, t) = E^{\mathbb{P}^{n-1}} f(\bar{x}, t)$, where $f = \widehat{g}$. A similar relation exists between the cone and the wave equation and also between the sphere and the Helmholtz equation.

The following theorem provides the first wave of restriction estimates that were ever obtained. They are due to Stein and Tomas in the case of the (hemi)sphere, and to Strichartz in the case of the paraboloid. What makes them special is the fact that the function f is estimated in L^2. The core of the argument relies on the TT^* method.

Theorem 2.1 (Stein [1993], Strichartz [1977], Tomas [1975]). *Let E be the extension operator associated with either $\mathbb{S}_{\pm}^{n-1}$ or $\mathbb{P}^{n-1}$. Then for each $p \ge \frac{2(n+1)}{n-1}$ and $f \in L^2(D)$ we have*

$$\|Ef\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_2.$$

There is an equivalent way to rephrase this theorem, using a rather standard local to global mechanism. The resulting inequality is an example of *discrete restriction estimate.*

Corollary 2.2. *Let $\mathfrak{M}$ be either $\mathbb{S}^{n-1}$ or $\mathbb{P}^{n-1}$. For each $R \geq 1$, each collection $\Lambda \subset \mathfrak{M}$ consisting of $\frac{1}{R}$-separated points, each sequence $a_\lambda \in \mathbb{C}$ and each ball B_R of radius R in $\mathbb{R}^n$ we have*

$$\|\sum_{\lambda\in\Lambda} a_\lambda e(\lambda\cdot x)\|_{L^{\frac{2(n+1)}{n-1}}(B_R)} \lesssim R^{\frac{n-1}{2}}\|a_\lambda\|_{l^2}. \tag{2}$$

If we introduce the normalized L^p norms

$$\|F\|_{L^p_\sharp(B)} := (\frac{1}{|B|}\int_B |F|^p)^{1/p},$$

then (2) says that

$$\|\sum_{\lambda\in\Lambda} a_\lambda e(\lambda\cdot x)\|_{L_\sharp^{\frac{2(n+1)}{n-1}}(B_R)} \lesssim R^{\frac{n-1}{2(n+1)}}\|a_\lambda\|_{l^2}. \tag{3}$$

The exponent of R is sharp, as can be seen by taking $a_\lambda \equiv 1$ and Λ a maximal $\frac{1}{R}$-separated set.

We will call the scale R of the spatial balls B_R the *uncertainty principle scale*, as it is the reciprocal of the scale that separates the frequency points λ. Since averages over large balls are controlled by averages over smaller balls, inequality (3) persists if B_R is replaced with $B_{R'}$ for $R' \geq R$. However, we will see that averaging the exponential sums over larger spatial balls will lead to improved estimates. This will be a direct consequence of the new decoupling phenomenon. In short, the waves $e(\lambda\cdot x)$ oscillate in different directions, and annihilate each other better if they are given more room to interact.

3 A first look at decouplings

Let $(f_j)_{j=1}^N$ be N elements of a Banach space X. The triangle inequality

$$\|\sum_{j=1}^N f_j\|_X \leq \sum_{j=1}^N \|f_j\|_X$$

is universal, it does not incorporate any possible cancelations between the f_j. When combined with the Cauchy–Schwarz inequality it leads to

$$\|\sum_{j=1}^N f_j\|_X \leq N^{\frac{1}{2}}(\sum_{j=1}^N \|f_j\|_X^2)^{1/2}.$$

But if X is a Hilbert space (e.g. $X = L^2(\mathbb{T})$) and if f_j are pairwise orthogonal (e.g. $f_j(x) = e(xj)$) then we have a stronger inequality (in fact an equality)

$$\|\sum_{j=1}^{N} f_j\|_X \le (\sum_{j=1}^{N} \|f_j\|_X^2)^{1/2}.$$

We will call such an inequality $l^2(X)$ *decoupling*. It is natural to ask if there is something analogous in $L^p(\mathbb{R}^n)$ when $p \neq 2$, in the absence of Hilbert space orthogonality.

The answer is yes. Our first example ($X = L^4[0,1]$) is due to the "bi-orthogonality" of the squares. Note that we lose N^ϵ (and some loss in N is in fact necessary in this case), but this will be acceptable in our definition of decoupling.

Theorem 3.1 (Discrete $l^2(L^4)$ decoupling for squares)**.** *For each* $\epsilon > 0$*, the following decoupling holds*

$$\|\sum_{j=1}^{N} a_j e(j^2 x)\|_{L^4[0,1]} \lesssim_\epsilon N^\epsilon (\sum_{j=1}^{N} \|a_j e(j^2 x)\|_{L^4[0,1]}^2)^{1/2} = N^\epsilon \|a_j\|_{l^2}.$$

Proof. We present the argument in the case $a_j = 1$, the general case requires only minor modifications. By raising to the fourth power, the left hand side equals

$$\int_0^1 \sum_{1\le j_i\le N} e((j_1^2 + j_2^2 - j_3^2 - j_4^2)x)dx = \sum_{1\le j_1, j_2 \le N} |\{(j_3, j_4) :\ j_3^2 + j_4^2 = j_1^2 + j_2^2\}| \lesssim_\epsilon N^{2+\epsilon}.$$

The last inequality follows since the equation

$$j_3^2 + j_4^2 = A$$

has $\lesssim_\epsilon A^\epsilon$ solutions, Grosswald [1985].

□

The second well known example relies on the "multi-orthogonality" of the sequence 2^j.

Theorem 3.2 (Discrete L^p decoupling for lacunary exponential sums)**.** *For* $1 \le p < \infty$ *and* $a_j \in \mathbb{C}$

$$\|\sum_{j=1}^{N} a_j e(2^j x)\|_{L^p[0,1]} \sim_p (\sum_{j=1}^{N} \|a_j e(2^j x)\|_{L^p[0,1]}^2)^{1/2} = \|a_j\|_{l^2}.$$

These easy examples are of arithmetic structure. We will develop tools that do not depend on this restriction. The reader will notice that what lies behind both examples is the fact that there is increasing level of separation between higher frequencies (squares, powers of 2). We will see that in higher dimensions quasi-uniform separation will suffice, as long as the frequencies lie on a curved manifold.

We close this section with one of the most important results in classical harmonic analysis, perhaps in the entire mathematics. It is a consequence and a continuous reformulation of Theorem 3.2.

Theorem 3.3 (Littlewood–Paley theorem). *Given $f : \mathbb{R} \to \mathbb{C}$, let*

$$P_j f(x) = \int_{I_j} \widehat{f}(\xi) e(x\xi) d\xi$$

be its Fourier projection on $I_j = [2^j, 2^{j+1}] \cup [-2^{j+1}, -2^j]$, for $j \in \mathbb{Z}$. Then for each $1 < p < \infty$

$$\|f\|_{L^p(\mathbb{R})} \sim_p \|(\sum_j |P_j f|^2)^{\frac{1}{2}}\|_{L^p(\mathbb{R})}.$$

For our purposes, it suffices to note that, when combined with Minkowski's inequality, the Littlewood–Paley theorem leads to the following $l^2(L^p)$ decoupling on the real line, for $p \geq 2$

$$\|f\|_{L^p(\mathbb{R})} \lesssim (\sum_j \|P_j f\|_{L^p(\mathbb{R})}^2)^{\frac{1}{2}}.$$

4 Fourier analytic decouplings

For a ball (or cube) B_R in $\mathbb{R}^n$ with center c and radius (side length) R, we will denote by $w_{B_R}(x)$ a weight of the form $(1 + \frac{|x-c|}{R})^{-C}$, for some large unspecified C. This can be thought of as being a smooth approximation of 1_{B_R}.

Fix a manifold $\mathfrak{M} = \mathfrak{M}^{\psi}$ as in (1) and let $f : D \to \mathbb{C}$. If we partition D into sets τ then we may write

$$E^{\mathfrak{M}} f = \sum_{\tau} E_{\tau}^{\mathfrak{M}} f.$$

Roughly speaking, $E_{\tau}^{\mathfrak{M}} f(x)$ has the oscillatory phase $e(x \cdot (\xi_\tau, \psi(\xi_\tau)))$, where ξ_τ is a point in τ. If $\mathfrak{M}$ has some curvature, which is the same as saying that ψ is "far" from being affine, then there will be lots of cancellations between the components $E_{\tau}^{\mathfrak{M}} f(x)$. This will be formalized by a Fourier decoupling, which (for now) takes the following rather vague conjectural form.

Conjecture 4.1 (Fourier decoupling). *Let $\mathfrak{M}$ be sufficiently curved. Then there is a critical index $p_c > 2$ and some $q \geq 1$ so that for each partition $\wp_\delta$ of the domain D into N "caps" τ of "size" δ we have*

$$\|E^{\mathfrak{M}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_{\tau\in\wp_\delta} \|E^{\mathfrak{M}}_\tau f\|^2_{L^p(w_{B_R})}\Big)^{1/2} \quad (l^2(L^p)\ \textit{decoupling})$$

or (alternatively)

$$\|E^{\mathfrak{M}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon} N^{\frac{1}{2}-\frac{1}{p}}\Big(\sum_{\tau\in\wp_\delta} \|E^{\mathfrak{M}}_\tau f\|^p_{L^p(w_{B_R})}\Big)^{1/p} \quad (l^p(L^p)\ \textit{decoupling})$$

for each ball B_R with radius $R \geq \delta^{-q}$ and each $2 \leq p \leq p_c$.

The presence of w_{B_R} on the right hand side is probably necessary, but completely harmless for applications.

Note that an l^2 decoupling always implies an l^p decoupling, due to the Cauchy-Schwarz inequality. However, sometimes the former is false and the latter is true. In most applications, an l^p decoupling is as good as an l^2 decoupling.

The shape of the "caps", the precise meaning of "size" as well as the values of q and of the critical exponent p_c depend on the manifold $\mathfrak{M}$. Due to orthogonality considerations, there is always a decoupling for $p = 2$, even for flat manifolds (hyperplanes). In this latter case however, considering $f \equiv 1$ shows that there is no decoupling outside L^2, so $p_c = 2$.

The formulation of Conjecture 4.1 is vague in many ways. There are interesting examples of manifolds that are known to simultaneously host different decoupling phenomena, corresponding to different values of p_c, q and for different types of caps. While, as observed in Section 2, restriction inequalities are associated with the uncertainty principle scale ($q = 1$), the most genuine decouplings will happen at spatial scales of magnitude $q \geq 2$.

The first to consider a Fourier decoupling was Wolff [2000]. He proved an $l^p(L^p)$ decoupling for the cone $\mathbb{C}o^2$ when $p > 74$, by masterfully combining Fourier analytic and incidence geometric arguments. In his theorem the caps are thin annular sectors of dimensions $\sim \delta$ and 1. Wolff also showed that his decoupling has consequences for the local smoothing of the solutions to the wave equation. Subsequent developments for the higher dimensional cone and other manifolds, prior to the work we are about the describe here, have appeared in Łaba and Wolff [2002], Laba and Pramanik [2005], Garrigós and Seeger [2009], Garrigós and Seeger [2010], Pramanik and Seeger [2007], Bourgain [2013] and Demeter [n.d.].

The first full range results for any manifold came in our joint work Bourgain and Demeter [2015] with Jean Bourgain.

Theorem 4.2. *Assume $\mathfrak{M}$ has positive definite second fundamental form (e.g. $\mathbb{S}^{n-1}$, $\mathbb{P}^{n-1}$). For any partition of the domain D into square-like caps τ with diameter δ we have*

$$\|E^{\mathfrak{M}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_\tau \|E^{\mathfrak{M}}_\tau f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

as long as $R \geq \delta^{-2}$ and $2 \leq p \leq \frac{2(n+1)}{n-1}$.

The proof of this for $\mathbb{P}^1$ will be presented in Section 6. Quite surprisingly, we were able to use this result for $\mathbb{P}^{n-1}$ as a black box, in order to derive the sharp result for the cone in all dimensions, thus closing the program initiated by Wolff. Let us get a glimpse into our argument for $\mathbb{C}o^2$. After a rotation, the equation of $\mathbb{C}o^2$ can be rewritten as $z = \frac{x^2}{y}$. The y-slices are parabolas with roughly the same curvature. This forces small pieces of the cone to be close to parabolic cylinders. One may combine the decoupling for the parabola with Fubini in the zero curvature direction to gradually separate the cone into smaller pieces. This argument is very different from Wolff's, in that it does not require any incidence geometry.

Theorem 4.3 (Bourgain and Demeter [2015]). *Let $\mathfrak{M} = \mathbb{C}o^{n-1}$ be the cone. For any partition of the domain $D = \{|\xi| \sim 1\}$ into sectors τ with angular width δ we have*

$$\|E^{\mathbb{C}o^{n-1}} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_\tau \|E^{\mathbb{C}o^{n-1}}_\tau f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

as long as $R \geq \delta^{-2}$ and $2 \leq p \leq \frac{2n}{n-2}$. The range for p is sharp.

Another milestone of decoupling theory was the resolution of curves with torsion, in collaboration with Bourgain and Guth. More precisely, consider $\Phi : [0,1] \to \mathbb{R}^n$,

$$\Phi(\xi) = (\phi_1(\xi), \dots, \phi_n(\xi))$$

with $\phi_i \in C^n([0,1])$ and such that the Wronskian $W(\phi_1', \dots, \phi_n')(\xi)$ is nonzero on $[0,1]$. One example is the moment curve Γ_n. Let E^Φ be the associated extension operator.

Theorem 4.4 (Bourgain, Demeter, and Guth [2016]). *Partition $[0,1]$ into intervals τ of length $\sim \delta$. Then*

$$\|E^{\Phi} f\|_{L^p(B_R)} \lesssim_\epsilon \delta^{-\epsilon}\Big(\sum_\tau \|E^{\Phi}_\tau f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

as long as $R \geq \delta^{-n}$ and $2 \leq p \leq n(n+1)$. The range for p is sharp.

The proof for $\Gamma_2 = \mathbb{P}^1$ appeared in Bourgain and Demeter [2015], while for $n \geq 3$ in Bourgain, Demeter, and Guth [2016]. The extension to the arbitrary Φ with torsion is explained in Section 4 of Bourgain and Demeter [2017].

Decouplings for a wide variety of other manifolds have been proved in Bourgain and Demeter [ibid.], Bourgain and Demeter [2016a], Bourgain and Demeter [2016b], Bourgain [2017], Bourgain, Demeter, and Guo [2017], Demeter, Guo, and Shi [n.d.], Bourgain and Watt [2017], Bourgain and Watt [n.d.], Guo and Oh [n.d.], and the list is rapidly growing.

5 Applications: Exponential sums on large balls

It turns out that there is a very simple mechanism that allows decouplings to imply exponential sum estimates that are often sharp. Essentially, one applies the decoupling to a weighted combination of (approximations of) Dirac deltas. In this regard each decoupling seems to be stronger than the exponential sum estimate it implies, the author is not aware of any argument that reverses the implication.

Theorem 5.1. *Let $\mathfrak{M} = \mathfrak{M}^{\psi}$. Consider a partition $\mathcal{P}_\delta$ as in Conjecture 4.1, with $N = |\mathcal{P}_\delta|$. Let $\xi_\tau \in \tau$ for each $\tau \in \mathcal{P}_\delta$ and let $\lambda_\tau = (\xi_\tau, \psi(\xi_\tau))$ be the corresponding point on $\mathfrak{M}$. Then for each $2 \leq p \leq p_c$, $a_\tau \in \mathbb{C}$ and each $R \geq \delta^{-q}$ we have*

$$\|\sum_{\tau \in \mathcal{P}_\delta} a_\tau e(\lambda_\tau \cdot x)\|_{L^p_\sharp(B_R)} \lesssim_\epsilon \delta^{-\epsilon} \|a_\tau\|_{l^2},$$

if the l^2 version of the decoupling in Conjecture 4.1 holds true, and

$$\|\sum_{\tau \in \mathcal{P}_\delta} a_\tau e(\lambda_\tau \cdot x)\|_{L^p_\sharp(B_R)} \lesssim_\epsilon \delta^{-\epsilon} N^{\frac{1}{2}-\frac{1}{p}} \|a_\tau\|_{l^p},$$

if the l^p version of the conjecture holds instead.

Proof. Apply the conjecture to functions of the form $f(\xi) = \sum_{\tau \in \mathcal{P}_\delta} a_\tau 1_{B(\xi_\tau, r)}(\xi)$ and let $r \to 0$. The computation is straightforward. □

One notable feature of this estimate is that it does not assume anything else about ξ_τ other than the separation guaranteed by the pairwise disjointness of the caps. In particular, the points need not belong to a rescaled lattice. This is indicative of the fact that our methods do not involve number theory, and that in fact sometimes they transcend the barrier that is currently accessible using number theoretic methods.

Let us now consider a few particular cases of interest.

5.1 Stricharz estimates. An application of Theorem 5.1 to $\mathbb{S}^{n-1}$ and $\mathbb{P}^{n-1}$ leads to the following corollary.

Corollary 5.2. *For each $R \geq 1$, each collection Λ consisting of $\frac{1}{R}$-separated points on either $\mathbb{S}^{n-1}$ or $\mathbb{P}^{n-1}$ and each ball $B_{R'}$ of radius $R' \geq R^2$ in $\mathbb{R}^n$ we have*

$$\|\sum_{\lambda\in\Lambda} a_\lambda e(\lambda\cdot x)\|_{L_\sharp^{\frac{2(n+1)}{n-1}}(B_{R'})} \lesssim_\epsilon R^\epsilon \|a_\lambda\|_{l^2}. \tag{4}$$

Comparing this with (3) shows that large ball averages get smaller.

This corollary leads to sharp Strichartz estimates in the periodic and quasi-periodic case. More precisely, fix $\frac{1}{2} < \theta_1, \ldots, \theta_{n-1} < 2$ either rational or irrational. For $\phi \in L^2(\mathbb{T}^{n-1})$ consider its Laplacian

$$\Delta\phi(x_1, \ldots, x_{n-1}) =$$

$$\sum_{(\xi_1,\ldots,\xi_{n-1})\in\mathbb{Z}^{n-1}} (\xi_1^2\theta_1 + \ldots + \xi_{n-1}^2\theta_{n-1})\hat{\phi}(\xi_1, \ldots, \xi_{n-1})e(\xi_1x_1 + \ldots + \xi_{n-1}x_{n-1})$$

on the torus $\prod_{i=1}^{n-1}\mathbb{R}/(\theta_i\mathbb{Z})$. Let also

$$e^{it\Delta}\phi(x_1, \ldots, x_{n-1}, t) =$$

$$\sum_{(\xi_1,\ldots,\xi_{n-1})\in\mathbb{Z}^{n-1}} \hat{\phi}(\xi_1, \ldots, \xi_{n-1})e(x_1\xi_1 + \ldots + x_{n-1}\xi_{n-1} + t(\xi_1^2\theta_1 + \ldots + \xi_{n-1}^2\theta_{n-1}))$$

be the solution of the Schrödinger equation in this context. We have the following result. When $p > \frac{2(n+1)}{n-1}$, the N^ϵ loss can be removed, see Bourgain and Demeter [2015] and Killip and Vişan [2016].

Theorem 5.3 (Strichartz estimates for rational and irrational tori, Bourgain and Demeter [2015]). *Let $\phi \in L^2(\mathbb{T}^{n-1})$ with $\operatorname{supp}(\hat{\phi}) \subset [-N, N]^{n-1}$. Then for each $\epsilon > 0$ and $p \geq \frac{2(n+1)}{n-1}$ we have*

$$\|e^{it\Delta}\phi\|_{L^p(\mathbb{T}^{n-1}\times[0,1])} \lesssim_\epsilon N^{\frac{n-1}{2}-\frac{n+1}{p}+\epsilon}\|\phi\|_2, \tag{5}$$

and the implicit constant does not depend on N.

Proof. It suffices to consider the case $p = \frac{2(n+1)}{n-1}$. For $-N \le \xi_1, \ldots, \xi_{n-1} \le N$ define $\eta_i = \frac{\theta_i^{1/2}\xi_i}{4N}$ and $a_\eta = \hat{\phi}(\xi)$. A simple change of variables shows that

$$(6)\quad \int_{\mathbb{T}^{n-1}\times[0,1]} |e^{it\Delta}\phi|^p \lesssim \frac{1}{N^{n+1}} \int_{\substack{|y_1|,\ldots,|y_{n-1}|\le 8N \\ y_n \in I_{N^2}}} \left| \sum_{\eta_1,\ldots,\eta_{n-1}} a_\eta e(y_1\eta_1 + \ldots + y_{n-1}\eta_{n-1} + y_n(\eta_1^2 + \ldots \eta_{n-1}^2)) \right|^p dy_1 \ldots dy_n$$

where I_{N^2} is an interval of length $\sim N^2$. By periodicity in the $y_1, \ldots, y_{n-1}$ variables we bound the above by

$$\frac{1}{N^{n+1}N^{n-1}} \int_{B_{N^2}} \left| \sum_{\eta_1,\ldots,\eta_{n-1}} a_\eta e(y_1\eta_1 + \ldots + y_{n-1}\eta_{n-1} + y_n(\eta_1^2 + \ldots \eta_{n-1}^2)) \right|^p dy_1 \ldots dy_n,$$

for some ball B_{N^2} of radius $\sim N^2$. Our result will follow once we note that the points

$$(\eta_1, \ldots, \eta_{n-1}, \eta_1^2 + \ldots \eta_{n-1}^2)$$

are $\sim \frac{1}{N}$ separated on $\mathbb{P}^{n-1}$ and then apply Corollary 5.2 with $R' \sim N^2$.

□

5.2 Diophantine inequalities and the Vinogradov Mean Value Theorem.

An application of Theorem 5.1 to the moment curve gives the following exponential sum estimate.

Corollary 5.4. *For each $1 \le i \le N$, let t_i be a point in $(\frac{i-1}{N}, \frac{i}{N}]$. Then for each $R \gtrsim N^n$ and each $p \ge 2$ we have*

$$\left(\frac{1}{|B_R|} \int |\sum_{i=1}^N a_i e(x_1 t_i + x_2 t_i^2 + \ldots + x_n t_i^n)|^p w_{B_R}(x) dx_1 \ldots dx_n\right)^{\frac{1}{p}} \lesssim$$

$$(7)\qquad \lesssim_\epsilon (N^\epsilon + N^{\frac{1}{2}(1-\frac{n(n+1)}{p})+\epsilon}) \|a_i\|_{l^2(\{1,\ldots,N\})},$$

and the implicit constant does not depend on N, R and a_i.

For each $1 \le i \le N$ consider some real numbers $i-1 < X_i \le i$. We do not insist that X_i are integers. Let $S_X = \{X_1, \dots, X_N\}$. For each $s \ge 1$, denote by $J_{s,n}(S_X)$ the number of solutions of the following system of inequalities

$$|X_1^i + \dots + X_s^i - (X_{s+1}^i + \dots + X_{2s}^i)| \le N^{i-n}, \quad 1 \le i \le n \tag{8}$$

with $X_i \in S_X$.

Corollary 5.5. *For each integer $s \ge 1$ and each S_X as above we have that*

$$J_{s,n}(S_X) \lesssim_\epsilon N^{s+\epsilon} + N^{2s-\frac{n(n+1)}{2}+\epsilon},$$

where the implicit constant does not depend on S_X.

Proof. Let $\phi : \mathbb{R}^n \to [0,\infty)$ be a positive Schwartz function with positive Fourier transform satisfying $\widehat{\phi}(\xi) \ge 1$ for $|\xi| \lesssim 1$. Define $\phi_N(x) = \phi(\frac{x}{N})$. Using the Schwartz decay, (7) with $a_i = 1$ implies that for each $s \ge 1$

$$\Big(\frac{1}{|B_{N^n}|}\int_{\mathbb{R}^n} \phi_{N^n}(x)\Big|\sum_{i=1}^N e(x_1t_i + \dots + x_nt_i^n)\Big|^{2s} dx_1 \dots dx_n\Big)^{\frac{1}{2s}}$$

$$\lesssim_\epsilon N^{\frac{1}{2}+\epsilon} + N^{1-\frac{n(n+1)}{4s}+\epsilon}, \tag{9}$$

whenever $t_i \in [\frac{i-1}{N}, \frac{i}{N})$. Apply (9) to $t_i = \frac{X_i}{N}$. Let now

$$\phi_{N,1}(x_1, x_2, \dots, x_n) = \phi\Big(\frac{x_1}{N^{n-1}}, \frac{x_2}{N^{n-2}}, \dots, x_n\Big).$$

It suffices to consider the case $s = \frac{n(n+1)}{2}$. After making a change of variables and expanding the product, the term

$$\int_{\mathbb{R}^n} \phi_{N^n}(x)\Big|\sum_{i=1}^N e(x_1t_i + \dots + x_nt_i^n)\Big|^{2s} dx_1 \dots dx_n$$

can be written as the sum over all $X_i \in S_X$ of

$$N^{\frac{n(n+1)}{2}} \int_{\mathbb{R}^n} \phi_{N,1}(x) e(x_1Z_1 + \dots + x_nZ_n) dx_1 \dots dx_n,$$

where

$$Z_i = X_1^i + \dots + X_s^i - (X_{s+1}^i + \dots + X_{2s}^i).$$

Each such term is equal to

$$N^{n^2}\widehat{\phi}(N^{n-1}Z_1, N^{n-2}Z_2, \ldots, Z_n).$$

Recall that this is always positive, and in fact greater than N^{n^2} at least $J_{s,n}(S_X)$ times. It now suffices to use (9).

□

The special case of Corollary 5.5 when $X_i = i$ and the inequalities (8) are replaced with equalities

$$X_1^i + \ldots + X_s^i = X_{s+1}^i + \ldots + X_{2s}^i, \quad 1 \le i \le n$$

was known as the Main Conjecture in the Vinogradov Mean Value Theorem. The case $n = 2$ is very easy, while the case $n = 3$ was only recently proved by Wooley using the efficient congruencing method, Wooley [2016]. The case $n \ge 4$ was proved for the first time in Bourgain, Demeter, and Guth [2016].

6 The proof of the decoupling theorem for the parabola

In this section we prove Theorem 4.2 for the parabola $\mathbb{P}^1$. The argument in higher dimensions is very similar, though technically slightly more complicated. We will denote by E the extension operator associated with $\mathbb{P}^1$.

It will be more convenient to think of B_R as being an arbitrary square (rather than ball) with side length R in $\mathbb{R}^2$. We will often partition big squares into smaller ones.

For $m \ge 0$ let $\mathbb{I}_m$ be the collection of the 2^m dyadic subintervals of $[0,1]$ of length 2^{-m}. Thus $\mathbb{I}_0$ consists of only $[0,1]$. Note that each $I \in \mathbb{I}_{m+1}$ is inside some $I' \in \mathbb{I}_m$ and each $I' \in \mathbb{I}_m$ has two "children" in $\mathbb{I}_{m+1}$, adjacent to each other.

Define $\mathrm{Dec}(n,p)$ to be the smallest constant such that the inequality

$$\|Ef\|_{L^p(w_{B_{4^n}})} \le \mathrm{Dec}(n,p)\Big(\sum_{I\in\mathbb{I}_n} \|E_I f\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/2}$$

holds true for each $f : [0,1] \to \mathbb{C}$. Minkowski's inequality shows that $\mathrm{Dec}(n,p)$ controls the decoupling on larger squares, too. By that we mean that the inequality

$$\|Ef\|_{L^p(w_{B_R})} \lesssim \mathrm{Dec}(n,p)\Big(\sum_{I\in\mathbb{I}_n} \|E_I f\|^2_{L^p(w_{B_R})}\Big)^{1/2}$$

holds for each $R \ge 4^n$, with the implicit constant in $\lesssim$ independent of R.

It will suffice to prove that

$$\mathrm{Dec}(n,p) \lesssim_{\epsilon,p} 2^{n\epsilon}, \quad 2 \le p \le 6. \tag{10}$$

The inequality $\mathrm{Dec}(n,2)\lesssim 1$ follows from simple orthogonality reasons. We start by explaining why the proof of (10) is not quite as immediate as one would wish, when $p>2$. Let A_p be the smallest constant that governs the decoupling into two intervals. In precise terms, assume

$$\|Ef\|_{L^p(B)}\le A_p(\|E_{J_1}f\|^2_{L^p(w_B)}+\|E_{J_2}f\|^2_{L^p(w_B)})^{1/2}$$

holds for each disjoint intervals $J_1,J_2\subset\mathbb{R}$ of arbitrary length $L\le 1$ that are adjacent to each other, each $f:J_1\cup J_2\to\mathbb{C}$ and for each square $B\subset\mathbb{R}^2$ with side length at least L^{-2}. One can check that $A_p>1$ for $p>2$.

Since each $I'\in\mathbb{I}_m$ has two children in $\mathbb{I}_{m+1}$, it is easy to see that

$$\mathrm{Dec}(m+1,p)\le A_p\mathrm{Dec}(m,p).$$

Iterating this, we get the very unfavorable estimate $\mathrm{Dec}(n,p)\le A_p^{n-1}\mathrm{Dec}(1,p)$. Indeed, note that $A_p>1$ forces $A_p^{n-1}\gg 2^{n\epsilon_0}$, for some $\epsilon_0>0$. This shows that we can not afford to lose A_p each time we go one level up (call this a *step*). Instead, we are going to make huge *leaps*, and rather than going from level m to level $m+1$ at a time, we will instead go from m to $2m$. The choice for the size of this leap is motivated by the fact that intervals in $\mathbb{I}_{2m}$ have length equal to those in $\mathbb{I}_m$ squared. We have a very efficient mechanism to decouple from scale δ to δ^2, namely the bilinear Kakeya inequality.

Each leap will combine two inequalities. One is a consequence of the bilinear Kakeya, the other one is a form of L^2 orthogonality. The loss for each application of the bilinear Kakeya is rather tiny, at most n^C (compare this with the loss A_p^m accumulated if instead we went from level m to $2m$ in m steps). From $\mathbb{I}_0$ to $\mathbb{I}_n$, we need $\log n$ such leaps, so the overall loss from the repeated use of bilinear Kakeya amounts to $n^{O(\log n)}$. This is easily seen to be $O(2^{n\epsilon})$ for each $\epsilon>0$, as desired. There is however a price we pay in our approach: in each leap we only decouple a $1-\kappa_p$ fraction of the operator. See Proposition 6.5 for a precise statement.

Here is a sketch of how we put things together, and we will limit attention to the hardest case $p=6$. First, we will do a trivial decoupling (Cauchy-Schwarz) to get from $\mathbb{I}_0$ to $\mathbb{I}_{\frac{n}{2^s}}$ by loosing only $2^{O(\frac{n}{2^s})}$. We will be able to choose s as large as we wish, so this loss will end up being controlled by $2^{n\epsilon}$. The transition from $\mathbb{I}_{\frac{n}{2^s}}$ to $\mathbb{I}_n$ will then be done in s leaps, by each time applying Proposition 6.5. Collecting all contributions, an a priori bound of the form

$$\mathrm{Dec}(n,6)\lesssim 2^{nA},\ \text{for some } A>0$$

will get upgraded to a stronger (assuming s is large enough) bound

$$\mathrm{Dec}(n,6)\lesssim 2^{n(A(1-\frac{s+1}{2^{s+1}})+\frac{1}{2^s-1})}.$$

Applying this bootstrapping argument will force A to get smaller and smaller, arbitrarily close to 0.

The leaps are performed using bilinear decouplings, in order to take advantage of the bilinear Kakeya phenomenon. The fact that there is no serious loss in bilinearization is proved in Proposition 6.2.

6.1 Parabolic rescaling and linear vs. bilinear decoupling. One of our main tools will be the following parabolic rescaling, that takes advantage of the affine invariance of the parabola.

Proposition 6.1. *Let $I = [t, t + 2^{-l}] \subset \mathbb{R}$ be an interval of length 2^{-l} and for $n > l$ let the collection $\mathbb{I}_n(I)$ consist of all subintervals of I of the form $[t + j2^{-n}, t + (j+1)2^{-n}]$, with $j \in \mathbb{N}$. Then for each f supported on I*

$$\|Ef\|_{L^p(w_{B_{4^n}})} \lesssim \mathrm{Dec}(n-l, p)\Big(\sum_{J \in \mathbb{I}_n(I)} \|E_J f\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/2}$$

Note that the upper bound $\mathrm{Dec}(n-l, p)$ is morally stronger than the trivial upper bound $\mathrm{Dec}(n, p)$.

Proof. The proof is a simple applications of affine change of variables. Indeed $L_I(\xi) = 2^l(\xi - t)$ maps $\mathbb{I}_n(I)$ to $\mathbb{I}_{n-l}$ and the square B_{4^n} to a parallelepiped that can be covered efficiently with squares $B_{4^{n-l}}$.

□

Define $\mathrm{BilDec}(n, p)$ to be the smallest constant such that the inequality

$$\||Ef_1 Ef_2|^{1/2}\|_{L^p(w_{B_{4^n}})} \le$$
$$\le \mathrm{BilDec}(n, p)\Big(\sum_{I \in \mathbb{I}_n(I_1)} \|E_I f_1\|^2_{L^p(w_{B_{4^n}})} \sum_{I \in \mathbb{I}_n(I_2)} \|E_I f_2\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/4}$$

holds true for all f_1, f_2 supported on I_1 and I_2, respectively.

It is immediate that

$$\mathrm{BilDec}(n, p) \le \mathrm{Dec}(n, p).$$

The next result is some sort of a converse.

Proposition 6.2. *For each $\epsilon > 0$*

$$\mathrm{Dec}(n, p) \lesssim_\epsilon 2^{n\epsilon}\big(1 + \max_{m \le n} \mathrm{BilDec}(m, p)\big).$$

Proof. It will suffice to prove that for each $k < n$

$$\mathrm{Dec}(n,p) \lesssim C^{\frac{n}{k}}(1 + C_k n \max_{m \le n} \mathrm{BilDec}(m,p)).$$

This will instead follow by iterating the inequality

$$\mathrm{Dec}(n,p) \le C\,\mathrm{Dec}(n-k,p) + C_k \max_{m \le n} \mathrm{BilDec}(m,p), \tag{11}$$

with C independent of n, k. Let us next prove this inequality.

Fix k and let f be supported on $[0,1]$. Since

$$Ef(x) = \sum_{I \in \mathbb{I}_k} E_I f(x),$$

it is not difficult to see that

$$|Ef(x)| \le 4 \max_{I \in \mathbb{I}_k} |E_I f(x)| + 2^{O(k)} \sum_{\substack{J_1, J_2 \in \mathbb{I}_k \\ 2J_1 \cap 2J_2 = \emptyset}} |E_{J_1} f(x) E_{J_2} f(x)|^{1/2}, \tag{12}$$

where the sum on the right is taken over all pairs of intervals $J_1, J_2 \in \mathbb{I}_k$ which are not neighbors. Fix such a pair $J_1 = [a, a + 2^{-k}]$, $J_2 = [b, b + 2^{-k}]$, and let m be a positive integer satisfying $2^{-m} \le b - a < 2^{-m+1}$. Since J_1, J_2 are not adjacent to each other, we must have $m \le k - 1$. It follows that the affine function $T(\xi) = \frac{\xi - a}{2^{-m+1}}$ maps J_1 to a dyadic subinterval of $[0, \frac{1}{4}]$ and J_2 to a dyadic subinterval of $[\frac{1}{2}, 1]$. Thus, parabolic rescaling shows that

$$\||E_{J_1} f E_{J_2} f|^{1/2}\|_{L^p(w_{B_{4^n}})}$$
$$\lesssim \mathrm{BilDec}(n-m+1,p)\Big(\sum_{I \in \mathbb{I}_n(J_1)} \|E_I f\|^2_{L^p(w_{B_{4^n}})} \sum_{I \in \mathbb{I}_n(J_2)} \|E_I f\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/4}$$
$$\le \mathrm{BilDec}(n-m+1,p)\Big(\sum_{I \in \mathbb{I}_n} \|E_I f\|^2_{L^p(w_{B_{4^n}})}\Big)^{1/2}. \tag{13}$$

Finally, invoking again Proposition 6.1 we get

$$\|\max_{I \in \mathbb{I}_k} |E_I f|\|_{L^p(w_{B_{4^n}})} \le \Big(\sum_{I \in \mathbb{I}_k} \|E_I f\|^2_{L^p(w_{B_{4^n}})}\Big)^{\frac{1}{2}}$$
$$\lesssim \mathrm{Dec}(n-k,p)\Big(\sum_{I \in \mathbb{I}_k} \sum_{I' \in \mathbb{I}_n(I)} \|E_{I'} f\|^2_{L^p(w_{B_{4^n}})}\Big)^{\frac{1}{2}}$$
$$= \mathrm{Dec}(n-k,p)\Big(\sum_{I' \in \mathbb{I}_n} \|E_{I'} f\|^2_{L^p(w_{B_{4^n}})}\Big)^{\frac{1}{2}}. \tag{14}$$

Now (11) follows by combining (12), (13) and (14).

□

6.2 A consequence of bilinear Kakeya. We start by recalling the following bilinear Kakeya inequality. While this inequality is rather trivial in two dimensions, its higher dimensional analogs that are needed in order to prove decouplings for the paraboloid $\mathbb{P}^{n-1}$, $n \geq 3$, are more complicated. The multilinear Kakeya inequality was first proved in Bennett, Carbery, and Tao [2006], and an easier proof appeared in Guth [2015].

Theorem 6.3. *Consider two families $\mathbb{T}_1, \mathbb{T}_2$ consisting of rectangles T in $\mathbb{R}^2$ having the following properties*

(i) each T has the short side of length $R^{1/2}$ and the long side of length equal to R pointing in the direction of the unit vector v_T

(ii) $v_{T_1} \wedge v_{T_2} \geq \frac{1}{100}$ for each $T_i \in \mathbb{T}_i$.

We have the following inequality

$$\int_{\mathbb{R}^2} \prod_{i=1}^{2} F_i \lesssim \frac{1}{R^2} \prod_{i=1}^{2} \int_{\mathbb{R}^2} F_i \tag{15}$$

for all functions F_i of the form

$$F_i = \sum_{T \in \mathbb{T}_i} c_T 1_T, \ c_T \in [0, \infty).$$

The implicit constant will not depend on $R, c_T, \mathbb{T}_i$.

Proof. The verification is immediate using the fact that $|T_1 \cap T_2| \lesssim R$ whenever $T_i \in \mathbb{T}_i$. □

If $I \subset \mathbb{R}$ is an interval of length 2^{-l} and $\delta = 2^{-k}$ with $k \geq l$, we will denote by $\text{Part}_\delta(I)$ the partition of I into intervals of length δ. Recall also that $I_1 = [0, \frac{1}{4}]$, $I_2 = [\frac{1}{2}, 1]$ and that $L^q_\sharp$ denotes the average integral in L^q.

The following result is part of a two-stage process. Note that, strictly speaking, this inequality is not a decoupling, since the size of the frequency intervals $I_{i,1}$ remains unchanged. However, the side length of the spatial squares increases from δ^{-1} to δ^{-2}. This will facilitate a subsequent decoupling, as we shall later see in Proposition 6.5.

Proposition 6.4. *Let $q \geq 2$ and $\delta < 1$. Let B be an arbitrary square in $\mathbb{R}^2$ with side length δ^{-2}, and let $\mathfrak{B}$ be the unique partition of B into squares Δ of side length δ^{-1}. Then* *for each* $g : [0,1] \to \mathbb{C}$ *we have*

$$\frac{1}{|\mathfrak{B}|}\sum_{\Delta\in\mathfrak{B}}\left[\prod_{i=1}^{2}\Bigg(\sum_{I_{i,1}\in\mathrm{Part}_\delta(I_i)}\|E_{I_{i,1}}g\|^2_{L^q_\sharp(w_\Delta)}\Bigg)^{\frac12}\right]^q \lesssim \left(\log\left(\frac{1}{\delta}\right)\right)^{O(1)}\left[\prod_{i=1}^{2}\Bigg(\sum_{I_{i,1}\in\mathrm{Part}_\delta(I_i)}\|E_{I_{i,1}}g\|^2_{L^q_\sharp(w_B)}\Bigg)^{\frac12}\right]^q. \tag{16}$$

Moreover, the implicit constant is independent of g, δ, B.

Proof. We will reduce the proof to an application of Theorem 6.3. Indeed, for each interval J of length δ, the Fourier transform of $E_J g$ is supported inside a $2\delta\times 2\delta^2$-rectangle. This in turn suggests that $|E_J g|$ is essentially constant on $\delta^{-1}\times\delta^{-2}$-rectangles dual to this rectangle. Note that due to the separation of I_1 and I_2, the rectangles corresponding to intervals $I_{1,1}\subset I_1$, $I_{2,1}\subset I_2$ satisfy the requirements in Theorem 6.3 with $R=\delta^{-2}$.

Since we can afford logarithmic losses in δ, it suffices to prove the inequality with the summation on both sides restricted to families of $I_{i,1}$ for which $\|E_{I_{i,1}}g\|_{L^q_\sharp(w_B)}$ have comparable size (within a multiplicative factor of 2), for each i. Indeed, the intervals $I'_{i,1}$ satisfying (for some large enough $C=O(1)$)

$$\|E_{I'_{i,1}}g\|_{L^q_\sharp(w_B)}\le\delta^C\max_{I_{i,1}\in\mathrm{Part}_\delta(I_i)}\|E_{I_{i,1}}g\|_{L^q_\sharp(w_B)}$$

can be easily dealt with by using the triangle inequality, since we automatically have

$$\max_{\Delta\in\mathfrak{B}}\|E_{I'_{i,1}}g\|_{L^q_\sharp(w_\Delta)}\le\delta^C\max_{I_{i,1}\in\mathrm{Part}_\delta(I_i)}\|E_{I_{i,1}}g\|_{L^q_\sharp(w_B)}.$$

This leaves only $\log_2(\delta^{-O(1)})$ sizes to consider.

Let us now assume that we have N_i intervals $I_{i,1}$, with $\|E_{I_{i,1}}g\|_{L^q_\sharp(w_B)}$ of comparable size. Since $q\ge 2$, by Hölder's inequality (16) is at most

$$\left(\prod_{i=1}^{2}N_i^{\frac12-\frac1q}\right)^q\frac{1}{|\mathfrak{B}|}\sum_{\Delta\in\mathfrak{B}}\left(\prod_{i=1}^{2}\left(\sum_{I_{i,1}}\|E_{I_{i,1}}g\|^q_{L^q_\sharp(w_\Delta)}\right)\right). \tag{17}$$

For each $I=I_{i,1}$ centered at c_I, consider the family $\mathfrak{F}_I$ of pairwise disjoint, mutually parallel rectangles T_I. They have the short side of length δ^{-1} and the longer side of length δ^{-2}, pointing in the direction of the normal $N(c_I)$ to the paraboloid $\mathbb{P}^1$ at c_I.

The function

$$F_I(x) := \|E_I g\|^q_{L^q_\sharp(w_{B(x,\delta^{-1})})}$$

can be thought of as being essentially constant on rectangles in $\mathfrak{F}_I$. This can be made precise, but we will sacrifice a bit of the rigor for the sake of keeping the argument simple enough. Thus we may write

$$\frac{1}{|\mathcal{B}|}\sum_{\Delta\in\mathcal{B}}\prod_i\left(\sum_{I_{i,1}}\|E_{I_{i,1}}g\|^q_{L^q_\sharp(w_\Delta)}\right) \approx \frac{1}{|B|}\int\prod_{i=1}^2 F_i,$$

with $F_i(x) = \sum_{I_{i,1}} F_{I_{i,1}}(x)$.

Applying Theorem 6.3 we can dominate the term on the right by

$$\frac{1}{|B|^2}\prod_i\int F_i.$$

Note also that

$$\frac{1}{|B|}\int F_i \approx \sum_{I_{i,1}}\|E_{I_{i,1}}g\|^q_{L^q_\sharp(w_B)}.$$

It follows that (17) is dominated by

$$\left(\prod_{i=1}^2 N_i^{\frac{1}{2}-\frac{1}{q}}\right)^q\prod_{i=1}^2\left(\sum_{I_{i,1}}\|E_{I_{i,1}}g\|^q_{L^q_\sharp(w_B)}\right). \tag{18}$$

Recalling the restriction we have made on $I_{i,1}$, (18) is comparable to

$$\left[\prod_{i=1}^2\left(\sum_{I_{i,1}}\|E_{I_{i,1}}g\|^2_{L^q_\sharp(w_B)}\right)^{1/2}\right]^q,$$

as desired. □

6.3 The leap: decoupling from scale δ to δ^2.

To simplify notation, we denote by B^r an arbitrary square in $\mathbb{R}^2$ with side length 2^r. Given $q, r \in \mathbb{N}$, $t \geq 2$ and g supported in $I_1 \cup I_2$ write

$$D_t(q, B^r, g) = \left[\left(\sum_{I\in\mathbb{I}_q(I_1)}\|E_I g\|^2_{L^t_\sharp(w_{B^r})}\right)\left(\sum_{I\in\mathbb{I}_q(I_2)}\|E_I g\|^2_{L^t_\sharp(w_{B^r})}\right)\right]^{\frac{1}{4}}.$$

For $s \le r$ we will denote by $\mathcal{B}_s(B^r)$ the partition of B^r into squares B^s. Define

$$A_p(q, B^r, s, g) = \left(\frac{1}{|\mathcal{B}_s(B^r)|} \sum_{B^s \in \mathcal{B}_s(B^r)} D_2(q, B^s, g)^p \right)^{\frac{1}{p}}.$$

Note that when $r = s$,

$$A_p(q, B^r, r, g) = D_2(q, B^r, g).$$

For $p \ge 4$, let $0 \le \kappa_p \le 1$ satisfy

$$\frac{2}{p} = \frac{1-\kappa_p}{2} + \frac{\kappa_p}{p},$$

that is

$$\kappa_p = \frac{p-4}{p-2}.$$

The following result shows how to decouple from scale $\delta = 2^{-q}$ to scale δ^2. Note also that only a $1-\kappa_p$ fraction gets decoupled.

Proposition 6.5. *We have for each $p \ge 4$, $r \ge q$ and each g supported in $I_1 \cup I_2$*

$$A_p(q, B^{2r}, q, g) \lesssim_\epsilon \delta^{-\epsilon} A_p(2q, B^{2r}, 2q, g)^{1-\kappa_p} D_p(q, B^{2r})^{\kappa_p}$$

Proof. By using elementary inequalities, it suffices to prove the proposition for $r = q$. By Hölder's inequality,

$$\|E_I g\|_{L^2_\sharp(w_{B^q})} \lesssim \|E_I g\|_{L^{\frac{p}{2}}_\sharp(w_{B^q})}.$$

Using this and Proposition 6.4 with $\delta = 2^{-q}$ we can write

$$A_p(q, B^{2q}, q, g) \lesssim_\epsilon \delta^{-\epsilon} \left[\left(\sum_{I \in \mathbb{I}_q(I_1)} \|E_I g\|^2_{L^{\frac{p}{2}}_\sharp(w_{B^{2q}})} \right) \left(\sum_{I \in \mathbb{I}_q(I_2)} \|E_I g\|^2_{L^{\frac{p}{2}}_\sharp(w_{B^{2q}})} \right) \right]^{\frac{1}{4}}. \tag{19}$$

Using Hölder's inequality again, we can dominate this by

$$\delta^{-\epsilon} \left[\left(\sum_{I \in \mathbb{I}_q(I_1)} \|E_I g\|^2_{L^2_\sharp(w_{B^{2q}})} \right) \left(\sum_{I \in \mathbb{I}_q(I_2)} \|E_I g\|^2_{L^2_\sharp(w_{B^{2q}})} \right) \right]^{\frac{1-\kappa_p}{4}}$$
$$\times \left[\left(\sum_{I \in \mathbb{I}_q(I_1)} \|E_I g\|^2_{L^p_\sharp(w_{B^{2q}})} \right) \left(\sum_{I \in \mathbb{I}_q(I_2)} \|E_I g\|^2_{L^p_\sharp(w_{B^{2q}})} \right) \right]^{\frac{\kappa_p}{4}}.$$

To further process the first term we invoke L^2 orthogonality for each $I \in \mathbb{I}_q$

$$\|E_I g\|^2_{L^2_\sharp(w_{B^{2q}})} \lesssim \sum_{J \in \mathbb{I}_{2q}(I)} \|E_J g\|^2_{L^2_\sharp(w_{B^{2q}})}.$$

□

6.4 Putting everything together. We will now prove inequality (10). It will suffice to work with $n = 2^u$, $u \in \mathbb{N}$.

Iterating Proposition 6.5 s times leads to the following multi-scale inequality, for each $p \geq 4$.

Proposition 6.6. *For each g supported on $I_1 \cup I_2$ and $s \leq u$ we have*

$$A_p\left(\frac{n}{2^s}, B^{2n}, \frac{n}{2^s}, g\right) \lesssim_{s,\epsilon} 2^{\epsilon s n} A_p\left(n, B^{2n}, n, g\right)^{(1-\kappa_p)^s} \prod_{l=1}^{s} D_p\left(\frac{n}{2^l}, B^{2n}, g\right)^{\kappa_p(1-\kappa_p)^{s-l}}.$$

Via one application of Cauchy–Schwarz we see that

$$\||Ef_1 Ef_2|^{1/2}\|_{L^p_\sharp(B^{2n})} = \left(\frac{1}{|\mathcal{B}_{\frac{n}{2^s}}(B^{2n})|} \sum_{B \in \mathcal{B}_{\frac{n}{2^s}}(B^{2n})} \||Ef_1 Ef_2|^{1/2}\|^p_{L^p_\sharp(B)}\right)^{1/p} \leq$$

$$\leq 2^{\frac{n}{2^{s+1}}} \left[\frac{1}{|\mathcal{B}_{\frac{n}{2^s}}(B^{2n})|} \sum_{B \in \mathcal{B}_{\frac{n}{2^s}}(B^{2n})} \left(\sum_{I \in \mathbb{I}_{\frac{n}{2^s}}(I_1)} \|E_I f_1\|^2_{L^p_\sharp(B)} \sum_{I \in \mathbb{I}_{\frac{n}{2^s}}(I_2)} \|E_I f_2\|^2_{L^p_\sharp(B)}\right)^{p/4}\right]^{1/p}$$

holds true for all f_1, f_2 supported on I_1 and I_2, respectively.

At this point we need to invoke the following reverse Hölder's inequality

$$\|E_I f_i\|_{L^p_\sharp(B)} \lesssim \|E_I f_i\|_{L^2_\sharp(B)}, \tag{20}$$

for each square B with side length $2^{\frac{n}{2^s}}$ and each I of length $2^{-\frac{n}{2^s}}$. This is a consequence of the fact that $|E_I f_i|$ is essentially constant on B.

We conclude as follows.

Proposition 6.7. *The inequality*

$$\||Ef_1 Ef_2|^{1/2}\|_{L^p_\sharp(B^{2n})} \leq 2^{\frac{n}{2^{s+1}}} A_p\left(\frac{n}{2^s}, B^{2n}, \frac{n}{2^s}, g\right)$$

holds true when f_1, f_2 are the restrictions of g to I_1, I_2, respectively.

To combine the last two propositions, we need one more inequality, a consequence of Proposition 6.1.

Proposition 6.8. *For each $l \leq u$*

$$D_p\left(\frac{n}{2^l}, B^{2n}, g\right) \lesssim \mathrm{Dec}\left(n - \frac{n}{2^l}, p\right) D_p\left(n, B^{2n}, g\right).$$

The following result is now rather immediate.

Theorem 6.9. *Assume f_1, f_2 are the restrictions of g to I_1 and I_2, respectively. Then for each $s \leq u$*

$$\||Ef_1 Ef_2|^{1/2}\|_{L^p_\sharp(B^{2n})} \lesssim_{s,\epsilon} 2^{\epsilon s n} 2^{\frac{n}{2^{s+1}}} D_p\left(n, B^{2n}, g\right) \prod_{l=1}^{s} \mathrm{Dec}\left(n - \frac{n}{2^l}, p\right)^{\kappa_p (1-\kappa_p)^{s-l}}$$

Proof. Using Hölder and Minkowski's inequality in $l^{\frac{p}{2}}$ we find that

$$A_p\left(n, B^{2n}, n, g\right) \lesssim D_p\left(n, B^{2n}, g\right).$$

Combine this with the previous three propositions. $\square$

Since this inequality holds for arbitrary g we can take the supremum to get the following inequality.

Corollary 6.10. *For each $s \leq u$*

$$\mathrm{BilDec}(n, p) \lesssim_{s,\epsilon} 2^{\epsilon s n} 2^{\frac{n}{2^{s+1}}} \prod_{l=1}^{s} \mathrm{Dec}\left(n - \frac{n}{2^l}, p\right)^{\kappa_p (1-\kappa_p)^{s-l}}.$$

We are now ready to finalize the proof of inequality (10). Take $p = 6$, and note that $\kappa_6 = \frac{1}{2}$. The case $p < 6$ would follow very similarly since $\kappa_p < \frac{1}{2}$.

We will use a bootstrapping argument. Assume $\mathrm{Dec}(n, 6) \lesssim 2^{nA}$ holds for some A and all n. For example, it is easy to see that $A = \frac{1}{2}$ works. We will show that a smaller value of A always works, too. Corollary 6.10 implies that for each s and each n

$$\mathrm{BilDec}(n, 6) \lesssim_s 2^{n\left(Ac_s + \frac{1}{2^s}\right)},$$

where

$$c_s = \sum_{l=1}^{s} \left(1 - \frac{1}{2^l}\right) \frac{1}{2^{s-l+1}} = 1 - \frac{s+1}{2^{s+1}}.$$

Combining this with Proposition 6.2 we may write

$$\text{Dec}\,(n,6) \lesssim_s 2^{n\left(Ac_s+\frac{1}{2^{s-1}}\right)}. \tag{21}$$

Define

$$\mathfrak{A} := \{A > 0 :\ \text{Dec}\,(n,6) \lesssim 2^{nA}\}$$

and let $A_0 = \inf \mathfrak{A}$. Note that $\mathfrak{A}$ is either (A_0, ∞) or $[A_0, \infty)$. We claim that A_0 must be zero, which will finish the proof of our theorem. Indeed, if $A_0 > 0$ then

$$Ac_s + \frac{1}{2^{s-1}} < A_0$$

for some $A \in \mathfrak{A}$ sufficiently close to A_0 and s sufficiently large. This combined with (21) contradicts the definition of A_0.

References

Jonathan Bennett, Anthony Carbery, and Terence Tao (2006). "On the multilinear restriction and Kakeya conjectures". *Acta Math.* 196.2, pp. 261–302. MR: 2275834 (cit. on pp. 1558, 1573).

J. Bourgain (2013). "Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces". *Israel J. Math.* 193.1, pp. 441–458. MR: 3038558 (cit. on p. 1563).

– (2017). "Decoupling, exponential sums and the Riemann zeta function". *J. Amer. Math. Soc.* 30.1, pp. 205–224. MR: 3556291 (cit. on p. 1565).

Jean Bourgain and Ciprian Demeter (2015). "The proof of the l^2 decoupling conjecture". *Ann. of Math. (2)* 182.1, pp. 351–389. MR: 3374964 (cit. on pp. 1563–1566).

– (2016a). "Decouplings for surfaces in $\mathbb{R}^4$". *J. Funct. Anal.* 270.4, pp. 1299–1318. MR: 3447712 (cit. on p. 1565).

– (2016b). "Mean value estimates for Weyl sums in two dimensions". *J. Lond. Math. Soc. (2)* 94.3, pp. 814–838. MR: 3614930 (cit. on p. 1565).

– (2017). "Decouplings for curves and hypersurfaces with nonzero Gaussian curvature". *J. Anal. Math.* 133, pp. 279–311. MR: 3736493 (cit. on p. 1565).

Jean Bourgain, Ciprian Demeter, and Shaoming Guo (2017). "Sharp bounds for the cubic Parsell-Vinogradov system in two dimensions". *Adv. Math.* 320, pp. 827–875. MR: 3709122 (cit. on p. 1565).

Jean Bourgain, Ciprian Demeter, and Larry Guth (2016). "Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three". *Ann. of Math. (2)* 184.2, pp. 633–682. MR: 3548534 (cit. on pp. 1564, 1565, 1569).

Jean Bourgain and Larry Guth (2011). "Bounds on oscillatory integral operators based on multilinear estimates". *Geom. Funct. Anal.* 21.6, pp. 1239–1295. MR: 2860188 (cit. on p. 1558).

Jean Bourgain and Nigel Watt (n.d.). "Mean square of zeta function, Gauss circle problem and divisor problem revisited". arXiv: 1709.04340 (cit. on p. 1565).

– (2017). "Decoupling for perturbed cones and mean square of $\zeta(\frac{1}{2}+it)$". *International Mathematics Research Notices*, rnx009 (cit. on p. 1565).

C. Demeter, S. Guo, and F. Shi (n.d.). "Sharp decouplings for three dimensional manifolds in $\mathbb{R}^5$". To appear in Revista Matematica Iberoamericana (cit. on p. 1565).

Ciprian Demeter (n.d.). "Incidence theory and restriction estimates". arXiv: 1401.1873 (cit. on p. 1563).

Gustavo Garrigós and Andreas Seeger (2009). "On plate decompositions of cone multipliers". *Proc. Edinb. Math. Soc. (2)* 52.3, pp. 631–651. MR: 2546636 (cit. on p. 1563).

– (2010). "A mixed norm variant of Wolff's inequality for paraboloids". In: *Harmonic analysis and partial differential equations*. Vol. 505. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 179–197. MR: 2664568 (cit. on p. 1563).

Emil Grosswald (1985). *Representations of integers as sums of squares*. Springer-Verlag, New York, pp. xi+251. MR: 803155 (cit. on p. 1561).

Shaoming Guo and Changkeun Oh (n.d.). "Remarks on Wolff's inequality for hypersurfaces". arXiv: 1602.05861 (cit. on p. 1565).

Larry Guth (2015). "A short proof of the multilinear Kakeya inequality". *Math. Proc. Cambridge Philos. Soc.* 158.1, pp. 147–153. MR: 3300318 (cit. on p. 1573).

Rowan Killip and Monica Vişan (2016). "Scale invariant Strichartz estimates on tori and applications". *Math. Res. Lett.* 23.2, pp. 445–472. MR: 3512894 (cit. on p. 1566).

Izabella Laba and Malabika Pramanik (2005). *Wolff's inequality for hypersurfaces* (cit. on p. 1563).

Izabella Łaba and Thomas Wolff (2002). "A local smoothing estimate in higher dimensions". *J. Anal. Math.* 88. Dedicated to the memory of Tom Wolff, pp. 149–171. MR: 1956533 (cit. on p. 1563).

Malabika Pramanik and Andreas Seeger (2007). "L^p regularity of averages over curves and bounds for associated maximal operators". *Amer. J. Math.* 129.1, pp. 61–103. MR: 2288738 (cit. on p. 1563).

Elias M. Stein (1993). *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Vol. 43. Princeton Mathematical Series. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, pp. xiv+695. MR: 1232192 (cit. on p. 1559).

Robert S. Strichartz (1977). "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations". *Duke Math. J.* 44.3, pp. 705–714. MR: 0512086 (cit. on p. 1559).

Peter A. Tomas (1975). "A restriction theorem for the Fourier transform". *Bull. Amer. Math. Soc.* 81, pp. 477–478. MR: 0358216 (cit. on p. 1559).

T. Wolff (2000). "Local smoothing type estimates on L^p for large p". *Geom. Funct. Anal.* 10.5, pp. 1237–1288. MR: 1800068 (cit. on p. 1563).

Trevor D. Wooley (2016). "The cubic case of the main conjecture in Vinogradov's mean value theorem". *Adv. Math.* 294, pp. 532–561. MR: 3479572 (cit. on p. 1569).

Received 2017-11-02.

Ciprian Demeter
Department of Mathematics
Indiana University
Bloomington IN
demeterc@indiana.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1579–1600)

PLURIPOTENTIAL THEORY AND COMPLEX DYNAMICS IN HIGHER DIMENSION

Tien-Cuong Dinh

Abstract

Positive closed currents, the analytic counterpart of effective cycles in algebraic geometry, are central objects in pluripotential theory. They were introduced in complex dynamics in the 1990s and become now a powerful tool in the field. Challenging dynamical problems involve currents of any dimension. We will report recent developments on positive closed currents of arbitrary dimension, including the solutions to the regularization problem, the theory of super-potentials and the theory of densities. Applications to dynamics such as properties of dynamical invariants (e.g. dynamical degrees, entropies, currents, measures), solutions to equidistribution problems, and properties of periodic points will be discussed.

Contents

The author was supported by Start-Up Grant R-146-000-204-133 and Tier 1 Grant R-146-000-248-114 from the National University of Singapore.

MSC2010: primary 32Uxx; secondary 37F10, 37F75.

Keywords: positive closed current, super-potential, tangent current, dynamical system, dynamical degree, entropy, equidistribution, periodic point.

1 Introduction

Let X be a compact Kähler manifold of dimension k. Let $f : X \to X$ be a dynamical system associated with a dominant holomorphic map, or more generally, a meromorphic map or correspondence, i.e. multivalued map. As a basic example, one can consider the complex affine space $\mathbb{C}^k$ as the complement of a projective hyperplane in the complex projective space $\mathbb{P}^k$. Then, any polynomial map from $\mathbb{C}^k$ to $\mathbb{C}^k$ extends to a meromorphic map from $\mathbb{P}^k$ to $\mathbb{P}^k$.

Denote by $f^n := f \circ \cdots \circ f$ (n times) the iterate of order n of f. The aim of the theory of complex dynamics is to study the longtime asymptotic behaviour of the sequence $(f^n)_{n\geq 0}$. This includes not only the study of the orbits of points, sets, currents, measures, under the action of f, but also the dynamical invariants such as dynamical degrees, entropies, Green currents, equilibrium measures, and the distribution of periodic points, etc.

Complex dynamics in dimension 1 has a long history, going back to the works by Fatou and Julia in 1920s, see e.g. Berteloot and Mayer [2001] and Carleson and Gamelin [1993]. In 1965, Brolin considered the harmonic measure of the Julia set of a polynomial in one complex variable which turns out to be a fundamental dynamical object, see Brolin [1965]. In 1981, Sibony considered the Green functions associated with Brolin's measures of polynomials of fixed degree. They can be obtained as the rate of escaping to infinity of the orbits of points in $\mathbb{C}$ under the action of the polynomials, see Sibony [1984, 1999]. Sibony also considered these Green functions in a family which constitute the Green function for some dynamical systems in higher dimension. Hubbard extended this notion of Green function to complex Hénon maps on $\mathbb{C}^2$, see Hubbard [1986]. In 1990, Sibony considered positive closed currents associated to these Green functions and their intersection, see Bedford, Lyubich, and Smillie [1993b, p.78] and also Sibony [1999].

Green currents and their intersections turn out to be fundamental objects in dynamics and pluripotential theory becomes a powerful tool in the field. The theory of complex dynamics of several variables has been developed quickly, see for example, the works by Bedford, Lyubich, and Smillie [1993a,b] and Bedford and Smillie [1991, 1992] and Fornæss and Sibony [1992, 1994a,b,c, 1995a] among others. One can observe that many works only involve currents of bi-degree $(1, 1)$ and their intersections because pluripotential theory has been developed first in this setting. However, some very basic questions already show the necessity of using positive closed currents of arbitrary bi-degree.

We will see in this survey different applications of such currents. Let's illustrate here their important role in the following basic picture. The periodic points of period n of f are the solutions of the equation $f^n(z) = z$. They can be identified with the intersection of the graph Γ_n of f^n with the diagonal Δ of $X \times X$. When n goes to infinity, for interesting dynamical systems, the volume of Γ_n tends to infinity. So in order to study the distribution of periodic points when n tends to infinity, it is necessary to consider the positive closed

(k,k)-current $[\Gamma_n]$ associated with Γ_n. Indeed, in this way, one can normalize $[\Gamma_n]$ to have mass 1 and consider the limit as n tends to infinity.

It is worth noting that in general Γ_n is not a complete intersection of hypersurfaces in $X \times X$: we may need more than k hypersurfaces in order to get Γ_n as their intersection. For example, the diagonal of $\mathbb{P}^k \times \mathbb{P}^k$, with $k \geq 2$, which is the graph of the identity map on $\mathbb{P}^k$, is not a complete intersection. More generally, the current associated with Γ_n is rarely the intersection of positive closed $(1,1)$-currents. So it is not enough to use $(1,1)$-currents to study Γ_n. Furthermore, computing the limit of a sequence of intersections of $(1,1)$-currents requires strong conditions on these currents which are not always available in the dynamical setting.

We will focus our discussion on the recent developments of pluripotential theory for currents of arbitrary bi-degree and their applications to dynamics. We refer the reader to the non-exhaustive list of references at the end of the paper, in particular the surveys Dinh and Sibony [2010a, 2017], Fornæss [1996], and Sibony [1999], for a more complete panorama of the theory of complex dynamics in higher dimension.

In Section 2, we will recall basic facts on currents and discuss the problem of approximating positive closed currents by appropriate smooth differential forms. As consequences, we will give some calculus on positive closed currents. Dynamical degrees, topological and algebraic entropies will be introduced together with the famous Gromov's inequality saying that the topological entropy is bounded from above by the algebraic one. The regularization theorem is a key point in the proofs.

In Section 3, we will introduce the notion of super-potentials which are canonical functions associated with positive closed (p,p)-currents. They play the role of quasi-plurisubharmonic functions which are used as quasi-potentials for positive closed $(1,1)$-currents. An intersection theory for positive closed currents of arbitrary bi-degrees will be presented. We then state some theorems in dynamics on the equidistribution of orbits of points and varieties. Unique ergodicity property and rigidity for dynamical currents will be discussed.

In Section 4, we will introduce the theory of densities for positive closed currents. A basic example of the theory is the case of two analytic subsets whose intersection is larger than expected, in terms of dimension. The densities are introduced in order to measure the dimension excess for the intersection of positive closed currents, see Fulton [1998] for an algebraic counterpart. Applications to dynamics concerning the distribution or the counting of periodic points will be considered.

Finally, in Section 5, some open problems in dynamics will be stated. They are related to the discussions in the previous sections and will require new ideas from pluripotential theory or from complex geometry. We expect that the solutions to these questions will provide new tools for complex dynamics.

Acknowledgments. The paper was partially prepared during a visit of the author at the University of Tokyo. He would like to express the gratitude to this university and to Keiji Oguiso for their hospitality and support. The author also thanks Viet-Anh Nguyen, Nessim Sibony, Tuyen-Trung Truong and Duc-Viet Vu for their collaborations which are largely reported in this paper.

2 Regularization of currents, dynamical degrees and entropies

In this section, we will discuss a regularization theorem for positive closed currents and its applications. We refer the reader to Demailly [2012], Hörmander [1990], Siu [1974] for basic notions and results of pluripotential theory and to Voisin [2002] for Hodge theory on compact Kähler manifolds.

Let X be a compact Kähler manifold of dimension k and let ω be a Kähler form on X. Let T be a positive closed (p,p)-current on X. The pairing $\langle T, \omega^{k-p}\rangle$, i.e. the value of T at the test form ω^{k-p}, depends only on the (Hodge or de Rham) cohomology classes of T and ω. Moreover, this quantity is comparable with the mass of T which is, by definition, the norm of T as a linear operator on the space of continuous test $(k-p,k-p)$-forms. Therefore, a large part of the computations with positive closed currents reduces to a computation with cohomology classes which is often simpler.

Positive closed currents can be seen as positive closed differential forms with distribution coefficients. In general, they are singular and calculus with them requires suitable regularization processes. The following result gives us a regularization with a control of the positivity loss, see Demailly [1992] and Dinh and Sibony [2004]. The loss of positivity is unavoidable in general. For simplicity, we also call $\|T\| := \langle T, \omega^{k-p}\rangle$ *the mass* of T.

Theorem 2.1 (Demailly for $p=1$, Dinh–Sibony for $p \geq 1$). *Let (X,ω) be a compact Kähler manifold. There is a constant $c>0$ depending only on X and ω satisfying the following property. If T is a positive closed (p,p)-current on X, there are positive closed (p,p)-currents T^+ and T^- which can be approximated by smooth positive closed (p,p)-forms and such that*

$$T = T^+ - T^- \qquad \text{and} \qquad \|T^\pm\| \leq c\|T\|.$$

This result still holds for larger classes of currents, e.g. positive dd^c-closed currents. It is the analytic counterpart of the known fact in algebraic geometry that any cycle can be represented as the difference of movable effective cycles. The regularization process used in the proof preserves good properties of T when they exist. We will give now two consequences of the regularization theorem. They are used to prove the properties of dynamical degrees and entropies that we will discuss later.

Corollary 2.2. *Let X, ω be as above and let U be an open subset of X. Let $T_1, \dots, T_n$ be positive closed currents on X of mass at most equal to 1 whose total bi-degree is at most (k, k). Assume that $T_1, \dots, T_{n-1}$ are given by smooth positive closed forms on U; so the intersection (wedge-product) $T_1 \wedge \dots \wedge T_n$ is a well-defined positive closed current on U. Then, there is a positive closed current S on X such that $T_1 \wedge \dots \wedge T_n \leq S$ on U and the mass of S on X is bounded by a constant depending only on X, ω.*

In the dynamical setting, we need to work with positive closed forms which are smooth outside an analytic subset of X. This corollary allows us to show that the integrals involving such singular forms do not explode near the set of singularities.

Recall that a meromorphic map from X to X is a holomorphic map f from a dense Zariski open set Ω of X to X whose graph in $\Omega \times X$ is a Zariski open set of an irreducible analytic subset Γ of dimension k in $X \times X$. For simplicity, we call Γ the graph of the meromorphic map $f : X \to X$. We assume that f is *dominant*, that is, the image of f contains a non-empty open subset of X, see Oguiso [2016b,a, 2017] and Oguiso and Truong [2015] for some recent examples.

Denote by π_1 and π_2 the two canonical projections from $X \times X$ to X. So the map π_1 restricted to Γ is generically 1:1. Let $I(f)$ be the set of points $x \in X$ such that $\Gamma \cap \pi_1^{-1}(x)$ is not a single point, or equivalently, of positive dimension. This is the *indeterminacy set* of f which is an analytic subset of codimension at least 2 in X. It is non-empty when f is not holomorphic on X.

Consider two dominant meromorphic maps f and f' from X to X. We can define the composition $f' \circ f$ as a holomorphic map on a suitable Zariski open set of X and then extend it to a meromorphic map from X to X. By composing f with itself, we obtain the iterates of f.

Let S be a (p, q)-current on X. Define formally the *pull-back* of S by f by

$$f^*(S) := (\pi_1)_*(\pi_2^*(S) \wedge [\Gamma]),$$

when the last expression makes sense. Since the operators π_i^* and $(\pi_i)_*$ are well-defined on all currents, the last definition is meaningful when the wedge-product $\pi_2^*(S) \wedge [\Gamma]$ is meaningful. Similarly, the *push-forward* operator f_* is defined by

$$f_*(S) := (\pi_2)_*(\pi_1^*(S) \wedge [\Gamma]),$$

when the last expression makes sense.

Consider the particular case of a smooth differential (p, q)-form ϕ on X. The wedge-product $\pi_2^*(\phi) \wedge [\Gamma]$ is well-defined because $\pi_2^*(\phi)$ is smooth. So $f^*(\phi)$ is well-defined in the sense of currents. Moreover, the value of $f^*(\phi)$ at a point x is roughly the sum of the values of $\pi_2^*(\phi)$ on the fiber $\pi_1^{-1}(x) \cap \Gamma$. We can check that $f^*(\phi)$ is in general an

L^1 form and it may be singular at the indeterminacy set $I(f)$. So we cannot iterate the operator f^* on smooth forms.

Recall that the Hodge cohomology group $H^{p,q}(X,\mathbb{C})$ of X can be defined using either smooth forms or singular currents. When ϕ is closed or exact then $f^*(\phi)$ is also closed or exact. Therefore, the above operator f^* induces a linear map from $H^{p,q}(X,\mathbb{C})$ to itself, that we still denote by f^*. The operator f_* on $H^{p,q}(X,\mathbb{C})$ is defined similarly. We can iterate those operators as for every linear operator on a vector space but in general we don't have $(f^n)^* = (f^*)^n$ on $H^{p,q}(X,\mathbb{C})$.

Consider an arbitrary positive closed (p,p)-current T on X. The pull-back $f^*(T)$ and the push-forward $f_*(T)$ of T are not always well-defined. We can however define a *strict transform* of T by f in the following way. Choose a Zariski open set Ω of X such that π_2 restricted to $\Gamma \cap \pi_2^{-1}(\Omega)$ defines a unramified covering over Ω. Then the pull-back of T by π_2 is well-defined on $\Gamma \cap \pi_2^{-1}(\Omega)$. We can show using Theorem 2.1 that it has finite mass and then its extension by 0 is a positive closed current on $X \times X$, according to a theorem of Skoda [1982]. The push-forward of the last current by π_1 is a positive closed (p,p)-current of X that we denote by $f^\bullet(T)$. We define $f_\bullet(T)$ in a similar way.

In the following result, the norms of the operators f^* and f_* are considered using a fixed norm on the vector space $H^{p,p}(X,\mathbb{C})$.

Corollary 2.3. *There is a constant $c > 0$ depending only on X, ω and the norm on $H^{p,p}(X,\mathbb{C})$ such that*

$$\|f^\bullet(T)\| \leq c\|T\|\|f^* : H^{p,p}(X,\mathbb{C}) \to H^{p,p}(X,\mathbb{C})\|$$

and

$$\|f_\bullet(T)\| \leq c\|T\|\|f_* : H^{p,p}(X,\mathbb{C}) \to H^{p,p}(X,\mathbb{C})\|.$$

This result is clear when T is a smooth form. We then deduce the general case using Theorem 2.1. Note that the operators $f^\bullet$ and $f_\bullet$ depend on the choice of a Zariski open set. However, when we work with L^1 forms for example, this choice is not important. Note also that the constants involved in the above results do not depend on T nor on f. In the proofs of the results below, they will intervene under the form $c^{1/n}$ and their role will be negligible when n goes to infinity.

As mentioned above, we don't have in general $(f^n)^* = (f^*)^n$ on $H^{p,q}(X,\mathbb{C})$. However, we can show that the following quantities are always well-defined.

Definition 2.4. We call *dynamical degree* of order p of f the following limit

$$d_p(f) := \lim_{n\to\infty} \|(f^n)^* : H^{p,p}(X,\mathbb{C}) \to H^{p,p}(X,\mathbb{C})\|^{1/n},$$

and *algebraic entropy* of f the following quantity

$$h_a(f) := \max_{0\leq p\leq k} \log d_p(f).$$

The last dynamical degree $d_k(f)$ is also called *topological degree* because it is equal to the number of points in $f^{-1}(a)$ for a generic point a in X.

Note that by Poincaré duality, we also have

$$d_p(f) := \lim_{n\to\infty} \|(f^n)_* : H^{k-p,k-p}(X,\mathbb{C}) \to H^{k-p,k-p}(X,\mathbb{C})\|^{1/n}.$$

Theorem 2.5 (Dinh–Sibony)**.** *The limit in the above definition of $d_p(f)$ always exists. It is finite and doesn't depend on the choice of the norm on $H^{p,p}(X,\mathbb{C})$. Moreover, the dynamical degrees and the algebraic entropy are bi-meromorphic invariants of the dynamical system: if $\pi : X' \to X$ is a bi-meromorphic map between compact Kähler manifolds, then*

$$d_p(\pi^{-1}\circ f\circ \pi) = d_p(f) \quad \textit{and} \quad h_a(\pi^{-1}\circ f\circ \pi) = h_a(f).$$

We also have for $n \geq 1$ that $d_p(f^n) = d_p(f)^n$ and $h_a(f^n) = nh_a(f)$.

When X is a projective space, the first statement was used by Fornæss–Sibony for $p = 1$ in order to construct the Green dynamical $(1,1)$-current Fornæss and Sibony [1994c]. Also for projective spaces, it was extended by Russakovskii–Shiffman for higher degrees Russakovskii and Shiffman [1997]. In this case, the group $H^{p,p}(X,\mathbb{C})$ is of dimension 1 and the action of $(f^n)^*$ is just the multiplication by an integer $d_{p,n}$. Therefore, we easily get $d_{p,n+m} \leq d_{p,n}d_{p,m}$ which implies the existence of the limit of $(d_{p,n})^{1/n}$ as n tends to infinity.

The proof of the above theorem in the general case uses in an essential way a computation with positive closed currents and Theorem 2.1 plays a crucial role. We refer to Dinh and Sibony [2004, 2005] for details and Dinh, Nguyên, and Truong [2012], Esnault and Srinivas [2013], and Truong [2016] for related results. We also obtained in these works the following result, which is due to Gromov for holomorphic maps Gromov [2003].

Theorem 2.6 (Gromov, Dinh–Sibony)**.** *Let X and f be as above. Then the topological entropy $h_t(f)$ of f is bounded from above by its algebraic entropy $h_a(f)$. In particular, the topological entropy of f is finite.*

The topological entropy is an important dynamical invariant. It measures the rate of divergence of the orbits of points. The formal definition for meromorphic maps is the same as the Bowen's definition for continuous maps, except that we don't consider orbits which reach the indeterminacy set. Therefore, it is not obvious that the entropy of a meromorphic map is finite. Note also that when f is a holomorphic map, the above result combined with a theorem by Yomdin [1987] implies that the topological entropy is indeed equal to the algebraic one. This property still holds for large families of meromorphic maps. We don't know if in general, there is always a map $\hat{f}$ bi-meromorphically conjugate to f such that $h_t(\hat{f}) = h_a(\hat{f})$, see Problem 5.1 below.

Observe that the action of f^n on $H^{p,q}(X, \mathbb{C})$ is not explicitly used in the above property of entropies when $p \neq q$. This can be explained by the following inequality from Dinh [2005]

$$\limsup_{n\to\infty} \|(f^n)^* : H^{p,q}(X, \mathbb{C}) \to H^{p,q}(X, \mathbb{C})\|^{1/n} \leq \sqrt{d_p(f)d_q(f)}.$$

Let T be a positive closed (p, p)-current on X, for example, the current of integration on a complex subvariety of codimension p. Applying Corollary 2.3 to f^n instead of f, we obtain that the mass of $(f^n)^\bullet(T)$ is bounded by a constant times $(d_p(f) + \epsilon)^n$ for every $\epsilon > 0$. Similarly, the mass of $(f^n)_\bullet(T)$ is bounded by a constant times $(d_{k-p}(f) + \epsilon)^n$. We see that dynamical degrees measures the growth of the degree and volume of varieties under the action of f or its inverse f^{-1}. So dynamical degrees are fundamental invariants in the study of the orbits of varieties. They play, with some variants, an important role in the problem of classification of meromorphic dynamical systems using invariant meromorphic fibrations, see Amerik and Campana [2008], Dinh, Nguyên, and Truong [2012], Nakayama and Zhang [2009], Oguiso [2016a], and Zhang [2009a,b] for details

Finally, recall that a direct consequence of the mixed Hodge–Riemann theorem applied to (resolutions of singularities of) the graphs of f^n, see e.g. Dinh and Nguyên [2006] and Gromov [1990], implies that, the function $p \mapsto \log d_p(f)$ is concave. Equivalently, we have

$$d_p(f)^2 \geq d_{p-1}(f)d_{p+1}(f) \quad \text{for} \quad 1 \leq p \leq k-1.$$

In particular, we have $1 \leq d_p(f) \leq d_1(f)^p$, $h_a(f) > 0$ if and only if $d_1(f) > 1$, and there are two numbers r and s with $0 \leq r \leq s \leq k$ such that

$$1 = d_0(f) < \cdots < d_r(f) = \cdots = d_s(f) > \cdots > d_k(f).$$

The maximal dynamical degree $d_r(f)$ is also called *the main dynamical degree*. The algebraic entropy of f is then equal to $\log d_r(f)$.

3 Super-potentiel theory and equidistribution problems

Super-potentials have been introduced in order to deal with positive closed currents of arbitrary bi-degree. Let T be a positive closed (p, p)-current on a compact Kähler manifold X as above. Any analytic set of pure codimension p in X defines by integration a positive closed (p, p)-current. So the current T can be seen as a generalization of analytic sets of codimension p.

When $p = 1$, the current T can be seen as a generalization of hypersurfaces. Locally, we can write $T = dd^c u$, where u is a plurisubharmonic (p.s.h. for short) function. This function is unique up to an additive pluriharmonic function which is real analytic. Globally, if α is a smooth closed real $(1,1)$-form on X, in the cohomology class of T, by the classical $\partial\overline{\partial}$-lemma, one can write $T = \alpha + dd^c u$. Here, u is a quasi-p.s.h. function on X, that is, u is locally the sum of a p.s.h. function and a smooth function. It is uniquely determined by T and α, up to an additive constant. In particular, there is a unique function u such that $\max u = 0$. Recall that $d^c := \frac{1}{2\pi i}(\partial - \overline{\partial})$ and $dd^c = \frac{i}{\pi}\partial\overline{\partial}$.

For the case of higher bi-degree, the current T corresponds to a generalized algebraic cycle of higher co-dimension. We still can write T in a similar way but u will be a current of bi-degree $(p-1, p-1)$. It doesn't satisfy a similar uniqueness property and there is no intrinsic choice for u. Super-potentials are canonical functions defined on some infinite dimensional spaces. They play the role of quasi-potentials as quasi-p.s.h. functions do for bi-degree $(1,1)$. For simplicity, we will not introduce this notion in full generality and refer the reader to Dinh and Sibony [2009, 2010c] for details.

Let $\mathfrak{D}_q(X)$ denote the real vector space spanned by positive closed (q,q)-currents on X. Define the $*$-*norm* on this space by $\|R\|_* := \min(\|R^+\| + \|R^-\|)$, where $R^\pm$ are positive closed (q,q)-currents satisfying $R = R^+ - R^-$. We consider this space of currents with the following topology : a sequence $(R_n)_{n\geq 0}$ in $\mathfrak{D}_q(X)$ converges in this space to R if $R_n \to R$ weakly and if $\|R_n\|_*$ is bounded independently of n. On any $*$-bounded subset of $\mathfrak{D}_q(X)$, this topology coincides with the classical weak topology for currents. By Theorem 2.1, the subspace $\widetilde{\mathfrak{D}}_q(X)$ of real closed smooth (q,q)-forms is dense in $\mathfrak{D}_q(X)$ for the considered topology.

Let $\mathfrak{D}_q^0(X)$ and $\widetilde{\mathfrak{D}}_q^0(X)$ denote the linear subspaces in $\mathfrak{D}_q(X)$ and $\widetilde{\mathfrak{D}}_q(X)$ respectively of currents whose cohomology classes in $H^{q,q}(X,\mathbb{R})$ vanish. Their co-dimensions are equal to the dimension of $H^{q,q}(X,\mathbb{R})$ which is finite. Fix a real smooth and closed (p,p)-form α in the cohomology class of T in $H^{p,p}(X,\mathbb{R})$. We will consider the super-potential of T which is the real function $\mathcal{U}_T$ on $\widetilde{\mathfrak{D}}^0_{k-p+1}(X)$ defined by

$$\mathcal{U}_T(R) := \langle T - \alpha, U_R\rangle \qquad \text{for} \qquad R \in \widetilde{\mathfrak{D}}^0_{k-p+1}(X),$$

where U_R is any smooth form of bi-degree $(k-p, k-p)$ such that $dd^c U_R = R$. This form always exists because the cohomology class of R vanishes. Note that since the cohomology class of $T-\alpha$ vanishes, we can write $T - \alpha = dd^c U_T$ for some current U_T. By Stokes theorem, we have

$$\mathcal{U}_T(R) = \langle dd^c U_T, U_R\rangle = \langle U_T, dd^c U_R\rangle = \langle U_T, R\rangle.$$

We deduce from these identities that $\mathcal{U}_T(R)$ doesn't depend on the choice of U_R and U_T. However, $\mathcal{U}_T$ depends on the reference form α. Note also that if T is smooth, it is not necessary to take R and U_R smooth.

For simplicity, we will not consider other super-potentials of T. They are some affine extensions of $\mathcal{U}_T$ to some subspaces of $\mathfrak{D}_{k-p+1}(X)$. The following notions do not depend on the choice of super-potential nor on the reference form α. We say that T has a *bounded super-potential* if $\mathcal{U}_T$ is bounded on each $*$-bounded subset of $\widetilde{\mathfrak{D}}^0_{k-p+1}(X)$. We say that T has a *continuous super-potential* if $\mathcal{U}_T$ can be extended to a continuous function on $\mathfrak{D}^0_{k-p+1}(X)$ with respect to the topology previously introduced.

As the definition of super-potentials introduces a new space $\mathfrak{D}_{k-p+1}(X)$, their calculus is not immediate. Recently, with Nguyen and Vu, we proved that if a positive closed current is bounded by another one with bounded or continuous super-potentials, then it satisfies the same property Dinh, Nguyên, and Truong [2017b]. The result plays a role in some constructions of dynamical Green currents and the study of periodic points. Super-potentials also permit to build an intersection theory, see Bedford and Taylor [1982], Demailly [2012], and Fornæss and Sibony [1995b] for the case of bi-degree $(1,1)$. In the dynamical setting, they allow us to define invariant measures as intersections of dynamical Green currents.

Consider two positive closed currents T and S on X of bi-degree (p,p) and (q,q) respectively. Assume that $p+q\leq k$ and that T has a continuous super-potential. So $\mathcal{U}_T$ is defined on whole $\mathfrak{D}^0_{k-p+1}(X)$. We can define the wedge-product $T\wedge S$ by

$$\langle T\wedge S,\phi\rangle := \langle \alpha\wedge S,\phi\rangle + \mathcal{U}_T(S\wedge dd^c\phi)$$

for every smooth real test form ϕ of bi-degree $(k-p-q,k-p-q)$. Note that $S\wedge dd^c\phi$ belongs to $\mathfrak{D}^0_{k-p+1}(X)$ because it is equal to $dd^c(S\wedge\phi)$. It is not difficult to check that $T\wedge S$ is equal to the usual wedge-product of T and S when one of them is smooth. The current $T\wedge S$ is positive and closed, see Dinh, Nguyên, and Truong [2017b], Dinh and Sibony [2009, 2010c], and Vu [2016b] for details.

In this short survey, we will not be able to discuss all properties of super-potentials. Let us focus our discussion in a key property which is crucial in the solution of equidistribution problems. It also illustrates how one can use super-potentials in a similar way that one can do with quasi-p.s.h. functions.

It is not difficult to show that quasi-p.s.h. functions are integrable with respect to the Lebesgue measure on X. However, we have the following much stronger property, see e.g. Dinh, Nguyên, and Sibony [2010], Kaufmann [2017], and Vu [2016a]. It implies that quasi-p.s.h. functions are L^p for all $1\leq p<\infty$.

Theorem 3.1 (Skoda)**.** *Let X and ω be as above. Let α be a smooth real closed $(1,1)$-form on X. There are constants $\lambda>0$ and $c>0$ such that if T is any positive closed $(1,1)$-current in the cohomology class of α and u is the quasi-p.s.h. function satisfying $dd^cu=T-\alpha$ and $\max u=0$, then we have*

$$\int_X e^{\lambda|u|}\omega^k\leq c.$$

It is not clear how to generalize this result to super-potentials because there is no natural measure on the domain of definition of super-potentials. The following result of Dinh and Sibony [2009, 2010c] gives us an answer to this question.

Theorem 3.2 (Dinh–Sibony). *Let X and ω be as above. Let α be a smooth real closed (p,p)-form on X. There is a constant $c>0$ such that if T is any positive closed (p,p)-current in the cohomology class of α and $\mathcal{U}_T$ is its super-potential defined above, then*

$$|\mathcal{U}_T(R)| \le c(1+\log^+\|R\|_{\mathcal{C}^1}),$$

for $R \in \mathfrak{D}^0_{k-p+1}(X)$ with $\|R\|_* = 1$, where $\log^+ := \max(\log, 0)$.

If we remove $\log^+$ from the statement, the obtained estimate is much weaker and easy to prove. So the contribution of $\log^+$ here is similar to the role of the exponential in Theorem 3.1. Several applications of super-potentials in dynamics have been obtained. We will only present here two results and refer the reader to Ahn [2016], De Thélin and Vigny [2010], and Dinh and Sibony [2009, 2010c] for some other applications, in particular, for dynamics of automorphisms of compact Kähler manifolds.

Let $\mathcal{H}_d$ denote the family of all holomorphic self-maps of $\mathbb{P}^k$ such that the first dynamical degree is an integer $d \ge 2$. This can be identified to a Zariski open subset of a projective space. A generic map from $\mathbb{C}^k$ to $\mathbb{C}^k$ whose components are polynomials of degree d can be extended to a holomorphic self-map of $\mathbb{P}^k$. The following result was obtained in Dinh and Sibony [2009], see also Ahn [2016] for some extension.

Theorem 3.3 (Dinh–Sibony). *There is an explicit dense Zariski open subset $\mathcal{H}'_d$ of $\mathcal{H}_d$ such that for every f in $\mathcal{H}'_d$ and every analytic subset V of pure codimension p and of degree $\deg(V)$ of $\mathbb{P}^k$ we have*

$$\lim_{n\to\infty} \frac{1}{d^{pn}\deg(V)}(f^n)^*[V] = T^p,$$

where T^p is the p-th power of the dynamical Green $(1,1)$-current T of f. Moreover, the convergence is uniform on V and exponentially fast with respect to some natural distances on the space of positive currents.

We have not yet introduced the Green current T. This is a positive closed $(1,1)$-current on $\mathbb{P}^k$, invariant by $d^{-1}f^*$, with unit mass and continuous potentials. The power T^p is well-defined and is called *Green (p,p)-current* of f. The last theorem gives us a construction of T by pulling back a hypersurface V by f^n (case $p=1$). However, T was originally constructed for every $f \in \mathcal{H}_d$ by pulling back smooth positive closed $(1,1)$-forms, see Dinh and Sibony [2010a] and Sibony [1999] for details.

Note that Theorem 3.3 still holds if we replace $[V]$ by any positive closed (p, p)-current. So f satisfies the unique ergodicity for currents. The result is however not true for every map f in $\mathcal{H}_d$. In general, there may exist exceptional analytic sets V for which the convergence in the theorem doesn't hold, see Conjecture 5.2 below. There are satisfactory equidistribution results for general maps f only when $p = k$ and $p = 1$, that is, when V is a point or a hypersurface. The full result for $p = k$ was obtained by Dinh and Sibony [2003, 2010b] generalizing results obtained by Fornæss and Sibony [1994c] and Briend and Duval [2001]. The case $p = 1$ was obtained in Dinh and Sibony [2008] and Taflin [2011] generalizing results obtained earlier by Fornæss and Sibony [1994b], Russakovskii and Shiffman [1997] and Favre and Jonsson [2003] (for $p = 1, k = 2$). The same property for dynamics in one variable, except the rate of convergence, has been proved by Brolin [1965], Freire, Lopes, and Mañé [1983] and Ljubich [1983].

The convergence of currents in Theorem 3.3 is equivalent to the convergence of their super-potentials. The rate of the convergence of super-potentials implies the rate of convergence of currents with respect to some natural distances for positive currents. These distances are analogous to the classical Kantorovich–Wasserstein distance for measures.

We discuss now the second result where, as for the last result, super-potentials and Theorem 3.2 play crucial roles in the proof. Let f be a polynomial automorphism of $\mathbb{C}^k$. We extend it to a birational map on the projective space $\mathbb{P}^k$. Denote by $I(f)$ and $I(f^{-1})$ the indeterminacy sets of f and f^{-1} respectively. They are analytic subsets of the hyperplane at infinity $\mathbb{P}^k \setminus \mathbb{C}^k$. The following notion was introduced under the name of regular automorphisms in Sibony [1999].

Definition 3.4 (Sibony). We say that f is a *Hénon-type automorphism* if f is not an automorphism of $\mathbb{P}^k$ and $I(f) \cap I(f^{-1}) = \varnothing$.

This is a large family of maps. In dimension 2, all polynomial automorphisms of $\mathbb{C}^2$ are conjugated either to Hénon-type maps as in Definition 3.4 or to elementary maps whose dynamics is simple to study, see Friedland and Milnor [1989]. Consider a Hénon-type map f as above. It is known that there is an integer p such that $\dim I(f) = k - p - 1$ and $\dim I(f^{-1}) = p - 1$. The action of f on cohomology is simple and $d_p(f)$ is the main dynamical degree.

It is also known that the set $I(f^{-1})$ is attractive for f. Let $\mathcal{U}(f)$ denote the basin of $I(f^{-1})$ which is an open neighbourhood of $I(f^{-1})$ in $\mathbb{P}^k$. The set $\mathcal{K}(f) := \mathbb{C}^k \setminus \mathcal{U}(f)$ is the set of all points $z \in \mathbb{C}^k$ whose orbits by f are bounded in $\mathbb{C}^k$. The closure $\overline{\mathcal{K}(f)}$ of $\mathcal{K}(f)$ in $\mathbb{P}^k$ is known to be the union of $\mathcal{K}(f)$ with $I(f)$. The following result was obtained in Dinh and Sibony [2009] generalizing results by Bedford, Lyubich, and Smillie [1993b] and Fornæss and Sibony [1994c], where the case of dimension 2, except the rate of convergence, was considered.

Theorem 3.5 (Dinh–Sibony). *Let V be an analytic subset of pure dimension $k-p$ and of degree* $\deg(V)$ *in $\mathbb{P}^k$ such that $V \cap I(f^{-1}) = \varnothing$. Then* $\deg(V)^{-1} d_p(f)^{-n}(f^n)^*[V]$ *converges exponentially fast to a positive closed (p,p)-current $T(f)$ with support in $\overline{\mathcal{K}(f)}$. Moreover, the set $\overline{\mathcal{K}(f)}$ is rigid in the sense that $T(f)$ is the unique positive closed (p,p)-current of mass 1 with support in $\overline{\mathcal{K}(f)}$.*

Note that the rigidity of $\overline{\mathcal{K}(f)}$ implies the convergence of $\deg(V)^{-1} d_p(f)^{-n}(f^n)^*[V]$ to $T(f)$ because $f^{-n}(V)$ converges to $\overline{\mathcal{K}(f)}$. However, it doesn't imply the rate of convergence. The current $T(f)$ is the dynamical Green (p,p)-current of f. It was constructed by Sibony as a power of the dynamical Green $(1,1)$-current. The later has been obtained by pulling back smooth positive closed $(1,1)$-forms, see Sibony [1999] and Taflin [2011] for details. Observe that Theorem 3.5 still holds if we replace $[V]$ by any positive closed (p,p)-current whose support is disjoint from $I(f^{-1})$. The result can be applied for f^{-1} instead of f since f^{-1} is also a Hénon-type automorphism of $\mathbb{C}^k$. We refer the reader to the survey Dinh and Sibony [2014] for a more complete panorama on rigidity property in dynamics.

4 Theory of densities of currents and periodic points

The theory of densities has been introduced in order to study the intersection between positive closed currents of arbitrary dimension. These currents may not admit an intersection in the classical sense. In particular, the theory permits to measure the dimension excess of the intersection and to understand what happens for the limit of such intersections. Such situations appear in several dynamical problems. We will not report on the theory in full generality and refer the reader to Dinh and Sibony [2012, 2018] for details. Some applications in dynamics will be discussed at the end of this section.

Consider the case of two positive closed currents : the first one is a general positive closed (p,p)-current T and the second one is the current of integration on a submanifold V of X. We want to understand the densities of T, i.e. the repartition of mass in various directions, along V via a notion of tangent current. The case where V is a point corresponds to the classical theory of Lelong number for positive closed currents. The rough idea is to dilate the manifold X in the normal directions to V. When the dilation factor tends to infinity, the image of T by the dilation admits limits that we will call tangent currents of T along V. They may not be unique but belong to the same cohomology class. However, in general, there is no natural dilations in the normal directions to V and tangent currents are defined in a more sophisticated way.

Let E denote the normal vector bundle to V in X and $\overline{E}$ its canonical compactification. Denote by $A_\lambda : \overline{E} \to \overline{E}$ the map induced by the multiplication by λ on fibers of E with $\lambda \in \mathbb{C}^*$. We also identify V with the zero section of E. The tangent currents to T along

V will be positive closed (p, p)-currents on $\overline{E}$ which are V*-conic*, i.e. invariant under the action of A_λ.

Let τ be a diffeomorphism between a neighbourhood of V in X and a neighbourhood of V in E whose restriction to V is identity. Assume that τ is *admissible* in the sense that the endomorphism of E induced by the differential of τ is the identity map from E to E. Using exponential maps associated with a Kähler metric on X, it is not difficult to show that such maps exist. Here is a main result of the theory of densities.

Theorem 4.1. *Let $X, V, T, E, \overline{E}, A_\lambda$ and τ be as above. Then the family of currents $T_\lambda := (A_\lambda)_* \tau_*(T)$ is relatively compact and any limit current, for $\lambda \to \infty$, is a positive closed (p, p)-current on E whose trivial extension is a positive closed (p, p)-current on $\overline{E}$. Moreover, if S is such a current, it is V-conic, i.e. invariant under $(A_\lambda)_*$, and its cohomology class in $H^{p,p}(\overline{E}, \mathbb{R})$ does not depend on the choice of τ and S.*

Note that T_λ is not of bi-degree (p, p) in general and one cannot talk about its positivity. The above theorem not only states the existence of a unique cohomology class, but it claims that it can be computed using any admissible τ. The result still holds and we obtain the same family of limit currents using local admissible diffeomorphisms. This flexibity is very useful in the analytic calculus with tangent currents and densities while the use of global admissible diffeomorphisms is convenient for calculus on cohomology.

We say that S is a *tangent current* to T along V. Its cohomology class is called the *total tangent class* of T along V. Note that this notion generalizes a notion of tangent cone in the algebraic setting where T is also given by a manifold. It measures the densities of T along V. The cohomology ring of $\overline{E}$ is generated by the cohomology ring of V and the tautological $(1, 1)$-class on $\overline{E}$. Therefore, we can decompose the cohomology class of S and associate to it cohomology classes of different degrees on V. These classes represent different parts of the tangent class of T along V.

Note also that for the general case of two arbitrary positive closed currents T and T' on X (the manifold V is replaced by a general current T'), the densities between T and T' are determined by the densities between the tensor product $T \otimes T'$ on $X \times X$ and the diagonal of $X \times X$. As already mentioned above, we will not develop the general case in this report.

It is important to estimate or compute the densities. The following particular case of Dinh and Sibony [2012, Th.4.11] is used in the proofs of the dynamical properties presented below. It is analogous to a result by Siu for Lelong numbers.

Theorem 4.2. *Let T_n be a sequence of positive closed (p, p)-current converging to a positive closed (p, p)-current T on X. Let V be a submanifold of X and denote by κ_n, κ the total tangent classes of T_n, T along V. Let c be the cohomology class of a projective subspace of a fiber of $\overline{E}$. Assume that $\kappa = \lambda c$ for some non-negative constant λ. Then* *any cluster value of* κ_n *has the form* $\lambda' c$ *with a constant* $0 \leq \lambda' \leq \lambda$. *In particular, if* $\lambda = 0$ *then* κ_n *tends to* 0.

We will give now some applications in dynamics where the last two statements play a crucial role in the proof. For the following result, see Bedford, Lyubich, and Smillie [1993a] and Dinh and Sibony [2016].

Theorem 4.3 (Bedford–Lyubich–Smillie for $k = 2$, Dinh–Sibony for $k \geq 2$)**.** *Under the hypotheses of Theorem 3.5, let* P_n *denote the set of periodic points of period* n *of* f *in* $\mathbb{C}^k$. *Then the points in* $P_n(f)$ *are asymptotically equidistributed with respect to the equilibrium measure* μ *of* f. *More precisely, if* δ_a *denotes the Dirac mass at a point* a, *then*

$$\lim_{n\to\infty} d_p(f)^{-n} \sum_{a\in P_n(f)} \delta_a = \mu.$$

The result still holds if we replace $P_n(f)$ *by the set of saddle periodic points of period* n.

The measure μ was constructed by Sibony [1999]. With the notations of Theorem 3.5, it is equal to the intersection of the Green current $T(f)$ of f and the Green current $T(f^{-1})$ of f^{-1}. It has support in the compact set $\mathcal{K}(f) \cap \mathcal{K}(f^{-1})$ in $\mathbb{C}^k$. If Δ denotes the diagonal of $\mathbb{P}^k \times \mathbb{P}^k$, then μ can be identified with the intersection between the current $[\Delta]$ and the tensor product $T(f) \otimes T(f^{-1})$.

Let Γ_n denote the graph of f^n in $\mathbb{P}^k \times \mathbb{P}^k$. The set $P_n(f)$ can be identified with the intersection of Γ_n and Δ in $\mathbb{C}^k \times \mathbb{C}^k$. Denote for simplicity $d := d_p(f)$. We can show that the positive closed (k,k)-currents $d^{-n}[\Gamma_n]$ converge to the current $T(f) \otimes T(f^{-1})$. Therefore, Theorem 4.3 is equivalent to the identity

$$\lim_{n\to\infty} \big([\Delta] \wedge d^{-n}[\Gamma_n]\big) = [\Delta] \wedge \big(\lim_{n\to\infty} d^{-n}[\Gamma_n]\big)$$

on $\mathbb{C}^k \times \mathbb{C}^k$.

In the general setting of the theory of currents, the two operations of intersection and of taking the limit, even when they are well-defined, may not commute. In our setting, the last identity requires a transversality property described below for the intersection between Γ_n and Δ which is, in some sense, uniform in n. To establish this property requires a delicate analysis using in particular a result by de Thélin [2008].

Let $\mathrm{Gr}(\mathbb{P}^k \times \mathbb{P}^k, k)$ denote the Grassmannian bundle over $\mathbb{P}^k \times \mathbb{P}^k$ where each point corresponds to a pair $(x, [v])$ of a point $x \in \mathbb{P}^k \times \mathbb{P}^k$ and the direction $[v]$ of a simple tangent k-vector v of $\mathbb{P}^k \times \mathbb{P}^k$ at x. Let $\widetilde{\Gamma}_n$ denote the set of points $(x, [v])$ in $\mathrm{Gr}(\mathbb{P}^k \times \mathbb{P}^k, k)$ with $x \in \Gamma_n$ and v a k-vector not transverse to Γ_n at x. Let $\widehat{\Delta}$ denote the lift of Δ to $\mathrm{Gr}(\mathbb{P}^k \times \mathbb{P}^k, k)$, i.e. the set of points $(x, [v])$ with $x \in \Delta$ and v tangent to Δ. The intersection $\widetilde{\Gamma}_n \cap \widehat{\Delta}$ corresponds to the non-transverse points of intersection between Γ_n

and Δ. Note that $\dim \widetilde{\Gamma}_n + \dim \widehat{\Delta}$ is smaller than the dimension of $\mathrm{Gr}(\mathbb{P}^k \times \mathbb{P}^k, k)$ and the intersection of subvarieties of such dimensions are generically empty.

We show that the currents $d^{-n}[\widetilde{\Gamma}_n]$ cluster towards a positive closed current whose tangent currents along $\widehat{\Delta}$ vanish. This together with Theorem 4.2 implies that the intersection between Γ_n and Δ is asymptotically transverse as n goes to infinity. As mentioned above, this is the key point in the proof of Theorem 4.3.

We will end this section with another application of the theory of densities. Let f be a general dominant meromorphic map from X to X. When the periodic points of period n of f are isolated, then their number, counting with multiplicity, can be obtained using Lefschetz fixed point formula. In general, this set may have components of positive dimension, but we still want to study the distribution of isolated periodic points, in particular, to count them. The following result was recently obtained in Dinh, Nguyên, and Truong [2017a] as a consequence of Theorem 4.2 and some properties of the sequence Γ_n.

Theorem 4.4 (Dinh–Nguyen–Truong)**.** *Let f be a dominant meromorphic self-map on a compact Kähler manifold X. Let $h_a(f)$ be its algebraic entropy and $P_n(f)$ its number of isolated periodic points of period n counted with multiplicity. Then we have*

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(f) \leq h_a(f).$$

In particular, f is an Artin–Mazur map, i.e., its number of isolated periodic points of period n grows at most exponentially fast with n.

Note that there are smooth real maps on compact manifolds which are not Artin–Mazur maps, see e.g. Artin and Mazur [1965] and Kaloshin [2000]. For large families of meromorphic maps or correspondences, we can obtain a sharp upper bound for the cardinality of $P_n(f)$ which is equal to $1 + o(1)$ times the number given by the Lefschetz fixed point formula Dinh, Nguyên, and Truong [2015, 2017a,b]. This is a crucial step in the study of the equidistribution property for these points. Lower bounds for the cardinality of $P_n(f)$ were also obtained in some cases using other ideas from dynamics. We refer to Cantat [2001], Diller, Dujardin, and Guedj [2010], Dujardin [2006], Favre [1998], Iwasaki and Uehara [2010], Jonsson and Reschke [2015], Saito [1987], and Xie [2015] for lower bounds and related results.

5 Some open problems

In this section, we will state three open problems which are related to our discussion in the previous three sections. We think that they are important problems in complex dynamics. They require new ideas and may provide new techniques that can be used to solve other questions.

The following problem is related to Theorem 2.6. It may require some ideas from complex analysis together with the techniques used by Yomdin [1987]. It was already briefly mentioned in Section 2.

Problem 5.1. Let $f : X \to X$ be a dominant meromorphic map. Does there always exist a bi-meromorphic map $\pi : X' \to X$ between compact Kähler manifolds such that the algebraic and topological entropies of $\pi^{-1} \circ f \circ \pi$ are equal ?

Some trivial examples show that we don't have this equality without modifying the manifold X, see Guedj [2005].

The following conjecture was stated in Dinh and Sibony [2008]. It may requires a deep understanding on the space of positive closed currents which is of infinite dimension. Let f be an endomorphism of $\mathbb{P}^k$ of algebraic degree $d \geq 2$. A proper analytic subset of $\mathbb{P}^k$ is said to be *totally invariant* if it is invariant by both f and f^{-1}. They appear as exceptional sets, where the multiplicity of f is large. Recall that the family of all these analytic sets is either empty or finite, see Dinh and Sibony [2010b].

Conjecture 5.2. *Let T be the Green $(1,1)$-current of f and let p be an integer with $2 \leq p \leq k-1$. Then $(\deg V)^{-1} d^{-pn} (f^n)^*[V]$ converge to T^p for every analytic subset V of $\mathbb{P}^k$ of pure codimension p which is generic. Here, V is generic if either $V \cap E = \varnothing$ or $\operatorname{codim} V \cap E = p + \operatorname{codim} E$ for any irreducible component E of a totally invariant proper analytic subset of $\mathbb{P}^k$.*

Finally, the following problem seems to be very challenging. The current approach to get the equidistribution of periodic points in Theorem 4.3 contains different steps. Several of them are quantifiable but some of them need to be substituted by new ideas from pluripotential theory.

Problem 5.3. Study the rate of convergence of periodic points of Hénon-type maps toward the equilibrium measure.

References

Taeyong Ahn (2016). "Equidistribution in higher codimension for holomorphic endomorphisms of $\mathbb{P}^k$". *Trans. Amer. Math. Soc.* 368.5, pp. 3359–3388. MR: 3451880 (cit. on p. 1589).

Ekaterina Amerik and Frédéric Campana (2008). "Fibrations méromorphes sur certaines variétés à fibré canonique trivial". *Pure Appl. Math. Q.* 4.2, Special Issue: In honor of Fedor Bogomolov. Part 1, pp. 509–545. MR: 2400885 (cit. on p. 1586).

Michael Artin and Barry Mazur (1965). "On periodic points". *Ann. of Math. (2)* 81, pp. 82–99. MR: 0176482 (cit. on p. 1594).

Eric Bedford, Mikhail Lyubich, and John Smillie (1993a). "Distribution of periodic points of polynomial diffeomorphisms of $\mathbf{C}^2$". *Invent. Math.* 114.2, pp. 277–288. MR: 1240639 (cit. on pp. 1580, 1593).

– (1993b). "Polynomial diffeomorphisms of $\mathbf{C}^2$. IV. The measure of maximal entropy and laminar currents". *Invent. Math.* 112.1, pp. 77–125. MR: 1207478 (cit. on pp. 1580, 1590).

Eric Bedford and John Smillie (1991). "Polynomial diffeomorphisms of $\mathbf{C}^2$: currents, equilibrium measure and hyperbolicity". *Invent. Math.* 103.1, pp. 69–99. MR: 1079840 (cit. on p. 1580).

– (1992). "Polynomial diffeomorphisms of $\mathbf{C}^2$. III. Ergodicity, exponents and entropy of the equilibrium measure". *Math. Ann.* 294.3, pp. 395–420. MR: 1188127 (cit. on p. 1580).

Eric Bedford and B. Allan Taylor (1982). "A new capacity for plurisubharmonic functions". *Acta Math.* 149.1-2, pp. 1–40. MR: 674165 (cit. on p. 1588).

François Berteloot and Volker Mayer (2001). *Rudiments de dynamique holomorphe*. Vol. 7. Cours Spécialisés [Specialized Courses]. Société Mathématique de France, Paris; EDP Sciences, Les Ulis, pp. vi+160. MR: 1973050 (cit. on p. 1580).

Jean-Yves Briend and Julien Duval (2001). "Deux caractérisations de la mesure d'équilibre d'un endomorphisme de $\mathrm{P}^k(\mathbf{C})$". *Publ. Math. Inst. Hautes Études Sci.* 93, pp. 145–159. MR: 1863737 (cit. on p. 1590).

Hans Brolin (1965). "Invariant sets under iteration of rational functions". *Ark. Mat.* 6, 103–144 (1965). MR: 0194595 (cit. on pp. 1580, 1590).

Serge Cantat (2001). "Dynamique des automorphismes des surfaces $K3$". *Acta Math.* 187.1, pp. 1–57. MR: 1864630 (cit. on p. 1594).

Lennart Carleson and Theodore W. Gamelin (1993). *Complex dynamics*. Universitext: Tracts in Mathematics. Springer-Verlag, New York, pp. x+175. MR: 1230383 (cit. on p. 1580).

Henry De Thélin and Gabriel Vigny (2010). "Entropy of meromorphic maps and dynamics of birational maps". *Mém. Soc. Math. Fr. (N.S.)* 122, pp. vi+98. MR: 2752759 (cit. on p. 1589).

Jean-Pierre Demailly (1992). "Regularization of closed positive currents and intersection theory". *J. Algebraic Geom.* 1.3, pp. 361–409. MR: 1158622 (cit. on p. 1582).

– (2012). "Complex analytic and differential geometry" (cit. on pp. 1582, 1588).

Jeffrey Diller, Romain Dujardin, and Vincent Guedj (2010). "Dynamics of meromorphic maps with small topological degree III: geometric currents and ergodic theory". *Ann. Sci. Éc. Norm. Supér. (4)* 43.2, pp. 235–278. MR: 2662665 (cit. on p. 1594).

Tien-Cuong Dinh (2005). "Suites d'applications méromorphes multivaluées et courants laminaires". *J. Geom. Anal.* 15.2, pp. 207–227. MR: 2152480 (cit. on p. 1586).

Tien-Cuong Dinh and Viêt-Anh Nguyên (2006). "The mixed Hodge-Riemann bilinear relations for compact Kähler manifolds". *Geom. Funct. Anal.* 16.4, pp. 838–849. MR: 2255382 (cit. on p. 1586).

Tien-Cuong Dinh, Viêt-Anh Nguyên, and Nessim Sibony (2010). "Exponential estimates for plurisubharmonic functions and stochastic dynamics". *J. Differential Geom.* 84.3, pp. 465–488. MR: 2669362 (cit. on p. 1588).

Tien-Cuong Dinh, Viêt-Anh Nguyên, and Tuyen Trung Truong (2012). "On the dynamical degrees of meromorphic maps preserving a fibration". *Commun. Contemp. Math.* 14.6, pp. 1250042, 18. MR: 2989646 (cit. on pp. 1585, 1586).

– (2015). "Equidistribution for meromorphic maps with dominant topological degree". *Indiana Univ. Math. J.* 64.6, pp. 1805–1828. MR: 3436236 (cit. on p. 1594).

– (2017a). "Growth of the number of periodic points for meromorphic maps". *Bulletin of the London Mathematical Society* 49.6, pp. 947–964 (cit. on p. 1594).

– (Oct. 2017b). "Super-potentials, densities of currents and number of periodic points for holomorphic maps". arXiv: 1710.01473 (cit. on pp. 1588, 1594).

Tien-Cuong Dinh and Nessim Sibony (2003). "Dynamique des applications d'allure polynomiale". *J. Math. Pures Appl. (9)* 82.4, pp. 367–423. MR: 1992375 (cit. on p. 1590).

– (2004). "Regularization of currents and entropy". *Ann. Sci. École Norm. Sup. (4)* 37.6, pp. 959–971. MR: 2119243 (cit. on pp. 1582, 1585).

– (2005). "Une borne supérieure pour l'entropie topologique d'une application rationnelle". *Ann. of Math. (2)* 161.3, pp. 1637–1644. MR: 2180409 (cit. on p. 1585).

– (2008). "Equidistribution towards the Green current for holomorphic maps". *Ann. Sci. Éc. Norm. Supér. (4)* 41.2, pp. 307–336. MR: 2468484 (cit. on pp. 1590, 1595).

– (2009). "Super-potentials of positive closed currents, intersection theory and dynamics". *Acta Math.* 203.1, pp. 1–82. MR: 2545825 (cit. on pp. 1587–1590).

– (2010a). "Dynamics in several complex variables: endomorphisms of projective spaces and polynomial-like mappings". In: *Holomorphic dynamical systems*. Vol. 1998. Lecture Notes in Math. Springer, Berlin, pp. 165–294. MR: 2648690 (cit. on pp. 1581, 1589).

– (2010b). "Equidistribution speed for endomorphisms of projective spaces". *Math. Ann.* 347.3, pp. 613–626. MR: 2640045 (cit. on pp. 1590, 1595).

– (2010c). "Super-potentials for currents on compact Kähler manifolds and dynamics of automorphisms". *J. Algebraic Geom.* 19.3, pp. 473–529. MR: 2629598 (cit. on pp. 1587–1589).

– (Mar. 2012). "Density of positive closed currents, a theory of non-generic intersections". arXiv: 1203.5810 (cit. on pp. 1591, 1592).

– (2014). "Rigidity of Julia sets for Hénon type maps". *J. Mod. Dyn.* 8.3-4, pp. 499–548. MR: 3345839 (cit. on p. 1591).

Tien-Cuong Dinh and Nessim Sibony (2016). "Equidistribution of saddle periodic points for Hénon-type automorphisms of $\mathbb{C}^k$". *Math. Ann.* 366.3-4, pp. 1207–1251. MR: 3563236 (cit. on p. 1593).

– (2017). "Equidistribution problems in complex dynamics of higher dimension". *Internat. J. Math.* 28.7, pp. 1750057, 31. MR: 3667901 (cit. on p. 1581).

– (2018). "Unique ergodicity for foliations in $\mathbb{P}^2$ with an invariant curve". *Invent. Math.* 211.1, pp. 1–38. MR: 3742755 (cit. on p. 1591).

Romain Dujardin (2006). "Laminar currents and birational dynamics". *Duke Math. J.* 131.2, pp. 219–247. MR: 2219241 (cit. on p. 1594).

Hélène Esnault and Vasudevan Srinivas (2013). "Algebraic versus topological entropy for surfaces over finite fields". *Osaka J. Math.* 50.3, pp. 827–846. MR: 3129006 (cit. on p. 1585).

Charles Favre (1998). "Points périodiques d'applications birationnelles de $\mathbf{P}^2$". *Ann. Inst. Fourier (Grenoble)* 48.4, pp. 999–1023. MR: 1656005 (cit. on p. 1594).

Charles Favre and Mattias Jonsson (2003). "Brolin's theorem for curves in two complex dimensions". *Ann. Inst. Fourier (Grenoble)* 53.5, pp. 1461–1501. MR: 2032940 (cit. on p. 1590).

John Erik Fornæss (1996). *Dynamics in several complex variables*. Vol. 87. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, pp. viii+59. MR: 1363948 (cit. on p. 1581).

John Erik Fornæss and Nessim Sibony (1992). "Complex Hénon mappings in $\mathbf{C}^2$ and Fatou-Bieberbach domains". *Duke Math. J.* 65.2, pp. 345–380. MR: 1150591 (cit. on p. 1580).

– (1994a). "Complex dynamics in higher dimension. I". *Astérisque* 222. Complex analytic methods in dynamical systems (Rio de Janeiro, 1992), pp. 5, 201–231. MR: 1285389 (cit. on p. 1580).

– (1994b). "Complex dynamics in higher dimensions". In: *Complex potential theory (Montreal, PQ, 1993)*. Vol. 439. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Notes partially written by Estela A. Gavosto. Kluwer Acad. Publ., Dordrecht, pp. 131–186. MR: 1332961 (cit. on pp. 1580, 1590).

– (1994c). "Complex dynamics in higher dimensions". In: *Complex potential theory (Montreal, PQ, 1993)*. Vol. 439. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Notes partially written by Estela A. Gavosto. Kluwer Acad. Publ., Dordrecht, pp. 131–186. MR: 1332961 (cit. on pp. 1580, 1585, 1590).

– (1995a). "Complex dynamics in higher dimension. II". In: *Modern methods in complex analysis (Princeton, NJ, 1992)*. Vol. 137. Ann. of Math. Stud. Princeton Univ. Press, Princeton, NJ, pp. 135–182. MR: 1369137 (cit. on p. 1580).

– (1995b). "Oka's inequality for currents and applications". *Math. Ann.* 301.3, pp. 399–419. MR: 1324517 (cit. on p. 1588).

Alexandre Freire, Artur Lopes, and Ricardo Mañé (1983). "An invariant measure for rational maps". *Bol. Soc. Brasil. Mat.* 14.1, pp. 45–62. MR: 736568 (cit. on p. 1590).

Shmuel Friedland and John Milnor (1989). "Dynamical properties of plane polynomial automorphisms". *Ergodic Theory Dynam. Systems* 9.1, pp. 67–99. MR: 991490 (cit. on p. 1590).

William Fulton (1998). *Intersection theory*. Second. Vol. 2. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, pp. xiv+470. MR: 1644323 (cit. on p. 1581).

Mikhaïl Gromov (1990). "Convex sets and Kähler manifolds". In: *Advances in differential geometry and topology*. World Sci. Publ., Teaneck, NJ, pp. 1–38. MR: 1095529 (cit. on p. 1586).

– (2003). "On the entropy of holomorphic maps". *Enseign. Math. (2)* 49.3-4, pp. 217–235. MR: 2026895 (cit. on p. 1585).

Vincent Guedj (2005). "Ergodic properties of rational mappings with large topological degree". *Ann. of Math. (2)* 161.3, pp. 1589–1607. MR: 2179389 (cit. on p. 1595).

Lars Hörmander (1990). *An introduction to complex analysis in several variables*. Third. Vol. 7. North-Holland Mathematical Library. North-Holland Publishing Co., Amsterdam, pp. xii+254. MR: 1045639 (cit. on p. 1582).

John H. Hubbard (1986). "The Hénon mapping in the complex domain". In: *Chaotic dynamics and fractals (Atlanta, Ga., 1985)*. Vol. 2. Notes Rep. Math. Sci. Engrg. Academic Press, Orlando, FL, pp. 101–111. MR: 858009 (cit. on p. 1580).

Katsunori Iwasaki and Takato Uehara (2010). "Periodic points for area-preserving birational maps of surfaces". *Math. Z.* 266.2, pp. 289–318. MR: 2678629 (cit. on p. 1594).

Mattias Jonsson and Paul Reschke (May 2015). "On the complex dynamics of birational surface maps defined over number fields". arXiv: 1505.03559 (cit. on p. 1594).

Vadim Yu. Kaloshin (2000). "Generic diffeomorphisms with superexponential growth of number of periodic orbits". *Comm. Math. Phys.* 211.1, pp. 253–271. MR: 1757015 (cit. on p. 1594).

Lucas Kaufmann (2017). "A Skoda-type integrability theorem for singular Monge-Ampère measures". *Michigan Math. J.* 66.3, pp. 581–594. MR: 3695353 (cit. on p. 1588).

Mkhail Ju. Ljubich (1983). "Entropy properties of rational endomorphisms of the Riemann sphere". *Ergodic Theory Dynam. Systems* 3.3, pp. 351–385. MR: 741393 (cit. on p. 1590).

Noboru Nakayama and De-Qi Zhang (2009). "Building blocks of étale endomorphisms of complex projective manifolds". *Proc. Lond. Math. Soc. (3)* 99.3, pp. 725–756. MR: 2551469 (cit. on p. 1586).

Keiji Oguiso (Dec. 2016a). "A criterion for the primitivity of a birational automorphism of a Calabi-Yau manifold and an application". To appear in *Math. Res. Lett.* arXiv: 1612.09016 (cit. on pp. 1583, 1586).

– (2016b). "Pisot Units, Salem Numbers, and Higher Dimensional Projective Manifolds with Primitive Automorphisms of Positive Entropy". *International Mathematics Research Notices* (cit. on p. 1583).

– (Apr. 2017). "A few explicit examples of complex dynamics of inertia groups on surfaces - a question of Professor Igor Dolgachev". To appear in *Transformation Groups*. arXiv: 1704.03142 (cit. on p. 1583).

Keiji Oguiso and Tuyen Trung Truong (2015). "Explicit examples of rational and Calabi-Yau threefolds with primitive automorphisms of positive entropy". *J. Math. Sci. Univ. Tokyo* 22.1, pp. 361–385. MR: 3329200 (cit. on p. 1583).

Alexander Russakovskii and Bernard Shiffman (1997). "Value distribution for sequences of rational mappings and complex dynamics". *Indiana Univ. Math. J.* 46.3, pp. 897–932. MR: 1488341 (cit. on pp. 1585, 1590).

Shuji Saito (1987). "General fixed point formula for an algebraic surface and the theory of Swan representations for two-dimensional local rings". *Amer. J. Math.* 109.6, pp. 1009–1042. MR: 919002 (cit. on p. 1594).

Nessim Sibony (1984). *Exposés à Orsay non publiés (1981)*. Cours UCLA (cit. on p. 1580).

– (1999). "Dynamique des applications rationnelles de $\mathbf{P}^k$". In: *Dynamique et géométrie complexes (Lyon, 1997)*. Vol. 8. Panor. Synthèses. Soc. Math. France, Paris, pp. ix–x, xi–xii, 97–185. MR: 1760844 (cit. on pp. 1580, 1581, 1589–1591, 1593).

Yum Tong Siu (1974). "Analyticity of sets associated to Lelong numbers and the extension of closed positive currents". *Invent. Math.* 27, pp. 53–156. MR: 0352516 (cit. on p. 1582).

Henri Skoda (1982). "Prolongement des courants, positifs, fermés de masse finie". *Invent. Math.* 66.3, pp. 361–376. MR: 662596 (cit. on p. 1584).

Johan Taflin (2011). "Equidistribution speed towards the Green current for endomorphisms of $\mathbb{P}^k$". *Adv. Math.* 227.5, pp. 2059–2081. MR: 2803794 (cit. on pp. 1590, 1591).

Henry de Thélin (2008). "Sur les exposants de Lyapounov des applications méromorphes". *Invent. Math.* 172.1, pp. 89–116. MR: 2385668 (cit. on p. 1593).

Tuyen Trung Truong (Jan. 2016). "Relative dynamical degrees of correspondence over a field of arbitrary characteristic". To appear in *J. Reine Angew. Math.* arXiv: 1501.01523 (cit. on p. 1585).

Claire Voisin (2002). *Hodge theory and complex algebraic geometry. I*. Vol. 76. Cambridge Studies in Advanced Mathematics. Translated from the French original by Leila Schneps. Cambridge University Press, Cambridge, pp. x+322. MR: 1967689 (cit. on p. 1582).

Duc-Viet Vu (Aug. 2016a). "Complex Monge-Ampère equation for measures supported on real submanifolds". To appear in *Math. Ann.* arXiv: 1608.02794 (cit. on p. 1588).

– (2016b). "Intersection of positive closed currents of higher bidegrees". *Michigan Math. J.* 65.4, pp. 863–872. MR: 3579190 (cit. on p. 1588).

Junyi Xie (2015). "Periodic points of birational transformations on projective surfaces". *Duke Math. J.* 164.5, pp. 903–932. MR: 3332894 (cit. on p. 1594).

Yosef Yomdin (1987). "Volume growth and entropy". *Israel J. Math.* 57.3, pp. 285–300. MR: 889979 (cit. on pp. 1585, 1595).

De-Qi Zhang (2009a). "Cohomologically hyperbolic endomorphisms of complex manifolds". *Internat. J. Math.* 20.7, pp. 803–816. MR: 2548399 (cit. on p. 1586).

– (2009b). "Dynamics of automorphisms on projective complex manifolds". *J. Differential Geom.* 82.3, pp. 691–722. MR: 2534992 (cit. on p. 1586).

Received 2017-12-02.

Tien-Cuong Dinh
Department of Mathematics
National University of Singapore
10 Lower Kent Ridge Road
Singapore 119076
matdtc@nus.edu.sg

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1601–1630)

THE INVERSE HULL OF 0-LEFT CANCELLATIVE SEMIGROUPS

Ruy Exel and Benjamin Steinberg

Abstract

Given a semigroup S with zero, which is left-cancellative in the sense that $st = sr \neq 0$ implies that $t = r$, we construct an inverse semigroup called the inverse hull of S, denoted $H(S)$. When S admits least common multiples, in a precise sense defined below, we study the idempotent semilattice of $H(S)$, with a focus on its spectrum. When S arises as the language semigroup for a subsift X on a finite alphabet, we discuss the relationship between $H(S)$ and several C*-algebras associated to X appearing in the literature.

1 Introduction

The goal of this note is to announce a series of results about semigroups, together with applications to C*-algebras, whose proofs will appear in later papers. The theory of semigroup C*-algebras has a long history, beginning with Coburn's work [Coburn 1967, 1969] where the C*-algebra of the additive semigroup of the natural numbers is studied in connection to Toeplitz operators. In [Murphy 1987] G. Murphy generalized this construction to the positive cone of an ordered group, and later to left cancellative semigroups ([Murphy 1991; Murphy 1994]). The C*-algebras studied by Murphy turned out to be too wild, even for nice looking semigroups such as $\mathbb{N} \times \mathbb{N}$, and this prompted Li [Li 2012] to introduce an alternative C*-algebra for a left cancellative semigroup. By definition a semigroup S is said to be left cancellative provided, for every r, s and t in S, one has that

$$st = sr \Rightarrow t = r. \tag{1.1}$$

Many interesting semigroups in the literature possess a zero element, namely an element 0 such that

$$s0 = 0s = 0,$$

MSC2010: primary 20M18; secondary 46L55.

for every s, and it is obvious that the presence of a zero prevents a semigroup from being left cancellative. In this work we focus on 0-*left cancellative* semigroups, meaning that (1.1) is required to hold only when the terms in its antecedent are supposed to be nonzero. This dramatically opens up the scope of applications including a wealth of interesting semigroups, such as those arising from subshifts and, more generaly, languages over a fixed alphabet. This also allows for the inclusion of categories and, more generally, the semigroupoids of [Exel 2008], once the multiplication is extended to all pairs of elements by setting undefined products to zero.

Starting with a 0-left cancellative semigroup S, the crucial point is first to build an inverse semigroup $\mathfrak{H}(S)$, which we call *the inverse hull* of S, by analogy with [Clifford and Preston 1961] and [Cherubini and Petrich 1987], from where one may invoke any of the now standard constructions of C*-algebras from inverse semigroups, such as the tight C*-algebra [Exel 2008] or Paterson's [Paterson 1999] universal C*-algebra. In fact this endeavor requires a lot more work regarding the passage from the original semigroup to its inverse hull, rather than the much better understood passage from there to the C*-algebras. Particularly demanding is the work geared towards understanding the idempotent semilattice of $\mathfrak{H}(S)$, which we denote by $\mathfrak{E}(S)$, as well as its spectrum. By a standard gadget $\mathfrak{E}(S)$ is put in correspondence with a subsemilattice of the power set of $S \setminus \{0\}$, whose members we call the *constructible sets*, by analogy with a similar concept relevant to Li's work in [Li 2012].

Central to the study of the spectrum of $\mathfrak{E}(S)$ is the notion of *strings*, which are motivated by the description of the unit space of graph groupoids in terms of paths in the graph.

Regarding the problem of understanding the spectrum of $\mathfrak{E}(S)$, we believe the present work represents only a modest beginning in a mammoth task lying ahead. This impression comes from situations in which similar spectra have been more or less understood, such as in [Exel and Laca 1999] and in [Dokuchaev and Exel 2017], illustrating the high degree of complexity one should expect.

It is only in our final section that we return to considering C*-algebras where we discuss, from the present perspective, how the Matsumoto and Carlsen-Matsumoto C*-algebras associated to a given subshift arise from the consideration of the inverse hull of the associated language semigroup. None of these correspond to the more well known tight or Paterson's universal C*-algebras, but we show that they instead arise from reductions of the Paterson groupoid to closed invariant subsets of its unit space which hitherto have not been identified.

2 Representations of semigroups

Let S be a semigroup, namely a nonempty set equipped with an associative operation.

A *zero element* for S is a (necessarily unique) element $0 \in S$, satisfying

$$s0 = 0s = 0, \quad \forall s \in S.$$

In what follows we will fix a semigroup S possessing a zero element.

Definition 2.1. Let Ω be any set. By a *representation of S on Ω* we shall mean any map

$$\pi : S \to \mathcal{I}(\Omega),$$

where $\mathcal{I}(\Omega)$ is the symmetric inverse semigroup[1] on Ω, such that

(i) π_0 is the empty map on Ω, and

(ii) $\pi_s \circ \pi_t = \pi_{st}$, for all s and t in S.

Given a set Ω, and any subset $X \subseteq \Omega$, let id_X denote the identity function on X, so that id_X an element of $E\big(\mathcal{I}(\Omega)\big)$, the idempotent semilattice of $\mathcal{I}(\Omega)$. One in fact has that

$$E\big(\mathcal{I}(\Omega)\big) = \{\mathrm{id}_X : X \subseteq \Omega\},$$

so we may identify $E\big(\mathcal{I}(\Omega)\big)$ with the meet semilattice $\mathcal{P}(\Omega)$ formed by all subsets of Ω.

Definition 2.2. Given a representation π of S, for every s in S we will denote the domain of π_s by F_s^π, and the range of π_s by E_s^π, so that π_s is a bijective mapping

$$\pi_s : F_s^\pi \to E_s^\pi.$$

When

$$\Omega = \Big(\bigcup_{s\in S} F_s^\pi\Big) \cup \Big(\bigcup_{s\in S} E_s^\pi\Big), \tag{2.3}$$

we will say that π is an *essential* representation. We will moreover let

$$f_s^\pi := \pi_s^{-1}\pi_s = \mathrm{id}_{E_s^\pi} \quad \text{and} \quad e_s^\pi := \pi_s\pi_s^{-1} = \mathrm{id}_{F_s^\pi}.$$

Let us fix, for the time being, a representation π of S on Ω. Whenever there is only one representation in sight we will drop the superscripts in F_s^π, E_s^π, f_s^π, and e_s^π, and adopt the simplified notations F_s, E_s, f_s, and e_s.

[1] The symmetric inverse semigroup on a set Ω is the inverse semigroup formed by all partially defined bijections on Ω.

The following may be proved easily.

Proposition 2.4. *Given s and t in S, one has that*

(i) $\pi_s e_t = e_{st}\pi_s$, *and*

(ii) $f_t \pi_s = \pi_s f_{ts}$.

Definition 2.5.

(i) The inverse subsemigroup of $\mathcal{I}(\Omega)$ generated by the set $\{\pi_s : s \in S\}$ will be denoted by $\mathcal{I}(\Omega, \pi)$.

(ii) Given any $X \in \mathcal{P}(\Omega)$ such that id_X belongs to $E\big(\mathcal{I}(\Omega,\pi)\big)$, we will say X is a *π-constructible subset.*

(iii) The collection of all π-constructible subsets of Ω will be denoted by $\mathcal{P}(\Omega,\pi)$. In symbols
$$\mathcal{P}(\Omega,\pi) = \big\{X \in \mathcal{P}(\Omega) : \mathrm{id}_X \in E\big(\mathcal{I}(\Omega,\pi)\big)\big\}.$$

Observe that E_s and F_s are π-constructible sets. For the special case of $s = 0$, we have $E_s = F_s = \varnothing$, so the empty set is π-constructible as well.

Since $\mathcal{P}(\Omega,\pi)$ corresponds to the idempotent semilattice of $\mathcal{I}(\Omega,\pi)$ by definition, it is clear that $\mathcal{P}(\Omega,\pi)$ is a semilattice, and in particular the intersection of two π-constructible sets is again π-constructible.

3 Cancellation properties for semigroups

Definition 3.1. Let S be a semigroup containing a zero element. We will say that S is 0-*left cancellative* if, for every $r, s, t \in S$,
$$st = sr \neq 0 \Rightarrow t = r,$$
and 0-*right cancellative* if
$$ts = rs \neq 0 \Rightarrow t = r.$$
If S is both 0-left cancellative and 0-right cancellative, we will say that S is 0-*cancellative*.

► In what follows we will fix a 0-left cancellative semigroup S. In a few occasions below we will also assume that S is 0-right cancellative.

For any s in S we will let
$$F_s = \{x \in S : sx \neq 0\},$$
and
$$E_s = \{y \in S : y = sx \neq 0, \quad \text{for some} \quad x \in S\}.$$

Observe that the correspondence "$x \to sx$" gives a map from F_s onto E_s, which is one-to-one by virtue of 0-left cancellativity.

Definition 3.2. For every s in S we will denote by θ_s the bijective mapping given by

$$\theta_s : x \in F_s \mapsto sx \in E_s.$$

Observing that 0 is neither in F_s, nor in E_s, we see that these are both subsets of

$$S' := S \setminus \{0\}, \tag{3.3}$$

so we may view θ_s as a partially defined bijection on S', which is to say that $\theta_s \in \mathcal{I}(S')$. We should also notice that when $s = 0$, both F_s and E_s are empty, so θ_s is the empty map.

Proposition 3.4. *The correspondence*

$$s \in S \mapsto \theta_s \in \mathcal{I}(S')$$

is a representation of S on S', henceforth called the regular representation of S.

Regarding the notations introduced in (2.2) in relation to the regular representation, notice that

$$F_s = F_s^\theta \quad \text{and} \quad E_s = E_s^\theta.$$

Definition 3.5. A semigroup S is called *right reductive* if it acts faithfully on the left of itself, that is, $sx = tx$ for all $x \in S$ implies $s = t$.

Of course every unital semigroup is right reductive. If S is a right reductive 0-left cancellative semigroup, then it embeds in $\mathcal{I}(S')$ via $s \mapsto \theta_s$.

Observe that if S is 0-right cancellative, then a single x for which $sx = rx$, as long as this is nonzero, is enough to imply that $s = t$. So, in a sense, right reductivity is a weaker version of 0-right cancellativity.

Definition 3.6. The *inverse hull* of S, henceforth denoted by $\mathfrak{H}(S)$, is the inverse subsemigroup of $\mathcal{I}(S')$ generated by the set $\{\theta_s : s \in S\}$. Thus, in the terminology of (2.5.i) we have

$$\mathfrak{H}(S) = \mathcal{I}(S', \theta).$$

The reader should compare the above with the notion of *inverse hull* defined in [Clifford and Preston 1961; Cherubini and Petrich 1987].

The collection of θ-constructible subsets of S' is of special importance to us, so we would like to give it a special notation:

Definition 3.7. The idempotent semilattice of $\mathfrak{H}(S)$, which we will tacitly identify with the semilattice of θ-constructible subsets of S', will be denoted by $\mathfrak{E}(S)$. Thus, in the terminology of (2.5.iii) we have

$$\mathfrak{E}(S) = \mathcal{P}(S', \theta).$$

It will be important to identify some properties of 0-left cancellative semigroups which will play a role later.

Proposition 3.8. *Let S be a 0-left cancellative semigroup.*

(i) *If e is an idempotent element in S and $s \in S \setminus \{0\}$, then $es \neq 0$ if and only if $es = s$, that is, $s \in eS \setminus \{0\}$.*

(ii) *If $s \in S \setminus \{0\}$, then $s \in sS$ if and only if $se = s$ for a necessarily unique idempotent e.*

(iii) *If $sS = S$ and S is right reductive, then S is unital and s is invertible.*

A semigroup S is said to have *right local units* if $S = SE(S)$, that is, for all $s \in S$, there exists an idempotent element e in S with $se = s$. A unital semigroup has right local units for trivial reasons. If S has right local units, then $sS = 0$ implies that $s = 0$. From Proposition 3.8 we obtain the following:

Corollary 3.9. *Let S be a 0-left cancellative semigroup. Then S has right local units if and only if $s \in sS$ for all $s \in S$.*

In a right reductive 0-left cancellative semigroup, the idempotents are orthogonal to each other:

Proposition 3.10. *Let S be a right reductive 0-left cancellative semigroup and suppose that $e \neq f$ are distinct idempotents of S. Then $ef = 0$.*

If S is a 0-left cancellative, right reductive semigroup with right local units, then for $s \in S \setminus \{0\}$, we denote by s^+ the unique idempotent with $ss^+ = s$. If S is unital, then $s^+ = 1$. If C is a left cancellative category, we can associate a semigroup $S(C)$ by letting $S(C)$ consist of the arrows of C together with a zero element 0. Products that are undefined in C are made zero in $S(C)$. It is straightforward to check that $S(C)$ is 0-left cancellative, right reductive and has right local units. If $f : c \to d$ is an arrow of C, then $f^+ = 1_c$.

4 Examples

If G is any group, and P is a subsemigroup of G, let $S = P \cup \{0\}$, where 0 is any element not belonging to P, with the obvious multiplication operation. Then S is clearly a 0-cancellative semigroup.

A more elaborate example is as follows: let Λ be any finite or infinite set, henceforth called the *alphabet*, and let Λ^+ be the free semigroup generated by Λ, namely the set of all finite words in Λ of positive length (and hence excluding the empty word), equipped with the multiplication operation given by concatenation.

Let L be a *language* on Λ, namely any nonempty subset of Λ^+. We will furthermore assume that L is *closed under prefixes and suffixes* in the sense that, for every α and β in Λ^+, one has

$$\alpha\beta \in L \Rightarrow \alpha \in L, \quad \text{and} \quad \beta \in L.$$

Define a multiplication operation on

$$S := L \cup \{0\}, \tag{4.1}$$

where 0 is any element not belonging to L, by

$$\alpha \cdot \beta = \begin{cases} \alpha\beta, & \text{if } \alpha, \beta \neq 0, \text{ and } \alpha\beta \in L, \\ 0, & \text{otherwise.} \end{cases}$$

One may then prove that S is a 0-cancellative semigroup.

One important special case of the above example is based on *subshifts*. Given an alphabet Λ, as above, let $X \subseteq \Lambda^{\mathbf{N}}$ be any nonempty subset invariant under the left shift, namely the mapping $\sigma : \Lambda^{\mathbf{N}} \to \Lambda^{\mathbf{N}}$ given by

$$\sigma(x_1x_2x_3\ldots) = x_2x_3x_4\ldots$$

Let $L \subseteq \Lambda^+$ be the *language of* X, namely the set of all finite words occuring in some infinite word belonging to X. Then L is clearly closed under prefixes and suffixes, and hence we are back in the conditions of the above example.

The fact that X is invariant under the left shift is indeed superfluous, as any nonempty subset $X \subseteq \Lambda^{\mathbb{N}}$ would lead to the same conclusion. However, languages arising from subshifts have been intensively studied in the literature, hence the motivation for the above example. The semigroup associated to the language of a shift was first studied in the early days of symbolic dynamics by Morse and Hedlund [Morse and Hedlund 1944].

Definition 4.2. (Munn [1964]) We will say that a semigroup S with zero is *categorical at zero* if, for every $r, s, t \in S$, one has that

$$rs \neq 0, \quad \text{and} \quad st \neq 0 \quad \Rightarrow \quad rst \neq 0.$$

The semigroup associated to a category is categorical at zero, whence the name. We may use the above ideas to produce a semigroup which is not categorical at zero: take any nonempty alphabet Λ, let L be the language consisting of all words of length at most two, and let $S = L \cup \{0\}$, as described above. If a, b and c, are members of Λ, we have that $abc = 0$, but ab and bc are nonzero, so S is not categorical at zero.

In contrast, notice that if S is the semigroup built as above from a Markov subshift, then S is easily seen to be categorical at zero.

Another interesting example is obtained from self-similar graphs [Exel and Pardo 2016]. Let G be a discrete group, $E = (E^0, E^1)$ be a graph with no sources, σ be an action of G by graph automorphisms on E, and

$$\varphi : G \times E^1 \to G$$

be a one-cocycle for the restriction of σ to E^1, which moreover satisfies

$$\sigma_{\varphi(g,e)}(x) = \sigma_g(x), \quad \forall g \in G, \quad \forall e \in E^1, \quad \forall x \in E^0.$$

Let E^* be the set of all finite paths[2] on E, including the vertices, which are seen as paths of length zero, and put

$$S = (E^* \times G) \cup \{0\}.$$

Given nonzero elements (α, g) and (β, h) in S, define

$$(\alpha, g)(\beta, h) = \begin{cases} \big(\alpha\sigma_g(\beta), \varphi(g, \beta)h\big), & \text{if } s(\alpha) = r\big(\sigma_g(\beta)\big), \\ 0, & \text{otherwise.} \end{cases}$$

One may then show that S is a categorical at zero, 0-left cancellative semigroup, which is 0-right cancellative if and only if (G, E) is pseudo-free in the sense of Exel and Pardo [ibid., Proposition5.6].

5 Least common multiples

We now wish to introduce a special class of semigroups, but for this we must first consider the question of divisibility.

Definition 5.1. Given s and t in a semigroup S, we will say that s *divides* t, in symbols

$$s \mid t,$$

[2] We adopt the functorial point of view so a path is a sequence $e_1 \dots e_n$ of edges, such that $s(e_i) = r(e_{i+1})$, for all i, as opposed to the also very popular "$r(e_i) = s(e_{i+1})$".

or that t is a *multiple* of s, when either $s = t$, or there is some u in S such that $su = t$. In other words,

$$t \in \{s\} \cup sS.$$

We observe that division is a reflexive and transitive relation, so it may be seen as a (not necessarily anti-symmetric) order relation.

For the strict purpose of simplifying the description of the division relation, regardless of whether or not S is unital, we shall sometimes employ the unitized semigroup

$$\tilde{S} := S \cup \{1\},$$

where 1 is any element not belonging to S, made to act like a unit for S. For every s and t in S we therefore have that

$$s \mid t \iff \exists u \in \tilde{S}, \quad su = t. \tag{5.2}$$

Having enlarged our semigroup, we might as well extend the notion of divisibility:

Definition 5.3. Given v and w in $\tilde{S}$, we will say that $v \mid w$ when there exists some u in $\tilde{S}$, such that $vu = w$.

Notice that if v and w are in S, then the above notion of divisibility coincides with the previous one by (5.2). Analysing the new cases where this extended divisibility may or may not apply, notice that:

$$\begin{aligned} \forall w \in \tilde{S}, &\qquad 1 \mid w, \\ \forall v \in \tilde{S}, &\quad v \mid 1 \iff v = 1. \end{aligned} \tag{5.4}$$

The introduction of $\tilde{S}$ brings with it several pitfalls, not least because $\tilde{S}$ might not be 0-left cancellative: when S already has a unit, say 1_S, then in the identity "$s1_S = s1$", we are not allowed to left cancel s, since $1_S \neq 1$. One should therefore be very careful when working with $\tilde{S}$.

Definition 5.5. Let S be a semigroup and let $s, t \in S$. We will say that an element $r \in S$ is a *least common multiple* for s and t when

(i) $sS \cap tS = rS$,

(ii) both s and t divide r.

Observe that when S has right local units then $r \in rS$, by (3.9), and hence condition (5.5.i) trivially implies (5.5.ii), so the former condition alone suffices to define least common multiples. However in a unitless semigroup condition (5.5.i) may hold while (5.5.ii) fails. Nevertheless, when $sS \cap tS = \{0\}$, then 0 is a least common multiple for s and t because s and t always divide 0.

Regardless of the existence of right local units, notice that condition (5.5.ii) holds if and only if $r\tilde{S} \subseteq s\tilde{S} \cap t\tilde{S}$, and therefore one has that r is a least common multiple for s and t if and only if

$$sS \cap tS = rS \subseteq r\tilde{S} \subseteq s\tilde{S} \cap t\tilde{S}. \tag{5.6}$$

One may of course think of alernative definitions for the concept of least common multiples, fiddling with the above ideas in many different ways. However (5.5) seems to be the correct choice, at least from the point of view of the theory we are about to develop.

Definition 5.7. We shall say that a semigroup S *admits least common multiples* if there exists a least common multiple for each pair of elements of S.

The language semigroup of (4.1) is easily seen to be an example of a semigroup admitting least common multiples.

Another interesting example is obtained from the quasi-lattice ordered groups of [Nica 1992], which we would now like to briefly describe.

Given a group G and a unital subsemigroup $P \subseteq G$, one defines a partial order on G via

$$x \le y \iff x^{-1}y \in P.$$

The quasi-lattice condition says that, whenever elements x and y in G admit a common upper bound, namely an element z in G such that $z \ge x$ and $z \ge y$, then there exists a least common upper bound, usually denoted $x \vee y$.

Under this situation, consider the semigroup $S = P \cup \{0\}$, obtained by adjoining a zero to P. Then, for every nonzero s in S, i.e. for s in P, one has that

$$sS = \{x \in P : x \ge s\} \cup \{0\},$$

so that the multiples of s are precisely the upper bounds of s in P, including zero.

If t is another nonzero element in S, one therefore has that s and t admit a nonzero commun multiple if and only if s and t admit a common upper bound in P, in which case $s \vee t$ is a least common multiple of s and t.

On the other hand, when s and t admit no common upper bound, then obviously $s \vee t$ does not exist, but still s and t admit a least common multiple in S, namely 0.

Summarizing we have the following:

Proposition 5.8. *Let (G, P) be a quasi-lattice ordered group. Then the semigroup $S := P \cup \{0\}$ admits least common multiples.*

► From this point on we will fix a 0-left cancellative semigroup S admitting least common multiples.

Proposition 5.9. *Given u and v in $\tilde{S}$, there exists w in $\tilde{S}$ such that*

(i) $uS \cap vS = wS$,

(ii) *both u and v divide w.*

Proof. When u and v lie in S, it is enough to take w to be a (usual) least common multiple of u and v. On the other hand, if $u = 1$, one takes $w = v$, and if $v = 1$, one takes $w = u$. □

Based on the above we may extend the notion of least common multiples to $\tilde{S}$, as follows:

Definition 5.10. Given u and v in $\tilde{S}$, we will say that an element w in $\tilde{S}$ is a least common multiple of u and v, provided (5.9.i-ii) hold. In the exceptional case that $u = v = 1$, only $w = 1$ will be considered to be a least common multiple of u and v, even though there might be another w in S satisfying (5.9.i-ii).

It is perhaps interesting to describe the exceptional situation above, where we are arbitrarily prohibiting by hand that an element of S be considered as a least common multiple of 1 and itself, even though it would otherwise satisfy all of the required properties. If $w \in S$ is such an element, then

$$wS = 1S \cap 1S = S,$$

so, in case we also assume that S is right-reductive, we deduce from (3.8.iii) that S is unital and w is invertible. Thus, in hindsight it might not have been such a good idea to add an external unit to S after all.

On the other hand, when s and t lie in S, it is not hard to see that any least common multiple of s and t in the new sense of (5.10) must belong to S, and hence it must also be a least common multiple in the old sense of (5.5).

Given a representation π of S on a set Ω, we will now concentrate our attention on giving a concrete description for the inverse semigroup $\mathcal{I}(\Omega, \pi)$ defined in (2.5), provided π satisfies certain special properties, which we will now describe.

Initially notice that if $s \mid r$, then the range of π_r is contained in the range of π_s because either $r = s$, or $r = su$, for some u in S, in which case $\pi_r = \pi_s \pi_u$. So, using the notation introduced in (2.2),

$$E_r^\pi \subseteq E_s^\pi.$$

When r is a common multiple of s and t, it then follows that

$$E_r^\pi \subseteq E_s^\pi \cap E_t^\pi.$$

Definition 5.11. A representation π of S is said to *respect least common multiples* if, whenever r is a least common multiple of elements s and t in S, one has that $E_r^\pi = E_s^\pi \cap E_t^\pi$.

As an example, notice that the regular representation of S, defined in (3.4), satisfies the above condition since the fact that $rS = sS \cap tS$ implies that

$$E_r^\theta = rS \setminus \{0\} = (sS \cap tS) \setminus \{0\} = (sS \setminus \{0\}) \cap (tS \setminus \{0\}) = E_s^\theta \cap E_t^\theta. \tag{5.12}$$

► From now on we will fix a representation π of S on a set Ω, assumed to respect least common multiples. Since π will be the only representation considered for a while, we will use the simplified notations F_s, E_s, f_s, and e_s.

There is a cannonical way to extend π to $\tilde{S}$ by setting

$$F_1 = E_1 = \Omega, \quad \text{and} \quad \pi_1 = \mathrm{id}_\Omega .$$

It is evident that π remains a multiplicative map after this extension. Whenever we find it convenient we will therefore think of π as defined on $\tilde{S}$ as above. We will accordingly extend the notations f_s and e_s to allow for any s in $\tilde{S}$, in the obvious way.

Proposition 5.13. *Let π be a representation of S on a set Ω. If π respects least common multiples then so does its natural extension to $\tilde{S}$. Precisely, if u and v are elements of $\tilde{S}$, and if $w \in \tilde{S}$ is a least common multiple of u and v, then $E_w = E_u \cap E_v$.*

Definition 5.14. Given a representation π of S, and given any nonempty finite subset $\Lambda \subseteq \tilde{S}$, we will let

$$F_\Lambda^\pi = \bigcap_{u \in \Lambda} F_u^\pi, \quad \text{and} \quad f_\Lambda^\pi = \prod_{u \in \Lambda} f_u^\pi .$$

When there is only one representation of S in sight, as in the present moment, we will drop the superscripts and use the simplified notations F_Λ and f_Λ.

We should remark that, since each f_s is the identity map on F_s, one has that f_Λ is the identity map on F_Λ.

Also notice that, since $f_1 = \mathrm{id}_\Omega$, the presence of 1 in Λ has no effect in the sense that $f_\Lambda = f_{\Lambda \cup \{1\}}$, for every Λ. Thus, whenever convenient we may assume that $1 \in \Lambda$.

As already indicated we are interested in obtaining a description for the inverse semigroup $\mathcal{I}(\Omega, \pi)$. In that respect it is interesting to observe that most elements of the form f_Λ belong to $\mathcal{I}(\Omega, \pi)$, but there is one exception, namely when $\Lambda = \{1\}$. In this case we have

$$f_{\{1\}} = \mathrm{id}_\Omega,$$

which may or may not lie in $\mathcal{I}(\Omega, \pi)$. However, when $\Lambda \cap S \neq \varnothing$, then surely

$$f_\Lambda \in \mathcal{I}(\Omega, \pi). \tag{5.15}$$

We should furthermore remark that, whenever we are looking at a term of the form $\pi_u f_\Lambda \pi_v^{-1}$, we may assume that $u, v \in \Lambda$, because

$$\begin{aligned} \pi_u f_\Lambda \pi_v^{-1} &= \pi_u \pi_u^{-1} \pi_u f_\Lambda \pi_v^{-1} \pi_v \pi_v^{-1} = \\ &= \pi_u f_u f_\Lambda f_v \pi_v^{-1} = \pi_u f_{\{u\}\cup\Lambda\cup\{v\}} \pi_v^{-1}, \end{aligned} \tag{5.16}$$

so Λ may be replaced by $\{u\} \cup \Lambda \cup \{v\}$ without altering the above term. Moreover, as in (5.15), observe that when $\Lambda \cap S \neq \varnothing$, then

$$\pi_u f_\Lambda \pi_v^{-1} \in \mathcal{I}(\Omega, \pi).$$

Theorem 5.17. *Let S be a 0-left cancellative semigroup admitting least common multiples. Also let π be a representation of S on a set Ω, assumed to respect least common multiples. Then the generated inverse semigroup $\mathcal{I}(\Omega, \pi)$ is given by*

$$\mathcal{I}(\Omega, \pi) = \Big\{\pi_u f_\Lambda \pi_v^{-1} : \Lambda \subseteq \tilde{S} \text{ is finite, } \Lambda \cap S \neq \varnothing, \text{ and } u, v \in \Lambda\Big\}.$$

With this we may describe the constructible sets in a very concrete way.

Proposition 5.18. *Under the assumptions of (5.17), the π-constructible subsets of Ω are precisely the sets of the form*

$$X = \pi_u(F_\Lambda),$$

where $\Lambda \subseteq \tilde{S}$ is a finite subset, $\Lambda \cap S \neq \varnothing$, and $u \in \Lambda$.

Recalling that the regular representation of S respects least common multiples, our last two results apply to give:

Corollary 5.19. *Let S be a 0-left cancellative semigroup admitting least common multiples. Then*

$$\mathfrak{H}(S) = \Big\{\theta_u f_\Lambda \theta_v^{-1} : \Lambda \subseteq \tilde{S} \quad \text{is finite,} \quad \Lambda \cap S \neq \varnothing, \quad \text{and} \quad u, v \in \Lambda\Big\},$$

and

$$\mathfrak{E}(S) = \Big\{u F_\Lambda : \Lambda \subseteq \tilde{S} \quad \text{is finite,} \quad \Lambda \cap S \neq \varnothing, \quad \text{and} \quad u \in \Lambda\Big\}.$$

6 Strings

This section is intended to introduce a device which will be highly useful in the study of the spectrum of $\mathfrak{E}(S)$.

► Throughout this section S will be a fixed 0-left cancellative semigroup.

Definition 6.1. A nonempty subset $\sigma \subseteq S$ is said to be a *string* in S, if

(i) $0 \notin \sigma$,

(ii) for every s and t in S, if $s \mid t$, and $t \in \sigma$, then $s \in \sigma$,

(iii) for every s_1 and s_2 in σ, there is some s in σ such that $s_1 \mid s$, and $s_2 \mid s$.

An elementary example of a string is the set of divisors of any nonzero element s in S, namely,

$$\delta_s = \{t \in S : t \mid s\}. \tag{6.2}$$

Strings often contain many elements, but there are some exceptional strings consisting of a single semigroup element. To better study these it is useful to introduce some terminology.

Definition 6.3. Given s in S we will say that s is:

(i) *prime*, if the only divisor of s is s, itself, or, equivalently, if $\delta_s = \{s\}$,

(ii) *irreducible*, if there are no two elements x and y in S such that $s = xy$, or, equivalently, if $s \notin S^2$.

It is evident that any irreducible element is prime, but there might be prime elements which are not irreducible. For example, in the semigroup $S = \{0, s, e\}$, with multiplication table given by

$\times$	0	e	s
0	0	0	0
e	0	e	0
s	0	s	0

one has that s is prime but not irreducible because $s = se \in S^2$.

Definition 6.4. A singleton $\{s\}$ is a string if and only if s is prime.

Proof. If s is prime then the singleton $\{s\}$ coincides with δ_s, and hence it is a string. Conversely, supposing that $\{s\}$ is a string, we have by (6.1.ii) that $\delta_s \subseteq \{s\}$, from where it follows that s is prime. □

Definition 6.5. The set of all strings in S will be denoted by $S^\star$.

From now on our goal will be to define an action of S on $S^\star$.

Proposition 6.6. *Let σ be a string in S, and let $r \in S$. Then*

(i) *if* 0 *is not in* $r\sigma$, *one has that*

$$r * \sigma := \{t \in S : t \mid rs, \quad \textit{for some} \ \ s \in \sigma\}$$

is a string whose intersection with rS *is nonempty.*

(ii) *If* σ *is a string whose intersection with* rS *is nonempty, then*

$$r^{-1} * \sigma := \{t \in S : rt \in \sigma\}$$

is a string, and 0 *is not in* $r(r^{-1} * \sigma)$.

It should be noted that, under the assumptions of (6.6.i), one has that

$$r\sigma \subseteq r * \sigma, \tag{6.7}$$

and in fact $r * \sigma$ is the hereditary closure of $r\sigma$ relative to the order relation given by division.

We may then define a representation of S on the set $S^\star$ of all strings in S, as follows:

Proposition 6.8. *For each* r *in* S, *put*

$$F^\star_r = \{\sigma \in S^\star : r\sigma \not\ni 0\}, \quad \textit{and} \quad E^\star_r = \{\sigma \in S^\star : \sigma \cap rS \neq \varnothing\}.$$

Also let

$$\theta^\star_r : F^\star_r \to E^\star_r$$

be defined by $\theta^\star_r(\sigma) = r * \sigma$, *for every* $\sigma \in F^\star_r$. *Then:*

(i) $\theta^\star_r$ *is bijective, and its inverse is the mapping defined by*

$$\sigma \in E^\star_r \mapsto r^{-1} * \sigma \in F^\star_r.$$

(ii) *Viewing* $\theta^\star$ *as a map from* S *to* $\mathcal{I}(S^\star)$, *one has that* $\theta^\star$ *is a representation of* S *on* $S^\star$.

Useful alternative characterizations of $F^\star_r$ and $E^\star_r$ are as follows:

Proposition 6.9. *Given* r *in* S, *and given any string* σ *in* $S^\star$, *one has that:*

(i) $\sigma \in F^\star_r \Leftrightarrow \sigma \subseteq F^\theta_r$,

(ii) $\sigma \in E^\star_r \Leftrightarrow \sigma \cap E^\theta_r \neq \varnothing$,

(iii) $\sigma \in E^\star_r \Rightarrow r \in \sigma$. *In addition, the converse holds provided* $r \in rS$ *(e.g. if* S *has right local units).*

Recall from (5.19) that, when S has least common multiples, every θ-constructible subset of S' has the form $\theta_u(F_\Lambda^\theta)$, where $\Lambda \subseteq \tilde{S}$ is finite, $\Lambda \cap S \neq \varnothing$, and $u \in \Lambda$. By analogy this suggests that it might also be useful to have a characterization of $\theta_u^\star(F_\Lambda^\star)$ along the lines of (6.9).

Proposition 6.10. *Let Λ be a finite subset of $\tilde{S}$ having a nonempty intersection with S, and let $u \in \Lambda$. Then $\theta_u^\star(F_\Lambda^\star)$ consists precisely of the strings σ such that*

$$\varnothing \neq \sigma \cap E_u^\theta \subseteq \theta_u(F_\Lambda^\theta).$$

After (6.8) we now have two natural representations of S, namely the regular representation θ acting on S', and $\theta^\star$ acting on $S^\star$.

Proposition 6.11. *The map*

$$\delta : s \in S' \mapsto \delta_s \in S^\star,$$

where δ_s is defined in (6.2), is covariant relative to θ and $\theta^\star$.

Observe that the union of an increasing family of strings is a string, so any string is contained in a maximal one by Zorn's Lemma.

Definition 6.12. The subset of $S^\star$ formed by all maximal strings will be denoted by S^∞.

Our next result says that S^∞ is invariant under $\theta^\star$.

Proposition 6.13. *For every r in S, and for every maximal string σ in $F_r^\star$, one has that $\theta_r^\star(\sigma)$ is maximal.*

Observe that the above result says that S^∞ is invariant under each $\theta_r^\star$, but not necessarily under $\theta_r^{\star\,-1}$.

An example to show that S^∞ may indeed be non invariant under $\theta_r^{\star\,-1}$ is as follows. Consider the language L on the alphabet $\Sigma = \{a, b\}$ given by

$$L = \{a, b, aa, ba\}.$$

Then, $\sigma = \{b, ba\}$ is a maximal string, while $\theta_b^{\star\,-1} = \{a\}$ is not maximal.

Definition 6.14. By a representation of a given inverse semigroup $\mathcal{S}$ on a set Ω we shall mean any map

$$\rho : \mathcal{S} \to \mathcal{I}(\Omega),$$

such that $\rho(0)$ is the empty map, and for every s and t in $\mathcal{S}$, one has that $\rho(st) = \rho(s)\rho(t)$, and $\rho(s^{-1}) = \rho(s)^{-1}$.

Proposition 6.15. *Let S be a 0-left cancellative semigroup admitting least common multiples. Then there exists a unique representation ρ of $\mathfrak{H}(S)$ on $S^\star$, such that the following diagram commutes.*

$$\begin{tikzcd}
S \arrow[rr, "\theta^\star"] \arrow[dr, "\theta"'] & & \mathfrak{I}(S^*) \\
& \mathfrak{H}(S) \arrow[ur, "\rho"'] &
\end{tikzcd}$$

Observing that a homomorphism of inverse semigroups must restrict to the corresponding idempotent semilattice s, we obtain the following:

Corollary 6.16. *Let S be a 0-left cancellative semigroup admitting least common multiples. Then there exists a semilattice representation*

$$\varepsilon : \mathfrak{E}(S) \to \mathcal{P}(S^\star),$$

such that

$$\varepsilon\big(\theta_u(F^\theta_\Lambda)\big) = \theta^\star_u(F^\star_\Lambda),$$

whenever Λ is a finite subset of $\tilde{S}$ intersecting S nontrivially, and $u \in \Lambda$.

Observing that $E^\theta_r = \theta_r(F^\theta_r)$, notice that

$$\varepsilon(E^\theta_r) = \theta^\star_r(F^\star_r) = E^\star_r.$$

7 The spectrum of the semilattice of constructible sets

Let us now fix a 0-left cancellative semigroup S admitting least common multiples. It is our goal in this section to study the spectrum of $\mathfrak{E}(S)$.

Recall that if $\mathcal{E}$ is a semilattice with zero, the spectrum of $\mathcal{E}$ is the set of all semilattice homomorphisms $\varphi : \mathcal{E} \to \{0, 1\}$, such that $\varphi(0) = 0$. Here $\{0, 1\}$ is equipped with its standard semilattice structure $0 < 1$.

Considering the representation

$$\varepsilon : \mathfrak{E}(S) \to \mathcal{P}(S^\star),$$

introduced in (6.16), and given $\sigma \in S^\star$, set

$$\varphi_\sigma : X \in \mathfrak{E}(S) \mapsto [\sigma \in \varepsilon(X)] \in \{0, 1\}.$$

It is clear that φ_σ is a semilattice homomorphism, so it is a character as long as it is nonzero.

The question of whether or not φ_σ is nonzero evidently boils down to the existence of some constructible set X for which $\sigma \in \varepsilon(X)$. By (5.19) it is easy to see that for every constructible set X, there is some r in S such that $X \subseteq E_r$, or $X \subseteq F_r$. Therefore $\varphi_\sigma = 0$ if and only if σ is never in any $E_r^\star$ nor in any $F_r^\star$, that is,

$$\sigma \nsubseteq F_r^\theta, \quad \text{and} \quad \sigma \cap E_r^\theta = \varnothing,$$

for every r in S, by (6.9).

The second condition above implies that every element in σ is irreducible, so it necessarily follows that σ is a singleton, say $\sigma = \{s\}$, where s is irreducible. In turn, the first condition above implies that s lies in no F_r^θ, whence $Ss = 0$.

Definition 7.1. (i) An element s in S will be called *degenerate* if s is irreducible and $Ss = \{0\}$.

(ii) A string σ will be called *degenerate* if $\sigma = \{s\}$, where s is a degenerate element.

(iii) The set of all non-degenerate strings will be denoted by $S_\sharp^\star$.

(iv) For every non-degenerate string σ, we shall denote by φ_σ the character of $\mathfrak{E}(S)$ given by

$$\varphi_\sigma(X) = [\sigma \in \varepsilon(X)], \quad \forall X \in \mathfrak{E}(S).$$

Suppose we are given φ_σ and we want to recover σ from φ_σ. In the special case in which S has right local units, we have that

$$\varphi_\sigma(E_s^\theta) = 1 \iff \sigma \in \varepsilon(E_s^\theta) = E_s^\star \overset{(6.9.\text{iii})}{\iff} s \in \sigma, \tag{7.2}$$

so σ is recovered as the set $\{s \in S : \varphi_\sigma(E_s^\theta) = 1\}$. Without assuming right local units, the last part of (7.2) cannot be trusted, but it may be replaced with

$$\cdots \overset{(6.9.\text{ii})}{\iff} \sigma \cap E_s^\theta \neq \varnothing, \tag{7.3}$$

so we at least know which E_s^θ have a nonempty intersection with σ.

Proposition 7.4. *Given any string σ, let the interior of σ be defined by*

$$\mathring{\sigma} := \{s \in S : \exists x \in S,\ sx \in \sigma\}.$$

Then

$$\mathring{\sigma} = \{s \in S : \varphi_\sigma(E_s^\theta) = 1\}.$$

Given any character φ of $\mathfrak{E}(S)$, regardless of whether or not it is of the form φ_σ as above, we may still consider the set

$$\sigma_\varphi := \{s \in S : \varphi(E_s^\theta) = 1\}, \tag{7.5}$$

so that, when $\varphi = \varphi_\sigma$, we get $\sigma_\varphi = \mathring{\sigma}$.

Proposition 7.6. *If φ is any character of $\mathfrak{E}(S)$, and σ_φ is nonempty, then σ_φ is a string closed under least common multiples.*

Based on (7.1.iv) we may define a map from the set of all non-degenerate strings to $\hat{\mathfrak{E}}(S)$, the spectrum of $\mathfrak{E}(S)$, by

$$\Phi : \sigma \in S_\sharp^\star \mapsto \varphi_\sigma \in \hat{\mathfrak{E}}(S), \tag{7.7}$$

but if we want the dual correspondence suggested by (7.5), namely

$$\varphi \mapsto \sigma_\varphi, \tag{7.8}$$

to give a well defined map from $\hat{\mathfrak{E}}(S)$ to $S^\star$, we need to worry about its domain because we have not checked that σ_φ is always nonempty, and hence σ_φ may fail to be a string. The appropriate domain is evidently given by the set of all characters φ such that σ_φ is nonempty but, before we formalize this map, it is interesting to introduce a relevant sub-semilattice of $\mathfrak{E}(S)$.

Proposition 7.9. *The subset of $\mathfrak{E}(S)$ given by*[3]

$$\mathfrak{E}_1(S) = \{sF_\Lambda^\theta,\ \Lambda \subseteq S \text{ is finite, and } s \in \Lambda\},$$

is an ideal of $\mathfrak{E}(S)$. Moreover, for every X in $\mathfrak{E}(S)$, one has that X lies in $\mathfrak{E}_1(S)$ if and only if $X \subseteq E_s^\theta$, for some s in S.

Whenever J is an ideal in a semilattice $\mathcal{E}$, there is a standard inclusion

$$\varphi \in \hat{J} \mapsto \tilde{\varphi} \in \hat{E},$$

where, for every x in E, one has that $\tilde{\varphi}(x) = 1$, if and only if there exists some y in J with $y \leq x$, and $\varphi(y) = 1$. The next result is intended to distinguish the elements of the copy of $\hat{\mathfrak{E}}_1(S)$ within $\hat{\mathfrak{E}}(S)$ given by the above correspondence.

Proposition 7.10. *Let φ be a character on $\mathfrak{E}(S)$. Then the following are equivalent:*

[3] This should be contrasted with (5.19), where the general form of an element of $\mathfrak{E}(S)$ is uF_Λ^θ, where u is in $\tilde{S}$, rather than S.

(i) $\varphi \in \hat{\mathfrak{E}}_1(S)$,

(ii) $\varphi(E_s^\theta) = 1$, *for some* s *in* S,

(iii) σ_φ *is nonempty, and hence it is a string by* (7.6).

If S admits right local units, then $\mathfrak{E}_1(S) = \mathfrak{E}(S)$, so σ_φ is a string for every character $\varphi \in \hat{\mathfrak{E}}(S)$.

The vast majority of non-degenerate strings σ lead to a character φ_σ belonging to $\hat{\mathfrak{E}}_1(S)$, but there are exceptions.

Proposition 7.11. *If* σ *is a non-degenerate string in* $S_\sharp^\star$ *then* φ_σ *does not belong to* $\hat{\mathfrak{E}}_1(S)$ *if and only if* $\sigma = \{s\}$, *where* s *is an irreducible element of* S.

By (7.10) we have that the largest set of characters on which the correspondence described in (7.8) produces a bona fide string is precisely $\hat{\mathfrak{E}}_1(S)$, so we may now formaly introduce the map suggested by that correspondence.

Definition 7.12. We shall let

$$\Sigma : \hat{\mathfrak{E}}_1(S) \mapsto S^\star,$$

be the map given by

$$\Sigma(\varphi) = \sigma_\varphi = \{s \in S : \varphi(E_s^\theta) = 1\}, \quad \forall \varphi \in \hat{\mathfrak{E}}_1(S).$$

For every string σ, excluding the exceptional ones discussed in (7.11), we then have that

$$\Phi(\sigma) = \varphi_\sigma \in \hat{\mathfrak{E}}_1(S),$$

and

$$\Sigma\big(\Phi(\sigma)\big) = \mathring{\sigma}, \tag{7.13}$$

by (7.4).

Definition 7.14. A string σ in S will be termed *open* if $\sigma = \mathring{\sigma}$.

The nicest situation is for open strings:

Proposition 7.15. *If* σ *is an open string, then*

(i) σ *is non-degenerate,*

(ii) $\Phi(\sigma) \in \hat{\mathfrak{E}}_1(S)$, *and*

(iii) $\Sigma\big(\Phi(\sigma)\big) = \sigma$.

Given that the composition $\Sigma \circ \Phi$ is so well behaved for open strings, we will now study the reverse composition $\Phi \circ \Sigma$ on a set of characters related to open strings.

Definition 7.16. A character φ in $\hat{\mathfrak{E}}(S)$ will be called an *open* character if σ_φ is a (nonempty) open string.

We remark that every open character belongs to $\hat{\mathfrak{E}}_1(S)$ by (7.10), although not all characters in $\hat{\mathfrak{E}}_1(S)$ are open.

By (7.15) it is clear that φ_σ is an open character for every open string σ.

If S admits right local units, we have seen that every string in $S^\star$ is open, and also that σ_φ is a string for every character. Therefore every character in $\hat{\mathfrak{E}}(S)$ is open.

The composition $\Phi \circ \Sigma$ is not as well behaved as the one discussed in (7.15), but there is at least some relationship between a character φ and its image under $\Phi \circ \Sigma$, as we shall now see.

Proposition 7.17. *Given any open character φ, one has that*

$$\varphi \leq \Phi\big(\Sigma(\varphi)\big).$$

This leads us to one of our main results.

Theorem 7.18. *Let S be a 0-left cancellative semigroup admitting least common multiples. Then, for every open, maximal string σ over S, one has that φ_σ is an ultra-character.*

The previous result raises the question as to whether σ_φ is a maximal string for every ultra-character φ, but this is not true in general. Consider for example the unital semigroup

$$S = \{1, a, aa, 0\},$$

in which $a^3 = 0$. The θ-constructible subsets of S are precisely

$E^\theta_1 = F^\theta_1 = \{1, a, aa\}$		
$F^\theta_a = \{1, a\}$	$E^\theta_a = aF^\theta_a = \{a, aa\}$	
$F^\theta_{aa} = \{1\}$	$aF^\theta_{aa} = \{a\}$	$E^\theta_{aa} = aaF^\theta_{aa} = \{aa\}$

List of θ-constructible sets

and there are three strings over S, namely

$\delta_1 = \{1\}$	$\delta_a = \{1, a\}$	$\delta_{aa} = \{1, a, aa\}$

Since the correspondence $s \mapsto \delta_s$ is a bijection from S' to $S^\star$, we see that $\theta^\star$ is isomorphic to θ, and in particular the $\theta^\star$-constructible subsets of $S^\star$, listed below, mirror the θ-constructible ones.

$E_1^\star = F_1^\star = \{\delta_1, \delta_a, \delta_{aa}\}$		
$F_a^\star = \{\delta_1, \delta_a\}$	$E_a^\star = aF_a^\star = \{\delta_a, \delta_{aa}\}$	
$F_{aa}^\theta = \{\delta_1\}$	$aF_{aa}^\star = \{\delta_a\}$	$E_{aa}^\theta = aaF_{aa}^\star = \{\delta_{aa}\}$

List of θ^*-constructible sets

Observe that the string $\sigma := \delta_a = \{1, a\}$ is a proper subset of the string $\{1, a, aa\}$, and hence σ is not maximal. But yet notice that φ_σ is an ultra-character, since $\{\delta_a\}$ is a minimal[4] member of $\mathcal{P}(S^\star, \theta^\star)$. We thus get an example of

"A string σ which is not maximal but such that φ_σ is an ultra-character."

On the other hand, since $\sigma = \sigma_{\varphi_\sigma}$, this also provides an example of

"An ultra-character φ such that σ_φ is not maximal."

This suggests the need to single out the strings which give rise to ultra-characters:

Definition 7.19. We will say that a string σ is *quasi-maximal* whenever φ_σ is an ultra-character. The set of all quasi-maximal strings will be denoted by $S^\propto$.

Adopting this terminology, the conclusion of (7.18) states that every open, maximal string is quasi-maximal.

Theorem 7.20. *Let S be a 0-left cancellative semigroup admitting least common multiples. Then, every open ultra-character on $\mathfrak{E}(S)$ is of the form φ_σ for some open, quasi-maximal string σ.*

The importance of quasi-maximal strings evidenced by the last result begs for a better understanding of such strings. While we are unable to provide a complete characterization, we can at least exhibit some further examples beyond the maximal ones.

To explain what we mean, recalll from (6.9.i) that a string σ belongs to some $F_\Lambda^\star$ if and only if σ is contained in F_Λ^θ. It is therefore possible that σ is maximal among all strings contained in F_Λ^θ, and still not be a maximal string. An example is the string $\{1, a\}$ mentioned above, which is maximal within F_a^θ, but not maximal in the strict sense of the word.

[4]Whenever e_0 is a nonzero minimal element of a semilattice E, the character $\varphi(e) = [e_0 \leq e]$ is an ultra-character.

Proposition 7.21. *Let Λ be a nonempty finite subset of S and suppose that σ is an open string such that $\sigma \subseteq F^\theta_\Lambda$. Suppose moreover that σ is maximal among the strings contained in F^θ_Λ, in the sense that for every string μ, one has that*

$$\sigma \subseteq \mu \subseteq F^\theta_\Lambda \Rightarrow \sigma = \mu.$$

Then φ_σ is an ultra-character, and hence σ is a quasi-maximal string.

8 Ground characters

In the last section we were able to fruitfully study open characters using strings, culminating with Theorem (7.20), stating that every open ultra-character is given in terms of a string. However nothing of interest was said about an ultra-character when it is not open. The main purpose of this section is thus to obtain some useful information about non-open ultra-characters. The main result in this direction is Theorem (8.11), below.

► Throughout this section we fix a 0-left cancellative semigroup S admitting least common multiples. For each s in S let

$$\hat{F}_s = \{\varphi \in \hat{\mathfrak{E}}(S) : \varphi(F^\theta_s) = 1\}, \quad \text{and} \quad \hat{E}_s = \{\varphi \in \hat{\mathfrak{E}}(S) : \varphi(E^\theta_s) = 1\},$$

and for every φ in $\hat{F}_s$, consider the character $\hat{\theta}_s(\varphi)$ given by

$$\hat{\theta}_s(\varphi)(X) = \varphi\big(\theta_s^{-1}(E^\theta_s \cap X)\big), \quad \forall X \in \mathfrak{E}(S).$$

Observing that

$$\hat{\theta}_s(\varphi)(E^\theta_s) = \varphi\big(\theta_s^{-1}(E^\theta_s)\big) = \varphi(F^\theta_s) = 1, \tag{8.1}$$

we see that $\hat{\theta}_s(\varphi)$ is indeed a (nonzero) character, and that $\hat{\theta}_s(\varphi)$ belongs to $\hat{E}_s$. As a consequence we get a map

$$\hat{\theta}_s : \hat{F}_s \to \hat{E}_s,$$

which is easily seen to be bijective, with inverse given by

$$\hat{\theta}_s^{-1}(\varphi)(X) = \varphi\big(\theta_s(F^\theta_s \cap X)\big), \quad \forall \varphi \in \hat{E}_s, \quad \forall X \in \mathfrak{E}(S).$$

We may then see each $\hat{\theta}_s$ as an element of $\mathcal{I}\big(\hat{\mathfrak{E}}(S)\big)$, and it is not hard to see that the correspondence

$$\hat{\theta} : s \in S \mapsto \hat{\theta}_s \in \mathcal{I}(\hat{\mathfrak{E}}(S))$$

is a representation of S on $\hat{\mathfrak{E}}(S)$.

All of this may also be deduced from the fact that any inverse semigroup, such as $\mathfrak{H}(S)$, admits a canonical representation on the spectrum of its idempotent semilattice (see Exel [2008, Section 10]), and that $\hat{\theta}$ may be obtained as the composition

$$S \xrightarrow{\theta} \mathfrak{H}(S) \longrightarrow \mathfrak{I}\big(\hat{\mathfrak{E}}(S)\big),$$

where the arrow in the right-hand-side is the canonical representation mentioned above.

Definition 8.2. We shall refer to $\hat{\theta}$ as the *dual representation* of S.

The following technical result gives some useful information regarding the relationship between the dual representation and the representation ρ of $\mathfrak{H}(S)$ described in (6.15).

Lemma 8.3. *Given s in S, and σ in $S^\star_\sharp$, one has that*

(i) $\varphi_\sigma \in \hat{F}_s \iff \sigma \in F^\star_s$,

(ii) *if the equivalent conditions in (i) are satisfied, then* $\hat{\theta}_s(\varphi_\sigma) = \varphi_{\theta^\star_s(\sigma)}$,

(iii) $\varphi_\sigma \in \hat{E}_s \iff \sigma \in E^\star_s$,

(iv) *if the equivalent conditions in (iii) are satisfied, then* $\hat{\theta}_s^{-1}(\varphi_\sigma) = \varphi_{\theta^{\star -1}_s(\sigma)}$.

Considering the representation $\theta^\star$ of S on $S^\star$, observe that $S^\star_\sharp$ is an invariant subset of $S^\star$, and it is easy to see that it is also invariant under the representation ρ of $\mathfrak{H}(S)$ described in (6.15). Together with the dual representation of $\mathfrak{H}(S)$ on $\hat{\mathfrak{E}}(S)$ mentioned above, we thus have two natural representations of $\mathfrak{H}(S)$, which are closeely related, as the following immediate consequence of the above result asserts:

Proposition 8.4. *The mapping*

$$\Phi : S^\star_\sharp \to \hat{\mathfrak{E}}(S)$$

of (7.7) is covariant relative to the natural representations of $\mathfrak{H}(S)$ referred to above.

The fact that the correspondence between strings and characters is not a perfect one (see e.g. (7.13) and (7.17)) is partly responsible for the fact that expressing the covariance properties of the map Σ of (7.12) cannot be done in the same straightforward way as we did for Φ in (8.4). Nevertheless, there are some things we may say in this respect.

Let us first treat the question of covariance regarding $\hat{\theta}_s^{-1}(\varphi)$. Of course, for this to be a well defined character we need φ to be in $\hat{E}_s$, meaning that $\varphi(E^\theta_s) = 1$, which is also equivalent to saying that $s \in \sigma_\varphi$. In particular characters with empty strings are immediately ruled out.

Lemma 8.5. *For every s in S, and every character φ in $\hat{E}_s$, one has that*

$$\sigma_{\hat{\theta}_s^{-1}(\varphi)} = \{p \in S : sp \in \sigma_\varphi\}.$$

The set appearing in the right hand side of the equation displayed in (8.5) is precisely the same set mentioned in definition (6.6.ii) of $s^{-1} * \sigma_\varphi$, except that this notation is reserved for the situation in which the intersection of σ with sS is nonempty, which precisely means that $s \in \mathring{\sigma}_\varphi$.

Proposition 8.6. *Pick s in S and let φ be any character in $\hat{E}_s$. Then $s \in \sigma_\varphi$, and moreover*

(i) *if s is in $\mathring{\sigma}_\varphi$, then $\sigma_\varphi \in E_s^\star$, and $\sigma_{\hat{\theta}_s^{-1}(\varphi)} = \theta_s^{\star-1}(\sigma_\varphi)$,*

(ii) *if s is not in $\mathring{\sigma}_\varphi$, then $\sigma_{\hat{\theta}_s^{-1}(\varphi)} = \varnothing$.*

Regarding the behavior of strings associated to characters of the form $\hat{\theta}_s(\varphi)$, we have:

Lemma 8.7. *For every s in S, and every character φ in $\hat{F}_s$, one has that $\hat{\theta}_s(\varphi)$ belongs to $\hat{\mathfrak{E}}_1(S)$ (and hence (7.10) implies that $\sigma_{\hat{\theta}_s}(\varphi)$ is a string), and moreover*

(i) *if σ_φ is nonempty, then $\sigma_\varphi \in F_s^\star$, and $\sigma_{\hat{\theta}_s(\varphi)} = \theta_s^\star(\sigma_\varphi)$,*

(ii) *if σ_φ is empty, then $\sigma_{\hat{\theta}_s(\varphi)} = \delta_s$.*

We may interpret the above result, and more specifically the identity

$$\sigma_{\hat{\theta}_s(\varphi)} = \theta_s^\star(\sigma_\varphi),$$

as saying that the correspondence $\varphi \mapsto \sigma_\varphi$ is covariant with respect to the actions $\hat{\theta}$ and $\theta^\star$, on $\hat{\mathfrak{E}}(S)$ and $S^\star$, respectively, except that the term "σ_φ" appearing is the right-hand-side above is not a well defined string since it may be empty, even though the left-hand-side is always well defined. In the problematic case of an empty string, (8.7.ii) then gives the undefined right-hand-side the default value of δ_s.

Definition 8.8. A character φ in $\hat{\mathfrak{E}}(S)$ will be called a *ground* character if σ_φ is empty.

By (7.10), the ground characters are precisely the members of $\hat{\mathfrak{E}}(S) \setminus \hat{\mathfrak{E}}_1(S)$.

Besides the ground characters, a character φ may fail to be open because σ_φ, while being a bona fide string, is not an open string. In this case we have that $\sigma_\varphi = \delta_s$, for some s in S such that $s \notin sS$.

Proposition 8.9. *Let φ be a character such that $\sigma_\varphi = \delta_s$, where s is such that $s \notin sS$. Then $\varphi \in \hat{E}_s$, and $\hat{\theta}_s^{-1}(\varphi)$ is a ground character.*

We may now give a precise characterization of non-open characters in terms of the ground characters:

Proposition 8.10. *Denote by* $\hat{\mathfrak{E}}_{\mathrm{op}}(S)$ *the set of all open characters on* $\mathfrak{E}(S)$*. Then*

$$\hat{\mathfrak{E}}(S) \setminus \hat{\mathfrak{E}}_{\mathrm{op}}(S) = \left\{\hat{\theta}_u(\varphi) : u \in \tilde{S},\ \varphi \ \text{ is a ground character in } \ \hat{F}_u\right\}.$$

Moreover for each ψ *in the above set, there is a unique pair* (u, φ)*, with* u *in* $\tilde{S}$*, and* φ *a ground character, such that* $\psi = \hat{\theta}_u(\varphi)$*.*

We may now combine several of our earlier results to give a description of all ultra-characters on $\mathfrak{E}(S)$.

Theorem 8.11. *Let* S *be a* 0*-left cancellative semigroup admitting least common multiples. Denote by* $\hat{\mathfrak{E}}_\infty(S)$ *the set of all ultra-characters on* $\mathfrak{E}(S)$*, and by*

$$\hat{\mathfrak{E}}_\infty^{\mathrm{op}}(S) = \hat{\mathfrak{E}}_{\mathrm{op}}(S) \cap \hat{\mathfrak{E}}_\infty(S),$$

namely the subset formed by all open ultra-characters. Then

(i) $\hat{\mathfrak{E}}_\infty^{\mathrm{op}}(S) = \{\varphi_\sigma : \sigma \text{ is an open, quasi-maximal string in } S\}$*, and*

(ii) $\hat{\mathfrak{E}}_\infty(S) \setminus \hat{\mathfrak{E}}_\infty^{\mathrm{op}}(S) = \{\hat{\theta}_u(\varphi) : u \in \tilde{S},\ \varphi \text{ is a ground, ultra-character in } \hat{F}_u\}$*.*

The upshot is that in order to understand all ultra-characters on $\mathfrak{E}(S)$, it only remains to describe the ground ultra-characters.

9 Subshift semigroups

By a subshift on a finite alphabet Σ one means a subset $\mathfrak{X} \subseteq \Sigma^{\mathbf{N}}$, which is closed relative to the product topology, and invariant under the left shift map

$$x_1x_2x_3\ldots \mapsto x_2x_3x_4\ldots$$

► Throughout this chapter we will let $\mathfrak{X}$ be a fixed subshift.

The language of $\mathfrak{X}$ is the set L formed by all finite words appearing as a block in some infinite word belonging to $\mathfrak{X}$. We will not allow the empty word in L, as sometimes done in connection with subshifts, so all of our words have strictly positive length.

In the present section we will be concerned with the semigroup

$$S_{\mathfrak{X}} = L \cup \{0\}, \tag{9.1}$$

equipped with the multiplication operation given by

$$\mu \cdot \nu = \begin{cases} \mu\nu, & \text{if } \mu, \nu \neq 0, \text{ and } \mu\nu \in L, \\ 0, & \text{otherwise,} \end{cases}$$

where $\mu\nu$ stands for the concatenation of μ and ν.

Given μ and ν in $S_{\mathfrak{X}}$, with ν nonzero, notice that $\mu \mid \nu$ if and only if μ is a prefix of ν. Given that divisibility is also well defined for the unitized semigroup $\tilde{S}_{\mathfrak{X}}$, and that 1 divides any $\mu \in S_{\mathfrak{X}}$, we will also say that 1 is a prefix of μ.

Some of the special properties of $S_{\mathfrak{X}}$ of easy verification are listed below:

Proposition 9.2.

(i) *$S_{\mathfrak{X}}$ is 0-left cancellative and 0-right cancellative,*

(ii) *$S_{\mathfrak{X}}$ admits least common multiples,*

(iii) *$S_{\mathfrak{X}}$ has no idempotent elements other than 0.*

A further special property of $S_{\mathfrak{X}}$ is a very strong uniqueness of the normal form for elements in $\mathfrak{H}(S_{\mathfrak{X}})$:

Proposition 9.3. *For $i = 1, 2$, let Λ_i be a finite subset of $\tilde{S}X$ intersecting $S_{\mathfrak{X}}$ non-trivially, and let $u_i, v_i \in \Lambda_i$ be such that*

$$\theta_{u_1} f_{\Lambda_1} \theta_{v_1}^{-1} = \theta_{u_2} f_{\Lambda_2} \theta_{v_2}^{-1} \neq 0.$$

Then $u_1 = u_2$, $v_1 = v_2$, and $F_{\Lambda_1} = F_{\Lambda_2}$.

Given the importance of strings, let us give an explicit description of these in the present context.

Proposition 9.4. *Given a (finite or infinite) word*

$$\omega = \omega_1\omega_2\omega_3\ldots,$$

on the alphabet Σ, assume ω to be admissible (meaning that ω belongs to L, if finite, or to $\mathfrak{X}$, if infinite) and consider the set δ_ω formed by all prefixes of ω having positive length, namely

$$\delta_\omega = \{\omega_1,\ \omega_1\omega_2,\ \omega_1\omega_2\omega_3,\ \ldots\}.$$

Then:

(i) *δ_ω is a string,*

(ii) δ_ω *is an open string if and only if* ω *is an infinite word,*

(iii) δ_ω *is a maximal string if and only if* ω *is an infinite word,*

(iv) *for any string* σ *in* $S_{\mathfrak{X}}$, *there exists a unique admissible word* ω *such that* $\sigma = \delta_\omega$.

Strings consist in one of our best instruments to provide characters on $\mathfrak{E}(S_{\mathfrak{X}})$. Now that we have a concrete description of strings in terms of admissible words, let us give an equally concrete description of the characters induced by strings.

Proposition 9.5. *Let* ω *be a given infinite admissible word, and let* X *be any* θ*-constructible set, written in normal form, namely* $X = uF^\theta_\Lambda$, *where* Λ *is a finite subset of* $\tilde{S}_{\mathfrak{X}}$, *intersecting* $S_{\mathfrak{X}}$ *nontrivially, and* $u \in \Lambda$. *Regarding the string* δ_ω, *and the associated character* φ_{δ_ω}, *the following are equivalent:*

(i) $\varphi_{\delta_\omega}(X) = 1$,

(ii) u *is a prefix of* ω, *and upon writing* $\omega = u\eta$, *for some infinite word* η, *one has that* $t\eta$ *is admissible (i.e. lies in* $\mathfrak{X}$*), for every* t *in* Λ.

In case the set X of the above result coincides with F^θ_μ, for some μ in L, we get the following simplification:

Proposition 9.6. *Let* ω *be an infinite admissible word, and let* $\mu \in L$. *Then*

$$\varphi_{\delta_\omega}(F^\theta_\mu) = 1 \Leftrightarrow \mu\omega \in \mathfrak{X}.$$

Let us now study a situation in which infinite words provide all ultra-characters.

Proposition 9.7. *The following are equivalent:*

(i) *for every finite subset* $\Lambda \subseteq \tilde{S}_{\mathfrak{X}}$, *such that* $\Lambda \cap S_{\mathfrak{X}} \neq \varnothing$, *one has that* F^θ_Λ *is either empty or infinite,*

(ii) *every nonempty constructible set is infinite,*

(iii) $\{E^\theta_a : a \in \Sigma\}$ *is a cover for* $\mathfrak{E}(S_{\mathfrak{X}})$,

(iv) $\mathfrak{E}(S_{\mathfrak{X}})$ *admits no ground ultra-characters,*

(v) *for every ultra-character* φ *on* $\mathfrak{E}(S_{\mathfrak{X}})$, *there exists an infinite admissible word* ω, *such that* $\varphi = \varphi_{\delta_\omega}$.

Let us conclude this section with an example to show that nonempty finite constructible sets may indeed exist and hence the equivalent conditions of (9.7) do not always hold. Consider the alphabet $\Sigma = \{a, b, c\}$ and let $\mathfrak{X}$ be the subshift on Σ consisting of all infinite words ω such that, in any block of ω of length three, there are no repeated letters. Alternatively, a set of forbidden words defining $\mathfrak{X}$ is the set of all words of length three with some repetition.

It is then easy to see that the language L of $\mathfrak{X}$ is formed by all finite words on Σ with the same restriction on blocks of length three described above.

Notice that $c \in F_{\{a,b\}}$, because both ac and bc are in L. However there is no element in $F_{\{a,b\}}$ other than c, because it is evident that neither a nor b lie in $F_{\{a,b\}}$, and for any x in Σ, either acx or bcx will involve a repetition. So *voilà* the finite constructible set:

$$F_{\{a,b\}} = \{c\}. \tag{9.8}$$

10 C*-algebras associated to subshifts

In this final section we will briefly discuss applications of the theory so far described to various C*-algebras associated to subshifts that have been studied starting with Matsumoto's original work [Matsumoto 1997].

Given a subshift $\mathfrak{X}$, as in the previous section, we will consider the 0-left cancellative semigroup $S_{\mathfrak{X}}$, as well as its inverse hull $\mathfrak{H}(S_{\mathfrak{X}})$. We may then consider several general constructions of C*-algebras from inverse semigroups, and our goal is to argue that many of these, once applied to $\mathfrak{H}(S_{\mathfrak{X}})$, produce all of the C*-algberas studied in the literature in connection with subshifts.

The constructions we have in mind share a common pattern in the following sense. Given an inverse semigroup $\mathcal{S}$ with zero, consider the standard action of $\mathcal{S}$ on $\widehat{E}(\mathcal{S})$, namely the dual of the idempotent semilattice of $\mathcal{S}$. We may then build the groupoid $\mathcal{G}_{\mathcal{S}}$ formed by all germs for this action. This groupoid is sometimes referred to via the suggestive notation

$$\mathcal{S} \ltimes \widehat{E}(\mathcal{S}).$$

The C*-algebra of $\mathcal{G}_{\mathcal{S}}$ is well known to be a quotient[5] of Paterson's universal C*-algebra for $\mathcal{S}$.

Whenever Y is a closed invariant subset of $\widehat{E}(\mathcal{S})$, we may restrict the action of $\mathcal{S}$ to Y, and consider the corresponding groupoid of germs $\mathcal{S} \ltimes Y$, which may also be seen as the reduction of $\mathcal{G}_{\mathcal{S}}$ to Y.

[5] Modulo the relation that identifies the zero of $\mathcal{S}$ with the zero of the corresponding C*-algebra.

Our point is that several C*-algebras studied in the literature in connection to the subshift $\mathfrak{X}$ are actually groupoid C*-algebras of the form

$$C^*\big(\mathfrak{H}(S_{\mathfrak{X}}) \ltimes Y\big),$$

where Y is a closed subset of $\hat{\mathfrak{E}}(S_{\mathfrak{X}})$, invariant under the standard action of $\mathfrak{H}(S_{\mathfrak{X}})$.

In order to describe the first relevant alternative for Y, let S be any 0-left cancellative semigroup and consider the representation ι of $\mathfrak{E}(S)$ on $\mathcal{P}(S')$ given by the inclusion of the former in the latter. We will say that a character φ of $\mathfrak{E}(S)$ is essentially tight (relative to the above representation ι) provided one has that

$$\varphi(X) = \bigvee_{i=1}^{n} \varphi(Y_i),$$

whenever $X, Y_1, \ldots, Y_n$ are in $\mathfrak{E}(S)$, and the symmetric difference

$$X \,\Delta \left(\bigcup_{i=1}^{n} Y_i\right)$$

is finite. The set of all essentially tight characters of $\mathfrak{E}(S)$ will be denoted by $\hat{\mathfrak{E}}_{\text{ess}}(S)$, and it may be shown that $\hat{\mathfrak{E}}_{\text{ess}}(S)$ is a closed invariant subset of $\hat{\mathfrak{E}}(S)$.

The second relevant alternative for Y is based on the set $\hat{\mathfrak{E}}_{\max}(S)$ defined by

$$\hat{\mathfrak{E}}_{\max}(S) = \{\varphi_\sigma : \sigma \text{ is a maximal string}\}.$$

In general $\hat{\mathfrak{E}}_{\max}(S)$ is not invariant under the action of $\mathfrak{H}(S)$, but when $\mathcal{S} = S_{\mathfrak{X}}$ for some subshift $\mathfrak{X}$, invariance is guaranteed. One may then take Y to be the closure of $\hat{\mathfrak{E}}_{\max}(S_{\mathfrak{X}})$.

The tight spectrum $\hat{\mathfrak{E}}_{\text{tight}}(S_{\mathfrak{X}})$ is a further alternative, and in some sense it is the most natural one given that many C*-algebras associated to inverse semigroups turn out to be the groupoid C*-algebra for the reduction of Paterson's universal groupoid to the tight spectrum of the idempotent semilattice.

These subsets are related to each other as follows

$$\begin{array}{ccc} \overline{\hat{\mathfrak{E}}_{\max}(S_{\mathfrak{X}})} & \subseteq & \hat{\mathfrak{E}}_{\text{tight}}(S_{\mathfrak{X}}) \\ \rotatebox{90}{$\supseteq$} & & \\ \hat{\mathfrak{E}}_{\text{ess}}(S_{\mathfrak{X}}) & & \end{array} \tag{10.1}$$

As already observed, all of these are subsets of $\mathfrak{E}(S_{\mathfrak{X}})$ which are closed and invariant under the action of $\mathfrak{H}(S_{\mathfrak{X}})$, so each gives rise to a reduced subgroupoid, which we will correspondingly denote by $\mathcal{G}_{\max}$, $\mathcal{G}_{\text{tight}}$ and $\mathcal{G}_{\text{ess}}$.

Theorem 10.2. *Given any subshift $\mathfrak{X}$, one has that*

(i) $C^*(\mathcal{G}_{\text{ess}})$ *is isomorphic to Matsumoto's C*-algebra introduced in [Matsumoto 1997].*

(ii) $C^*(\mathcal{G}_{\max})$ *is isomorphic to Carlsen-Matsumoto's C*-algebra introduced in Carlsen and Matsumoto [2004, Definition 2.1].*

There are many situations in which the inclusions in (10.1) reduce to equality, but examples may be given to show these are, in general, proper inclusions. The fact that $\overline{\hat{\mathfrak{E}}_{\max}(S_{\mathfrak{X}})}$ and $\hat{\mathfrak{E}}_{\text{ess}}(S_{\mathfrak{X}})$ may differ is related to the fact that Matsumoto's C*-algebra may be non-isomorphic to the Carlsen-Matsumoto one, but it may be shown that under condition $(*)$[6] of Carlsen and Matsumoto [ibid.], one has that $\overline{\hat{\mathfrak{E}}_{\max}(S_{\mathfrak{X}})} = \hat{\mathfrak{E}}_{\text{ess}}(S_{\mathfrak{X}})$, whence isomorphism holds.

References

Toke Meier Carlsen and Kengo Matsumoto (2004). "Some remarks on the C^*-algebras associated with subshifts". *Math. Scand.* 95.1, pp. 145–160. MR: 2091486 (cit. on p. 1631).

Alessandra Cherubini and Mario Petrich (1987). "The inverse hull of right cancellative semigroups". *J. Algebra* 111.1, pp. 74–113. MR: 913198 (cit. on pp. 1602, 1605).

A. H. Clifford and G. B. Preston (1961). *The algebraic theory of semigroups. Vol. I.* Mathematical Surveys, No. 7. American Mathematical Society, Providence, R.I., pp. xv+224. MR: 0132791 (cit. on pp. 1602, 1605).

L. A. Coburn (1967). "The C^*-algebra generated by an isometry". *Bull. Amer. Math. Soc.* 73, pp. 722–726. MR: 0213906 (cit. on p. 1601).

– (1969). "The C^*-algebra generated by an isometry. II". *Trans. Amer. Math. Soc.* 137, pp. 211–217. MR: 0236720 (cit. on p. 1601).

M. Dokuchaev and R. Exel (2017). "Partial actions and subshifts". *J. Funct. Anal.* 272.12, pp. 5038–5106. MR: 3639522 (cit. on p. 1602).

Ruy Exel (2008). "Inverse semigroups and combinatorial C^*-algebras". *Bull. Braz. Math. Soc. (N.S.)* 39.2, pp. 191–313. MR: 2419901 (cit. on pp. 1602, 1624).

Ruy Exel and Marcelo Laca (1999). "Cuntz-Krieger algebras for infinite matrices". *J. Reine Angew. Math.* 512, pp. 119–172. MR: 1703078 (cit. on p. 1602).

Ruy Exel and Enrique Pardo (2016). "The tight groupoid of an inverse semigroup". *Semigroup Forum* 92.1, pp. 274–303. MR: 3448414 (cit. on p. 1608).

[6] The reader should be warned that the description of condition $(*)$ in [Carlsen and Matsumoto 2004] is incorrect and must be amended by requiring that the sequence $\{\mu_i\}_i$, mentioned there, have an infinite range.

Xin Li (2012). "Semigroup C*-algebras and amenability of semigroups". *J. Funct. Anal.* 262.10, pp. 4302–4340. MR: 2900468 (cit. on pp. 1601, 1602).

Kengo Matsumoto (1997). "On C^*-algebras associated with subshifts". *Internat. J. Math.* 8.3, pp. 357–374. MR: 1454478 (cit. on pp. 1629, 1631).

Marston Morse and Gustav A. Hedlund (1944). "Unending chess, symbolic dynamics and a problem in semigroups". *Duke Math. J.* 11, pp. 1–7. MR: 0009788 (cit. on p. 1607).

W. D. Munn (1964). "Brandt congruences on inverse semigroups". *Proc. London Math. Soc. (3)* 14, pp. 154–164. MR: 0172952 (cit. on p. 1607).

G. J. Murphy (1987). "Ordered groups and Toeplitz algebras". *J. Operator Theory* 18.2, pp. 303–326. MR: 915512 (cit. on p. 1601).

– (1991). "Ordered groups and crossed products of C^*-algebras". *Pacific J. Math.* 148.2, pp. 319–349. MR: 1094493 (cit. on p. 1601).

Gerard J. Murphy (1994). "Crossed products of C^*-algebras by semigroups of automorphisms". *Proc. London Math. Soc. (3)* 68.2, pp. 423–448. MR: 1253510 (cit. on p. 1601).

A. Nica (1992). "C^*-algebras generated by isometries and Wiener-Hopf operators". *J. Operator Theory* 27.1, pp. 17–52. MR: 1241114 (cit. on p. 1610).

Alan L. T. Paterson (1999). *Groupoids, inverse semigroups, and their operator algebras*. Vol. 170. Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA, pp. xvi+274. MR: 1724106 (cit. on p. 1602).

Received 2017-12-06.

Ruy Exel
ruyexel@gmail.com

Benjamin Steinberg
bsteinberg@ccny.cuny.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1631–1656)

CONSTANT NONLOCAL MEAN CURVATURES SURFACES AND RELATED PROBLEMS

Mouhamed Moustapha Fall

Abstract

The notion of Nonlocal Mean Curvature (NMC) appears recently in the mathematics literature. It is an extrinsic geometric quantity that is invariant under global reparameterization of a surface and provide a natural extension of the classical mean curvature. We describe some properties of the NMC and the quasilinear differential operators that are involved when it acts on graphs. We also survey recent results on surfaces having constant NMC and describe their intimate link with some problems arising in the study of overdetermined boundary value problems.

1 Introduction

The concept of *mean curvature* of a surface goes back to Sophie Germain's work on elasticity theory in the seventeenth century. The mathematical formulation of the mean curvature was first derived by Young and then by Laplace in the eighteenth, see Finn [1986]. The mean curvature of a surface is an extrinsic measure of curvature which locally describes the curvature of surface in some ambient space. The notion of *nonlocal mean curvature* appeared recently and for the first time in the work of L. A. Caffarelli and Souganidis [2010] on cellular automata. It is also an extrinsic geometric quantity that is invariant under global reparameterization of a surface. While Constant Mean Curvature (CMC) surfaces are stationary points for the area functional under some constraints, Constant Nonlocal Mean Curvature (CNMC) surfaces are also critical points of a fractional order area functional, called fractional or nonlocal perimeter, under some constraints. For simplicity, the fractional perimeter of a bounded set is given by a Sobolev fractional seminorm of its indicator function. Moreover, up to a normalization, as the fractional parameter tends to a maximal value, it approaches the classical perimeter functional. In some phase transition

The author would like to thank Tobias Weth, Xavier Cabré and Enrico Valdinoci for useful discussions. The author's work is supported by the Alexander von Humboldt Foundation.

MSC2010: primary 35SXX; secondary 58J0, 58J55, 47G30, 35N05, 35S0.

problems driven by Lèvy-type diffusion, the sharp-interface energy is given by the fractional perimeter functional, see Savin and Valdinoci [2012]. Moreover, the rich structure of CNMC surfaces captures the geometry and distributions of some periodic equilibrium patterns observed in physical systems involving short and long range competitions, see M. M. Fall [2017].

In this note, we survey recent results on embedded surfaces with nonzero constant nonlocal mean curvature together with their counterparts within the classical theory of constant mean curvature surfaces such as the Alexandrov's classification theorem, rotationally symmetric and periodic surfaces. Unlike the study of CMC surfaces, where we can rely on the theory ordinary differential equations in certain cases, the nonlocal case does not provide such advantage. Neither does it provide, in general, an analytic verification of parameterized surfaces to have CNMC. Up to now, almost all nontrivial surfaces with CNMC appearing in the literature are either derived by variational methods or by means of topological bifurcation theory. The study of solutions to partial differential equations with overdetermined boundary values is surprisingly intimately related with the question of finding CNMC surfaces. We also describe a number of such phenomena, notably those that were found recently as formal consequences of the highly nonlinear and nonlocal equations which are involved.

2 Constant mean curvature hypersurfaces

Let Σ be an orientable C^2 hypersurface of $\mathbb{R}^N$ and denote by $\mathcal{V}_\Sigma : \Sigma \to \mathbb{R}^N$ the unit normal vector field on Σ. For every $p \in \Sigma$, we let $\{e_1; \dots; e_{N-1}\}$ be an orthonormal basis of the tangent plane $T_p\Sigma$ of Σ at p. The (normalized) *mean curvature* at p of Σ is given by

$$H(\Sigma; p) := \frac{1}{N-1} \sum_{i=1}^{N-1} \langle D\mathcal{V}_\Sigma(p)e_i, e_i\rangle. \tag{2.1}$$

Here and in the following, $\langle , \rangle$ and "$\cdot$" denote scalar product on $\mathbb{R}^N$. As a consequence, for a C^1-extension of $\mathcal{V}_\Sigma$ by a unit vector field $\widetilde{\mathcal{V}}_\Sigma$ in a neighborhood of p in $\mathbb{R}^N$, we have

$$H(\Sigma; p) := \frac{1}{N-1} \mathrm{div}_{\mathbb{R}^N} \widetilde{\mathcal{V}}_\Sigma(p).$$

Let Ω and E be two open subsets of $\mathbb{R}^N$ with $E \subset \Omega$. Then the *perimeter functional of E relative to* Ω (total variation of 1_E in Ω) is given by

$$P(E, \Omega) = |D1_E|(\Omega) := \sup\left\{\int_E \mathrm{div}\xi(x)\, dx \ : \ \xi \in C_c^\infty(\Omega; \mathbb{R}^N),\ |\xi| \le 1\right\}.$$

In the following, we will simply write $P(E) := P(E, \mathbb{R}^N)$.

Definition 2.1. *We consider a vector field $\zeta \in C_c^\infty(\mathbb{R}^N, \mathbb{R}^N)$ and define the flow $(Y_t)_{t\in\mathbb{R}}$ induced by ζ given by*

$$\begin{cases} \partial_t Y_t(x) = \zeta(Y_t(x)) & t \in \mathbb{R} \\ Y_0(x) = x & \text{for all } x \in \mathbb{R}^N. \end{cases} \tag{2.2}$$

For $E \subset \mathbb{R}^N$, we call the family of sets $E_t := Y_t(E)$, $t \in \mathbb{R}$, the variation of E with respect to the vector field ζ.

Physically, constant mean curvatures surfaces appear also when looking at surface at equilibrium enclosing a given volume in the absence of gravity. We have the well know formula for first variation of area, see Spivak [1979].

Proposition 2.2. *Let Ω and E be two bounded domains of $\mathbb{R}^N$, with E of class C^2. Let $\lambda \in \mathbb{R}$ and $(E_t)_t$ be a variation of E with respect to $\zeta \in C_c^\infty(\Omega; \mathbb{R}^N)$. Then the map $t \mapsto J(t) := P(E_t, \Omega) - \lambda|E_t \cap \Omega|$ is differentiable at zero. Moreover*

$$J'(0) = (N-1)\int_{\partial E} \{H(\partial E; p) - \lambda\}\, v(p)\, dV(p),$$

where $v(p) := \langle \zeta(p), \mathcal{V}_{\partial E}(p)\rangle$ and $\mathcal{V}_{\partial E}$ is the unit exterior normal vector field of E.

The first classification result of compact CMC surfaces is due to Jellet in the 18th in Jellet [1853], showing that a compact CMC surface, enclosing a star-shaped domain in $\mathbb{R}^3$ must be the standard sphere. An other classification is due to Aleksandrov [1958], which we record in the following

Theorem 2.3. *An embedded closed C^2 hypersurface in $\mathbb{R}^N$, with nonzero constant mean curvature is a finite union of disjoint round spheres with same radius.*

Alexandrov's result for embedded CMC hypersurfaces provides, in particular, a positive partial answer to a conjecture of H. Hopf, stating that: a compact orientable CMC surface must be a sphere. Hopf gave a positive answer to his conjecture in Hopf [1989] for the case of CMC immersions of S^2 into $\mathbb{R}^3$. As what concerns immersed hypersurfaces with genus larger than 1, Hsiang found in Hsiang [1982] a first counterexample in $\mathbb{R}^4$ to Hopf's conjecture. Later Wente constructed in Wente [1986] an immersion of the torus $\mathbb{T}^2$ in $\mathbb{R}^3$ having constant mean curvature. While Wente's CMC torus has genus $g = 1$, compact CMC immersions with any genus $g > 1$ in $\mathbb{R}^3$ were obtained by N. Kapouleas [1991]. For a recent survey on constant mean curvature surfaces, we refer the reader to Meeks, Pérez, and Tinaglia [2016] and the references therein.

Alexandrov introduced the method of moving plane in the proof of Theorem 2.3. Formally, the argument goes as follows. Consider a connected hypersurface Σ of class C^2 and pick a unit vector $e \in \mathbb{R}^N$ together with a hyperplane P_e which is perpendicular to e and not intersecting Σ. Then slide the plane along the e-direction, toward Σ, until one of the two properties occurs for the first time:

1. *interior touching:* the reflection of Σ with respect to P_e, called Σ_e, intersects Σ at a point $p_0 \notin P_e$,

2. *non-transversal intersection:* the vector $e \in T_{p_0}\Sigma$, for some $p_0 \in \Sigma \cap \Sigma_e$.

In either case, local comparison principles for elliptic equations applied to the graphs u and u_e, locally representing Σ and Σ_e, show that there exists an $r > 0$ such that $\Sigma \cap B_r(p_0) = \Sigma_e \cap B_r(p_0)$. Now by unique continuations property of elliptic PDEs, we find that $\Sigma = \Sigma_e$. Since e is arbitrary, this implies that Σ must be a round sphere. Indeed, letting H the mean curvature of Σ, then u and u_e satisfies

$$\operatorname{div}\frac{\nabla u}{\sqrt{1+|\nabla u|^2}} = \operatorname{div}\frac{\nabla u_e}{\sqrt{1+|\nabla u_e|^2}} = H \qquad \text{on } B,$$

for some centered ball $B \subset \mathbb{R}^{N-1} \simeq T_{p_0}\Sigma$. Moreover, making the ball smaller if necessary, we have $u \geq u_e$ on B. The main point here is that $w = u - u_e$ solves an elliptic equation of the form

$$\partial_i(a_{ij}(x)\partial_j w) = 0 \qquad \text{on } B,$$

for a symmetric matrix $(a_{ij}(x))_{1\leq i,j\leq N-1}$ with positive and C^1 coefficients, only depending on the partial derivatives of u and u_e. The comparison principles that is need is resumed in the following result.

Lemma 2.4. *Let $w \in C^2(B) \cap C^1(\overline{B})$ be a nonegative function on $\overline{B}$ and satisfy*

$$\partial_i(a_{ij}(x)\partial_j w) = 0 \qquad \text{on } B,$$

for some positive definite matrix a_{ij} of class C^1. Then the following holds.

i. If $w(x_0) = 0$, for some $x_0 \in \partial B$ then $w = 0$ in B.

ii. If $w(x_0) = \nabla w(x_0)\cdot \nu = 0$, for some $x_0 \in \partial B$ and ν a unit vector normal to $T_{x_0}\partial B$, then $w = 0$ in B.

The proof of Lemma 2.4 can be found in many text books, see e.g. Gilbarg and Trudinger [1983]. Lemma 2.4(i) and (ii) are now used, to deal with cases *1.* and *2.*, respectively, yielding $w = 0$ in B.

2.1 Rotationally symmetric constant mean curvature surfaces. An important class of CMC surfaces can be found in the family of unbounded rotationnaly symmetric ones. They were first studied in Delaunay [1841], which he described explicitly as surfaces of revolution of *roulettes of the conics*[1]. These surfaces are the *catenoids, unduloids, nodoids and right circular cylinders*. They have the form

$$\Sigma = \{(x(s), y(s)\sigma) \,:\, s \in \mathbb{R},\ \sigma \in S^1\} \subset \mathbb{R} \times \mathbb{R}^2.$$

where the rotating plane curve $s \mapsto (x(s), y(s))$ is parmeterized by arclength and y a positive function on $\mathbb{R}$. The explicit form of the mean curvature of Σ at a point $q(s,\sigma) = (x(s), y(s)\sigma)$ is given by

$$H(\Sigma; q(s,\sigma)) = \frac{1}{2}\frac{-x'(s) + y(s)\{x'(s)y''(s) - x''(s)y'(s)\}}{y(s)}. \tag{2.3}$$

Of particular interest to us in this note are the *embedded* surfaces with nonzero CMC, the so called *unduloids*. They constitute a smooth 1-parameter family of surfaces $(\Sigma_b)_{b\in(0,1)}$ varying from the straight cylinder Σ_0 to a translation invariant tangent spheres Σ_1. Their explicit form was found by Kenmotsu [1979],

$$\Sigma_b = \{(x_b(s), y_b(s)\sigma) \,:\, s \in \mathbb{R},\ \sigma \in S^1\},$$

where

$$x_b(s) = \int_0^s \frac{1 + b\sin(hr)}{\sqrt{1 + b^2 + 2b\sin(hr)}}\,dr \qquad \text{and} \qquad y_b(s) = \frac{1}{h}\sqrt{1 + b^2 + 2b\sin(hs)}.$$

It is clearly that Σ_b is invariant under the translations $z \mapsto z + x_b(2k\pi/h)e_1$, with $e_1 = (1,0,0)$ and $k \in \mathbb{Z}$. Moreover $\Sigma_0 = \mathbb{R} \times \frac{1}{h}S^1$, the straight cylinder of width $\frac{1}{h}$. Furthermore, with the change of variables $\cos(\frac{\theta}{2}) = \frac{h}{2}y_1(s)$, for $s \in (-\frac{\pi}{2h}, \frac{\pi}{2h})$ and $\theta = \theta(s) \in (0,\pi)$, we see that

$$x_1(s) = \frac{\sqrt{2}}{h} - \frac{2}{h}\sin(\theta) \qquad \text{and} \qquad y_1(s) = \frac{2}{h}\cos(\theta).$$

That is

$$\Sigma_1 = \left\{\frac{2}{h}\left(-\sin(\theta), \cos(\theta)\sigma\right) + \frac{\sqrt{2} + 4k}{h}e_1 \,:\, \theta \in [0,\pi],\ \sigma \in S^1,\ k \in \mathbb{Z}\right\},$$

where $e_1 = (1,0,0) \in \mathbb{R}^3$. Therefore Σ_1 is a family of tangent spheres with radius $\frac{2}{h}$, centered at the points $\frac{\sqrt{2}+4k}{h}e_1$, $k \in \mathbb{Z}$.

[1] http://www.mathcurve.com/courbes2d/delaunay/delaunay.shtml

Next, we observe that for $b \in (0,1)$, the map x_b is increasing on $\mathbb{R}$. Hence with the change of variable $t = x_b(s)$, and setting $\varphi(t) = y_b(x_b^{-1}(t))$, we deduce that

$$\Sigma_b = \{(t,\zeta) \in \mathbb{R} \times \mathbb{R}^2 \,:\, t \in \mathbb{R},\, |\zeta| = \varphi_b(t)\}.$$

Recently, Sicbaldi and Schlenk used in Schlenk and Sicbaldi [2012] the Crandall–Rabinowitz bifurcation theorem (see Crandall and Rabinowitz [1971]) to derive these surfaces for b close to 0. We record it here.

Theorem 2.5. *There exist $b_0, h_* > 0$ and a smooth curve $(-b_0, b_0) \ni b \mapsto \lambda(b)$ such that $\lambda(0) = 1$ and*

$$\varphi_b(t) = \frac{1}{h_*} + \frac{b}{\lambda(b)} \{\cos(\lambda(b)t) + v_a(\lambda(b)t)\},$$

where $v_b \to 0$ in $C^{2,\alpha}(\mathbb{R}/2\pi\mathbb{Z})$ as $b \to 0$ and $\int_{-\pi}^{\pi} v_b(t)\cos(t)\,dt = 0$ for every $b \in (-b_0, b_0)$.

Remark 2.6. *We note that the family of surfaces $(\Sigma_b)_{b>1}$ are the immersed constant mean curvature surfaces known as the nodoids.*

Expression (2.3) was first derived by M. Sturm, in an appended note in Delaunay [1841], and characterizes the Delauney surfaces variationally, as the extremals of surfaces of rotation having fixed volume while maximizing lateral area. A gereralization of Delauney surfaces in higher dimension is due to Hsiang and Yu [1981].

3 Constant Nonlocal Mean Curvature hypersurfaces

Similarly to the mean curvature, the fractional or Nonlocal Mean Curvature (NMC for short) is an extrinsic geometric quantity that is invariant under global immersion representing a surface. Let $\alpha \in (0,1)$. If Σ is a smooth oriented hypersurface in $\mathbb{R}^N$ with unit normal vector field $\mathcal{V}_\Sigma$, its nonlocal mean curvature of order α at a point $x \in \Sigma$ is defined as

$$H_\alpha(\Sigma; x) = \frac{2}{\alpha} \int_\Sigma \frac{(y-x)\cdot \mathcal{V}_\Sigma(y)}{|y-x|^{N+\alpha}}\, dV(y). \tag{3.1}$$

If Σ is of class $C^{1,\beta}$ for some $\beta > \alpha$ and we assume $\int_\Sigma (1+|y|)^{1-N-\alpha}\, dV(y) < \infty$, then the integral in (3.1) is absolutely convergent in the Lebesgue sense. The orientation is chosen here so that H_α is positive for a sphere.

We note that if Σ is of class C^2, then the normalized nonlocal mean curvature $\frac{1-\alpha}{\omega'_N} H_\alpha(\Sigma;\cdot)$ converges, as $\alpha \to 1$, locally uniformly to the classical mean curvature

$H(\Sigma;\cdot)$ defined in 2.1, see Dávila, del Pino, and Wei [n.d.] and Abatangelo and Valdinoci [2014]. Here ω'_N is the measure of the $(N-2)$-dimensional unit sphere.
There is an alternative expression for $H_\alpha(\Sigma;\cdot)$ in terms of a solid integral. Suppose that $\Sigma = \partial E$ for some open set $E \subset \mathbb{R}^N$ and $\mathcal{V}_{\partial E}$ is the normal exterior to E. Then, for all $x \in \Sigma$, we have

$$H_\alpha(\partial E; x) = PV \int_{\mathbb{R}^N} \frac{\tau_E(y)}{|y-x|^{N+\alpha}}\, dy = \lim_{\varepsilon\to 0} \int_{\mathbb{R}^N \setminus B_\varepsilon(x)} \frac{\tau_E(y)}{|y-x|^{N+\alpha}}\, dy, \tag{3.2}$$

where $\tau_E := 1_{\mathbb{R}^N\setminus E} - 1_E$, with 1_D denotes the characteristic function of a set $D \subset \mathbb{R}^N$. This can be derived using the divergence theorem and the fact that

$$\nabla_y \cdot (y-x)|y-x|^{-N-\alpha} = -\alpha |y-x|^{-N-\alpha}. \tag{3.3}$$

Remark 3.1. *The formula of NMC in (3.2) is comparable with the one of the mean curvature, when written in infinitesimal solid integral,*

$$H(\partial E; x) := C \lim_{\varepsilon\to 0} \frac{-1}{\varepsilon |B_\varepsilon(x)|} \int_{B_\varepsilon(x)} \tau_E(y)\, dy,$$

for some constant $C > 0$ depending only on N. Here the orientation is chosen so that the mean curvature is positive for a sphere. In particular the mean curvature of ∂E is an infinitesimal average, centered at ∂E, of the sum of the "-1" coming from inside E and the "$+1$" from outside E. After all by Young's law, the mean curvature measures pressure difference across the interface of two non mixing fluids at rest. On the other hand the nonlocal mean curvature is a weighted average of the sum of all the "-1" coming from inside E and all the "$+1$" from outside E.

Both forms of the NMC (3.1) and (3.2) turn out to be useful depending on the users interests. For instance, global comparison principles is easily proved using (3.2), while expression (3.1) (without principle value integration) seems more convenient to work with when dealing with regularity of the NMC operator acting on graphs. By noticing that if $E_1 \subset E_2$ then $\tau_{E_1} - \tau_{E_2} = 2 1_{E_1 \setminus E_2}$ on $\mathbb{R}^N$, we can state the following result.

Lemma 3.2. *Let E_1, E_2 be two open sets of class $C^{1,\beta}$, $\beta > \alpha$, in a neighborhood of $p \in \partial E_1 \cap \partial E_2$. If $E_1 \subset E_2$, then $H_\alpha(\partial E_2; p) \le H_\alpha(\partial E_1; p)$, with equality if and only if $E_1 = E_2$, up to set of zero Lebesgue measure.*

Local comparison principle does not hold true in general as in the classical case.

Alike the mean curvature, the nonlocal mean curvature appears in the first variation of nonlocal perimeter functional as well. The fractional perimeter of a bounded open set

$E \subset \mathbb{R}^N$ is given by

$$P_\alpha(E) = \frac{1}{2}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} \frac{|1_E(x) - 1_E(y)|^2}{|y-x|^{N+\alpha}}\,dxdy = \int_E \int_{\mathbb{R}^N\setminus E} \frac{1}{|y-x|^{N+\alpha}}\,dxdy.$$

We note that, if E is a set with finite perimeter, the normalized fractional perimeter $\frac{1-\alpha}{\omega_{N-1}} P_\alpha(E)$ converges, as $\alpha \to 1$, to $P(E)$, see Ambrosio, De Philippis, and Martinazzi [2011] and Dávila, del Pino, and Wei [n.d.]. Here ω_{N-1} is the measure of the unit ball in $\mathbb{R}^{N-1}$. Isoperimetric type inequalities with respect to this functional was investigated in Garsia and Rodemich [1974] and Frank and Seiringer [2008]. According to a result of Frank and Seiringer [2008], it is known that balls uniquely minimize P_α among all sets with a equal volume. A quantitative stability of the fractional isoperimetric inequality is proven by Fusco, Millot, and Morini [2011].

Provided E is bounded with Lipschitz boundary, using integration by parts and (3.3), we can rewrite the fractional perimeter as

$$P_\alpha(E) = \frac{1}{\alpha(1+\alpha)}\int_{\partial E}\int_{\partial E} \frac{\mathcal{V}_{\partial E}(y)\cdot(y-x)\mathcal{V}_{\partial E}(x)\cdot(x-y)}{|y-x|^{N+\alpha}}\,dV(y)dV(x). \tag{3.4}$$

Note that (3.4) provides a natural way to define a nonlocal or fractional measure of an orientable compact hypersurface.

In the theory of minimal surfaces, the Plateau's problem is to show the existence of a surface that locally minimize perimeter with a given boundary. For a nonlocal setting of Plateau's problem, in L. Caffarelli, Roquejoffre, and Savin [2010] the authors introduced the fractional perimeter in a reference open set Ω given by

$$P_{\alpha,\Omega}(E) = L(E\cap\Omega, E^c\cap\Omega) + L(E\cap\Omega, E\cap\Omega^c) + L(E^c\cap\Omega, E^c\cap\Omega^c),$$

where $A^c := \mathbb{R}^N \setminus A$ and the interaction functional is given by

$$L(A,B) := \int_B\int_A \frac{dxdy}{|x-y|^{N+\alpha}}.$$

We note that the interaction between $E^c\cap\Omega^c$ and $E\cap\Omega^c$ is left free. This, allows a well-posed setting of a (nonparametric) nonlocal Plateau's problem: existence of sets minimizing $P_{\alpha,\Omega}$ among all sets that coincide in Ω^c. The seminal paper L. Caffarelli, Roquejoffre, and Savin [ibid.] established the first existence and regularity results for nonlocal perimeter minimizing sets in a reference set Ω, and moreover the boundary of the minimizers have, in a viscosity sense, zero NMC in Ω. In the literature the boundary of such sets are called *s-minimal or nonlocal minimal hypersurfaces*.

The following formula for the first variation of fractional perimeter was found in Figalli, Fusco, Maggi, Millot, and Morini [2015] and L. Caffarelli, Roquejoffre, and Savin [2010].

Theorem 3.3. *Let Ω and E be two domains of $\mathbb{R}^N$, with E of class $C^{1,\beta}$, $\beta > \alpha$. Let $\lambda \in \mathbb{R}$ and $(E_t)_{t\in\mathbb{R}}$ be a variation of E with respect to $\zeta \in C_c^\infty(\Omega; \mathbb{R}^N)$. Then the map $t \mapsto J_\alpha(t) := P_{\alpha,\Omega}(E_t) - \lambda|E_t \cap \Omega|$ is differentiable at zero. Moreover*

$$J'_\alpha(0) = (N-1)\int_{\partial E} \{H_\alpha(\partial E; p) - \lambda\}\, v(p)\, dV(p),$$

where $v(p) := \langle \zeta(p), \mathcal{V}_{\partial E}(p)\rangle$.

The paper Figalli, Fusco, Maggi, Millot, and Morini [2015] contains also the second variation of fractional perimeter.
From now on, we say that Σ is a CNMC hypersurface if it is of class $C^{1,\beta}$, for some $\beta > \alpha$, and such that $H_\alpha(\Sigma, \cdot)$ is a *nonzero* constant on Σ.
Even though we will be mainly interested in this note to CNMC hyperusfrcaes, we find it important to make a brief digression in the theory of the nonlocal minimal hypersurfaces and those with vanishing NMC but not necessarily minimizing fractional perimeter, since it is also a hot topic nowadays with challenging open questions, see Bucur and Valdinoci [2016].
We first recall that besides the hyperplane, which trivially has zero nonlocal mean curvature, there are some nontrivial ones: the s-Lawson cones found by Dávila, del Pino, and Wei [n.d.]; the helicoid found by Cinti, Davila, and Del Pino [2016]. Moreover, local inversion arguments have been used in Dávila, del Pino, and Wei [n.d.], to derive, for α close to 1, two interesting examples of rotationally symmetric surfaces with zero nonlocal mean curvature. The first one posses the shape of the catenoid whereas the other is disconnected with two ends which are asymptotic to a cone of revolution.
Regularity theory, stability and Bernstein-type results in the nonlocal setting are also studied in the recent years. Recall that in the theory of constant mean curvature surfaces, the boundary of perimeter minimizing regions are smooth except a closed singular set of Hausdorff dimension at most $N-8$. Such regularity result, in its full generality, is still not known to be true in the nonlocal setting. The progress made in this direction so far parallel the classical regularity theory up to a *dimension shift*. This latter fact was discovered by Dávila, del Pino, and Wei [ibid.] proving, for α close to 0 and $N = 7$, that there are nonlocal perimeter minimizing (s-Lawson) cones in any reference set. On the other hand when α is close to 1, L. Caffarelli and Valdinoci [2011, 2013] proved the $C^{1,\gamma}$ regularity of nonlocal minimal hypersurfaces, except a set of $(N-8)$-Hausdorff dimension. This, together with the subsequent results of Barrios, Figalli, and Valdinoci [2014], leads, for α close to 1, to their C^∞ regularity up to dimension $N \le 7$. We note that for $N = 2$ and any $\alpha \in (0,1)$, Savin and Valdinoci [2013] established that nonlocal minimal curves are C^∞.
Decisive regularity estimates have been proved in L. Caffarelli, Roquejoffre, and Savin [2010], Figalli and Valdinoci [2017], Cinti, Serra, and Valdinoci [2016], Cabre, Cinti, and

Serra [2017], and Cabre and Cozzi [2017], see also the monograph Bucur and Valdinoci [2016] and the references there in.

We close this section by noting that nonlocal capillary problems have been investigated in Maggi and Valdinoci [2017] and Mihaila [2017] and nonlocal mean curvature flow in Sáez and Valdinoci [2015], Cinti, Sinestrari, and Valdinoci [2016], and Imbert [2009].

3.1 Expressions of the NMC of some globally parameterized hypersurfaces. In this section, we derive some formulae of some hypersurfaces which admit global parameterizations. We will use both expression of the NMC in (2.1) and (3.1). Depending on the problems under study, each expression has its own advantages. For instance to prove regularity of the NMC operator, it is in general more convenient to consider those formulae from (2.1), whereas those from (3.1) turns out to be useful in the study of qualitative properties of the NMC acting on graphs as a quasilinear nonlocal elliptic operator.

3.1.1 Without principle value integration. We have the following result.

Proposition 3.4. *Let $N = n + m \geq 1$, with $n, m \in \mathbb{N}$. Let $u : \mathbb{R}^m \to (0, +\infty)$ be a function of class $C^{1,\beta}$, $\beta > \alpha$, and satisfy*

$$\int_{\mathbb{R}^m} \frac{(1 + |\nabla u(\tau)|)u^{n-1}(\tau)}{(1 + |\tau| + u(\tau))^{N-1+\alpha}} d\tau < \infty.$$

Consider the set

$$E_u := \{(s, \zeta) \in \mathbb{R}^m \times \mathbb{R}^n \ : \ |\zeta| < u(s)\}.$$

(i) For $n \geq 2$, the NMC of Σ_u at the point $q = (s, u(s)e_1)$ is given by

$$\begin{aligned}(3.5)\quad & -\frac{\alpha}{2} H_\alpha(\Sigma_u; q) = \\ &= \int_{S^{n-1}} \int_{\mathbb{R}^m} \frac{\{u(s) - u(s-\tau) - \tau \cdot \nabla u(s-\tau)\} u^{n-1}(s-\tau)}{\{|\tau|^2 + (u(s) - u(s-\tau))^2 + u(s)u(s-\tau)|\sigma - e_1|^2\}^{(N+\alpha)/2}} d\tau d\sigma \\ & - \frac{u(s)}{2} \int_{S^{n-1}} \int_{\mathbb{R}^m} \frac{|\sigma - e_1|^2 u^{n-1}(s-\tau)}{\{|\tau|^2 + (u(s) - u(s-\tau))^2 + u(s)u(s-\tau)|\sigma - e_1|^2\}^{(N+\alpha)/2}} d\tau d\sigma,\end{aligned}$$

where $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^n$. (ii) For $n = 1$, the NMC of Σ_u at the point $q = (s, u(s))$ is given by

$$\begin{aligned}(3.6)\quad -\frac{\alpha}{2} H(\Sigma_u; q) = & \int_{\mathbb{R}^{N-1}} \frac{u(s) - u(s-\tau) - \tau \cdot \nabla u(s-\tau)}{\{|\tau|^2 + (u(s) - u(s-\tau))^2\}^{(N+\alpha)/2}} d\tau \\ & - \int_{\mathbb{R}^{N-1}} \frac{u(s) + u(s-\tau) + \tau \cdot \nabla u(s-\tau)}{\{|\tau|^2 + (u(s) + u(s-\tau))^2\}^{(N+\alpha)/2}} d\tau.\end{aligned}$$

Moreover, all integrals above converge absolutely in the Lebesgue sense.

Expression (3.5) and (3.6) are easily derived from (3.1), by change of variables, taking into account that the unit outer normal of $\partial\Sigma_u$ at the point $q = (s, u(s)\sigma)$ is given by

$$\mathcal{V}_{\partial E_u}(q) = \frac{1}{\sqrt{1+|\nabla u|^2(s)}}(-\nabla u(s), \sigma)$$

and the volume element is $u^{n-1}(s)\sqrt{1+|\nabla u|^2(s)}ds\,d\sigma$. The proof uses similar arguments as in Cabré, M. M. Fall, and Weth [2018]. We note that the first term in the right hand of (3.6) provides the expression of NMC of an euclidean graph $x_N = u(x')$ without principle value integral.

3.1.2 With principle value integration. We consider a function $u : \mathbb{R}^{N-1} \to \mathbb{R}$ of class $C^{1,\beta}$, for some $\beta > \alpha$. We let

$$E_u = \{(y,t) \in \mathbb{R}^{N-1} \times \mathbb{R} \,:\, t < u(y)\}$$

and consider the parametrization

$$\mathbb{R}^N \to \mathbb{R}^N \qquad (y,t) \mapsto \mathfrak{F}(y,t) = (y, u(y)+t)\,.$$

It is well known that this parametrization is *volume preserving*. We now compute the NMC of the graph u (denoted by $H(u)$) at the point $\mathfrak{F}(x,0) = (x, u(x)) \in \partial E_u$. From (3.1), making change of variables and using Fubini's theorem, we have

$$\begin{aligned} H(u)(x) := H_\alpha(\partial E_u; (x,u(x))) &= \int_{\mathbb{R}^N} \frac{\tau_{E_0}(y,t)}{|(x,u(x)) - \mathfrak{F}(y,t)|^{N+\alpha}} dy\,dt \\ &= \int_{\mathbb{R}^{N-1}} I(x,y)dy, \end{aligned}$$

where

$$\begin{aligned} I(x,y) &= [\int_{-\infty}^0 - \int_0^{+\infty}] \{|x-y|^2 + (t+u(y)-u(x))^2\}^{\frac{-(N+\alpha)}{2}} dt \\ &= |x-y|^{-(N+\alpha)}[\int_{-\infty}^0 - \int_0^{+\infty}] \left\{1 + \left(\frac{t+u(y)-u(x)}{|x-y|}\right)^2\right\}^{\frac{-(N+\alpha)}{2}} dt. \end{aligned}$$

We then make the change of variables $s = \frac{t+u(y)-u(x)}{|x-y|}$ and define

$$p_u(x,y) = \frac{u(y)-u(x)}{|x-y|}, \tag{3.7}$$

to get

$$I(x,y) = -|x-y|^{-(N-1+\alpha)}\Big[\int_{-\infty}^{p_u(x,y)} - \int_{p_u(x,y)}^{+\infty}\Big]\{1+s^2\}^{\frac{-(N+\alpha)}{2}}\,ds.$$

Letting

$$F(p) := \int_p^{+\infty} \frac{d\tau}{(1+\tau^2)^{\frac{(N+\alpha)}{2}}}, \tag{3.8}$$

we find that

$$H(u)(x) = \int_{\mathbb{R}^{N-1}} \frac{F(p_u(x,y)) - F(-p_u(x,y))}{|x-y|^{N-1+\alpha}}\,dy.$$

We then have the following formulae for the NMC of the hypersurface ∂E_u.

Proposition 3.5. *Let $u : \mathbb{R}^{N-1} \to \mathbb{R}$ of class $C^{1,\beta}$, $\beta > \alpha$. Consider the subgraph of u,*

$$E_u = \left\{(y,t) \in \mathbb{R}^{N-1} \times \mathbb{R} \;:\; t < u(y)\right\}.$$

At a point $q = (x, u(x)) \in \partial E_u$, we have

$$H(u)(x) := H_\alpha(\partial E_u; q) = PV\int_{\mathbb{R}^{N-1}} \frac{F(p_u(x,y)) - F(-p_u(x,y))}{|x-y|^{N-1+\alpha}}\,dy \tag{3.9}$$

$$= PV\int_{\mathbb{R}^{N-1}} \frac{u(x)-u(y)}{|x-y|^{N+\alpha}}\, q_u(x,y)\,dy, \tag{3.10}$$

where p_u and F are given by (3.7) and (3.8), respectively, and

$$q_u(x,y) = \int_{-1}^{1}\left(1 + \tau^2\frac{(u(x)-u(y))^2}{|x-y|^2}\right)^{-\frac{N+\alpha}{2}}\,d\tau.$$

Here, (3.10), is a consequence of the fundamental theorem of calculus, which yields

$$F(p) - F(-p) = p\int_{-1}^{1} F'(\tau p)\,d\tau.$$

As in the classical case, we may expect that the difference of the NMC operator of two graphs gives rise to a nonlocal symmetric operator allowing for local comparison principles.

Lemma 3.6. *Let Ω be an open set of $\mathbb{R}^{N-1}$ and $u, v \in C^{1,\beta}_{loc}(\Omega) \cap C(\mathbb{R}^{N-1})$. Then letting $w = u - v$, for every $x \in \Omega$, we have*

$$H(u)(x) - H(v)(x) = PV\int_{\mathbb{R}^{N-1}} \frac{w(x)-w(y)}{|x-y|^{N+\alpha}}\,\widetilde{q}_{u,v}(x,y)\,dy, \tag{3.11}$$

where $\widetilde{q}_{u,v}(x,y) := -2\int_0^1 F'\left(p_v(x,y)+\rho p_{u-v}(x,y)\right) d\rho$. *In particular, if*

$$H(u)(x) - H(v)(x) \geq 0 \qquad \textit{for every } x \in \Omega$$

and $u \geq v$ *on* $\mathbb{R}^{N-1} \setminus \Omega$, *then* $u - v$ *cannot attain its global minimum in* Ω *unless it is constant on* $\mathbb{R}^{N-1}$.

Proof. Since F' is even, by the fundamental theorem of calculus, we get

$$[F(a) - F(b)] - [F(-a) - F(-b)] = 2(a-b)\int_0^1 F'\left(b + \rho(a-b)\right) d\rho.$$

In view of (3.9), this gives (3.11) since $p_w(x,y) = -\frac{w(x)-w(y)}{|x-y|}$.
For the comparison principle, we suppose that, for some $x_0 \in \Omega$, we have $w(x_0) = \min_{x\in\mathbb{R}^{N-1}} w(x)$. We then have

$$\begin{aligned} 0 &\leq H(u)(x_0) - H(v)(x_0) \\ &= -2PV\int_{\mathbb{R}^{N-1}} \frac{w(x_0)-w(y)}{|x_0-y|^{N-1+\alpha}} \int_0^1 F'\left(p_v(x_0,y) + \rho p_w(x_0,y)\right) d\rho\, dy \leq 0. \end{aligned}$$

Therefore, since $F' < 0$ on $\mathbb{R}$, we deduce that $w \equiv w(x_0)$ on $\mathbb{R}^{N-1}$. □

In the case of generalized slabs, we have the following result, the proof is similar to one in Cabré, M. M. Fall, and Weth [2018].

Proposition 3.7. *Let* $u : \mathbb{R}^m \to (0, +\infty)$ *be a a function of class* $C^{1,\beta}$ *and* $N = m + n$. *Consider the open set*

$$E_u := \{(s,\zeta) \in \mathbb{R}^m \times \mathbb{R}^n \ : \ |\zeta| < u(s)\}.$$

Then at the point $q = (s, u(s)z)$, *we have*

$$H_\alpha(\partial E_u; q) = PV\int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \frac{-\tau_{E_1}(\bar{s}, z)}{\left((s-\bar{s})^2 + (u(s)e_1 - u(\bar{s})z)^2\right)^{\frac{N+\alpha}{2}}} u^n(\bar{s})\, dz\, d\bar{s}.$$

We note that in the expressions of the NMC with PV in Proposition 3.5 and 3.7, no growth control of u at infinity is required.

3.2 Bounded constant nonlocal mean curvature hypersurfaces. In addition to the cylinders, spheres, which are CNMC with nozero NMC, we shall see that there many more. However, this class is reduced by a nonlocal counterpart of the Alexandrov result on the characterization of spheres as the only closed embedded CMC-hypersurfaces.

Theorem 3.8 (Cabré, Fall, Moustapha, Solà-Morales, and Weth [n.d.] and Ciraolo, Figalli, Maggi, and Novaga [n.d.])**.** *Suppose that E is a nonempty bounded open set (not necessarily connected) with $C^{2,\beta}$-boundary for some $\beta > \alpha$ and with the property that $H_\alpha(\partial E;\cdot)$ is constant on ∂E. Then E is a ball.*

This result was obtained at the same time and independently by Cabré, Solà-Morales, Weth and the author in Cabré, Fall, Moustapha, Solà-Morales, and Weth [n.d.] and by Ciraolo, Figalli, Maggi, and Novaga [n.d.]. We mention that the paper Ciraolo, Figalli, Maggi, and Novaga [ibid.] contains also stability results with respect to this rigidity theorem.
This new characterization of the sphere relies on the Alexandrov's moving planes method discussed above. Indeed, pick a unit vector $e \in \mathbb{R}^N$ together with a hyperplane P_e which is perpendicular to e and not intersecting E. Call E_e the reflection of E with respect to the plane P_e. Then slide the plane along the e-direction, toward E, until one of the two properties occurs for the first time:

1. *interior touching:* there exists $x_0 \not\in P_e$, with $x_0 \in \partial E \cap \partial E_e$,
2. *non-transversal intersection:* $e \in T_{x_0}\partial E$, for some $x_0 \in \partial E$.

In both cases, we will find that $E_e = E$, and since e is arbitrary, this implies that E must be a unit ball. As mentioned earlier, in contrast to the classical case, there is no local comparison principle related to the fractional mean curvature. Moreover, while comparison principle holds for graphs (see Lemma 3.6), we dare not hope for a global parameterization of E by a graph! However, we may rely on a global comparison principles inherent to problem. The montonicity of the weight $t \mapsto K_N(t) := |t|^{-N-\alpha}$ will be crucial. Indeed, in case *(1)*, we have

$$\begin{aligned}
0 = H_\alpha(\partial E; x_0) - H_\alpha(\partial E_e; x_0) &= \frac{1}{2} PV \int_{\mathbb{R}^N} (\tau_E(y) - \tau_{E_e}(y)) K_N(x_0 - y)\, dy \\
&= \frac{1}{2} PV \int_{E\setminus E_e} (\tau_E(y) - \tau_{E_e}(y)) K_N(x_0 - y)\, dy + \\
&\qquad + \frac{1}{2} PV \int_{E_e\setminus E} (\tau_E(y) - \tau_{E_e}(y)) K_N(x_0 - y)\, dy \\
&= \frac{1}{2} PV \int_{E\setminus E_e} K_N(x_0 - y)\, dy - \frac{1}{2} PV \int_{E_e\setminus E} K_N(x_0 - y)\, dy \\
&= \frac{1}{2} PV \int_{E\setminus E_e} \left(K_N(x_0 - z) - K_N(x_0 - \Re(z))\right) dz.
\end{aligned}$$

We made the change of variable, $y = \Re_e(z)$, with $\Re_e$ being the reflection with respect to the plane P_e. Now since $E \setminus E_e$ is contained on the side of the plane containing x_0, we

find that $|x_0 - z| < |x_0 - \Re(z)|$ for every $z \in E \setminus E_e$, together with the monotonicity of K_N imply that $E = E_e$.
Case *2.* is far more tricky and requires a formula for directional derivative of $H_\alpha(\partial E; \cdot)$, see Proposition 3.9 below. Since $e \in T_{x_0}\partial E = T_{x_0}\partial E_e$, we have

$$
\begin{aligned}
0 &= \frac{1}{2}\big(\partial_e H_\alpha(\partial E; x_0) - \partial_e H_\alpha(\partial E_e; x_0)\big) \\
&= -\frac{N+\alpha}{2} PV \int_{\mathbb{R}^N} (x_0 \cdot e - y \cdot e)\big(\tau_E(y) - \tau_{E_e}(y)\big) K_{N+2}(x_0 - y)\, dy \\
&= -\frac{N+\alpha}{2} PV \int_{\mathbb{R}^N} (x_0 \cdot e - y \cdot e)(1_{E\setminus E_e}(y) - 1_{E_e \setminus E}(y)) K_{N+2}(x_0 - y)\, dy.
\end{aligned}
$$

Since the integrand does not change sign on $\mathbb{R}^N$, there must be $E = E_e$. We then conclude that in either case $E = E_e$, and since e is arbitrary, we deduce that E is a ball.
Of course to make all these argument rigorous, one should take into account that the integrals are defined in the principle value sense, see Cabré, Fall, Moustapha, Solà-Morales, and Weth [n.d.] and Ciraolo, Figalli, Maggi, and Novaga [n.d.] for more details.

The following formula for the derivatives of the nonlocal mean curvature was used above, see Cabré, Fall, Moustapha, Solà-Morales, and Weth [n.d.] and Ciraolo, Figalli, Maggi, and Novaga [n.d.].

Proposition 3.9. *If $E \subset \mathbb{R}^N$ is bounded and ∂E is of class $C^{2,\beta}$ for some $\beta > \alpha$, then H_E is of class C^1 on ∂E, and we have*

$$
\partial_v H_\alpha(\partial E; x) = -(N+\alpha)\, PV \int_{\mathbb{R}^N} \tau_E(y) |x - y|^{-(N+2+\alpha)} (x - y) \cdot v\, dy
$$

for $x \in \partial E$ and $v \in T_x \partial E$.

In order to reduce to a single sphere, connectedness is obviously a necessary assumption in Theorem 2.3 whereas Theorem 3.8 allow for disconnected sets E with finitely many connected components. Similar pheonomenon was also observed in the study of fractional overdetermined problems, see M. M. Fall and Jarohs [2015].

3.3 Unbounded constant nonlocal mean curvature hypersurfaces. When compactness is dropped, the cylinder provides a trivial example of surfaces with constant and nonzero local/nonlcal mean curvature. Obviously, CMC curves bounding a periodic domain in 2-dimension are parallel straight lines and hence has zero boundary mean curvature. This latter fact does not hold in the nonlocal case, since two parallel straight lines in $\mathbb{R}^2$ has positive CNMC. In fact, any slab $\{(s, \zeta) \in \mathbb{R}^n \times \mathbb{R}^m \ : \ |\zeta| = 1\}$ has a positive CNMC, as it can be easily seen from Proposition 3.4. It is therefore natural to ask if there

are CNMC hypersurfaces besides the cylinder?
The lack of ODE theory in the nonlocal framework make this question not trivial to answer. Up to now, all existence results use either variational methods or perturbation methods.

The first answer to the above question, for $N = 2$, has been given by Cabré, J. Solà-Morales, Weth and the author in Cabré, Fall, Moustapha, Solà-Morales, and Weth [n.d.], where a continuous branch of periodic CNMC sets bifurcating from the straight band was found. This was improved and generalized in Cabré, M. M. Fall, and Weth [2018] to higher dimensions. We note that in Dávila, del Pino, Dipierro, and Valdinoci [2016], it is established variationally the existence of periodic and cylindrically symmetric hypersurfaces in $\mathbb{R}^N$ which minimize the periodic fractional perimeter under a volume constraint. More precisely, Dávila, del Pino, Dipierro, and Valdinoci [ibid.] establishes the existence of a 1-periodic minimizer for every given volume within the slab $\{(s,\zeta) \in \mathbb{R} \times \mathbb{R}^{N-1} : -1/2 < s < 1/2\}$, which are CNMC hypersurfaces in the viscosity sense of L. Caffarelli, Roquejoffre, and Savin [2010], since it is not known if they are of class $C^{1,\beta}$ for some $\beta > \alpha$.

3.3.1 CNMC hypersurface of revolution. We consider hypersurfaces with constant nonlocal mean curvature of the form

$$\Sigma_u = \{(s,\zeta) \in \mathbb{R} \times \mathbb{R}^{N-1} : |\zeta| = u(s)\},$$

where $u : \mathbb{R} \to (0, \infty)$ is a positive and even function. The following result shows the existence of a smooth branch of sets which are periodic in the variable s and have all the same constant nonlocal mean curvature; they bifurcate from a straight cylinder $\Sigma_R := \{|\zeta| = R\}$. The radius R of the straight cylinder is chosen so that the periods of the new cylinders converge to 2π as they approach the straight cylinder. We state the counterpart of Theorem 2.5.

Theorem 3.10. *Let $N \geq 2$. For every $\alpha \in (0,1)$, $\beta \in (\alpha, 1)$ there exist $R, a_0 > 0$ and smooth curves $a \mapsto u_a$ and $a \mapsto \lambda(a)$, with $\lambda(0) = 1$, such that, for every $a \in (-a_0, a_0)$,*

$$\Sigma_{u_a} = \{(s,\zeta) \in \mathbb{R} \times \mathbb{R}^{N-1} : |\zeta| = u_a(s)\}$$

is a CNMC hypersurface of class $C^{1,\beta}$, with

$$H_\alpha(\Sigma_{u_a}; q) = H_\alpha(\Sigma_R; \cdot), \qquad \text{for all } q \in \partial\Sigma_{u_a}.$$

Moreover for every $a \in (-a_0, a_0)$, we have

$$u_a(s) = R + \frac{a}{\lambda(a)} \{\cos(\lambda(a)s) + v_a(\lambda(a)s)\}, \qquad u_{-a}(s) = u_a\Big(s + \frac{\pi}{\lambda(a)}\Big),$$

where $v_a \to 0$ in $C^{1,\beta}(\mathbb{R}/2\pi\mathbb{Z})$ as $a \to 0$ and $\int_{-\pi}^{\pi} v_a(t)\cos(t)\,dt = 0$ for every $a \in (-a_0, a_0)$.

The proof of Theorem 3.10 rests on the application the Crandall-Rabinowitz bifurcation theorem Crandall and Rabinowitz [1971], applied to the quasilinear type fractional elliptic equation, see Proposition 3.4,

$$H_\alpha(\Sigma_u;\cdot) - H_\alpha(\Sigma_R;\cdot) = 0 \qquad \text{in } \partial\Sigma_u.$$

To do so and to obtain a smooth branch of bifurcation parameter, we need the smoothness of $u \mapsto (s \mapsto H_\alpha(\Sigma_u;(s,u(s)e_1)))$ as map from an open subset of $C^{1,\beta}(\mathbb{R})$ taking values in $C^{0,\beta-\alpha}(\mathbb{R})$. This is a nontrivial task, and as such the use of the formula for the NMC without PV (see Proposition 3.4) become crucial.

The first question that may come to the reader's mind, is the existence of global continuous $(\Sigma_{u_a})_{a\in(0,1)}$, branches of nonlocal Delaunay hypersurfaces and the transition from unduloids to a periodic array of spheres as in the local case ($\alpha = 1$) discussed in Section 2.1. It is indeed an open question. However some key differences are expected in the embedded regime $a \in (0,1)$. Indeed, two results suggest that as the bifurcation parameter a varies from 0 to 1, the hypersurfaces Σ_{u_a} should approach an infinite compound of not-round-spheres. The first one being Dávila, del Pino, Dipierro, and Valdinoci [2016], where the authors proved that the enclosed sets of their (weak) CNMC Delauney hypersurface, as their constraint volume goes to zero, tend in measure (more precisely, in the so called Fraenkel asymmetry) to a periodic array of balls. The second one is a consequence of the work of Cabré, Weth and the author in Cabré, M. M. Fall, and Weth [2017], where it is proven that an array of periodic disjoint round spheres is not necessary a CNMC hypersurface. We detail a bit more on the latter result in the following section.

3.3.2 Near-sphere lattices with CNMC. In this section, we consider CNMC hypersurfaces given by an infinite compound of aligned round spheres, tangent or disconnected, enlightening possible limiting configuration of nonlocal unduloids Σ_{u_a} (from Theorem 3.10) as the bifurcation parameter becomes large. In a more general setting, we can look for CNMC hypersurfaces which are countable union of a certain bounded domain. To be precise, we assume $N \geq 2$ and let

$$S := S^{N-1} \subset \mathbb{R}^N$$

denote the unit centered sphere of $\mathbb{R}^N$. For $M \in \mathbb{N}$ with $1 \leq M \leq N$ we regard $\mathbb{R}^M$ as a subspace of $\mathbb{R}^N$ by identifying $x' \in \mathbb{R}^M$ with $(x',0) \in \mathbb{R}^M \times \mathbb{R}^{N-M} = \mathbb{R}^N$. Let $\{\mathbf{a}_1;\dots;\mathbf{a}_M\}$ be a basis of $\mathbb{R}^M$. By the above identification, we then consider the M-dimensional lattice

$$\mathcal{L} = \left\{\sum_{i=1}^{M} k_i \mathbf{a}_i \ : \ k = (k_1,\dots,k_M) \in \mathbb{Z}^M\right\} \tag{3.12}$$

as a subset of $\mathbb{R}^N$. In the case where $\{\mathbf{a}_1; \dots; \mathbf{a}_M\}$ is an orthogonal or an orthonormal basis, we say that $\mathcal{L}$ is a rectangular lattice or a square lattice, respectively. We define, for $r > 0$,

$$\mathcal{S}_0^r := S + r\mathcal{L} := \bigcup_{p \in \mathcal{L}} \Big(S + rp\Big) \subset \mathbb{R}^N.$$

For r large enough, the set $\mathcal{S}_0^r$ is the union of disjoint unit spheres centered at the lattice points in $r\mathcal{L}$. Consequently, $\mathcal{S}_0^r$ is a set of constant classical mean curvature (equal to one). In contrast, we shall see that the NMC function $H_\alpha(\mathcal{S}_0^r; \cdot)$ is in general *not* constant on $\mathcal{S}_0^r$. However nearby $\mathcal{S}_0^r$, one may expect CNMC sets for r large enough. Indeed, as $r \to \infty$, the hypersurface $\mathcal{S}_0^r$ tends to the single centered sphere S, which has CNMC. On the other hand, by invariance under translations $\mathbb{R}^N$, the linearized NMC operator about S has an N-dimensional kernel spanned by the coordinates functions $x_1, \dots, x_N$. Taking advantages on the invariance of the NMC operator by even reflections, we shall see that this program is realizable. Indeed, we consider the open set

$$\mathcal{O} := \{\varphi \in C^{1,\beta}(S) \; : \; \|\varphi\|_{L^\infty(S)} < 1, \; \varphi \text{ is even on } S\},$$

with $\beta \in (\alpha, 1)$, and the deformed sphere $S_\varphi := \{(1 + \varphi(\sigma))\sigma \; : \; \sigma \in S\}, \qquad \varphi \in \mathcal{O}$. Provided that $r > 0$ is large enough, the deformed sphere lattice

$$\mathcal{S}_\varphi^r := S_\varphi + r\mathcal{L} := \bigcup_{p \in \mathcal{L}} \Big(S_\varphi + rp\Big)$$

is a noncompact periodic hypersurface of class $C^{1,\beta}$. We have the following result.

Theorem 3.11 (Cabré, M. M. Fall, and Weth [2017]). *Let $\alpha \in (0,1)$, $\beta \in (\alpha, 1)$, $N \geq 2$, $1 \leq M \leq N$ and $\mathcal{L}$ be an M-dimensional lattice. Then, there exist $r_0 > 0$, and a C^2-curve $(r_0, +\infty) \to \mathcal{O}$, $r \mapsto \varphi_r$ such that for every $r \in (r_0, +\infty)$, the hypersurface $S_{\varphi_r} + r\mathcal{L}$ has constant nonlocal mean curvature given by $H_\alpha(S_{\varphi_r} + r\mathcal{L}; \cdot) \equiv H_\alpha(S; \cdot)$. Moreover if $1 \leq M \leq N - 1$, then the functions φ_r, $r > r_0$, are non-constant on S.*

As a consequence a periodic array of aligned spheres does not necessarily have constant nonlocal mean curvature. A Taylor expansion of the perturbation φ_r, as $r \to \infty$, shows how the near-spheres interact to form a CNMC hypersurface. To see this, we consider the linearized operator for the NMC operator acting on graphs on the unit sphere S, given by

$$\varphi \mapsto 2(L_\alpha\varphi - \lambda_1\varphi),$$

where

$$L_\alpha\varphi(\theta) = PV \int_S \frac{\varphi(\theta) - \varphi(\sigma)}{|\theta - \sigma|^{N+\alpha}} \, dV(\sigma).$$

The operator L_α can be seen as a *spherical fractional Laplacian*, and the above integral is understood in the principle value sense. It has the spherical harmonics as eigenfunctions corresponding to an increasing sequence of eigenvalues $\lambda_0 = 0 < \lambda_1 < \lambda_2 < \dots$.
The function φ_r in Theorem 3.11 expands as

$$\varphi_r(\theta) = r^{-N-\alpha}\left(-\kappa_0 + r^{-2}\left\{\kappa_1 \sum_{p\in\mathcal{L}_*} \frac{(\theta\cdot p)^2}{|p|^{N+\alpha+4}} - \kappa_2\right\} + o\left(r^{-2}\right)\right) \quad \text{for } \theta \in S$$

as $r \to +\infty$, where $\mathcal{L}_* := \mathcal{L}\setminus\{0\}$, for some positive constants $\kappa_0, \kappa_2, \kappa_3$, only depending on $N, \alpha, \lambda_1, \lambda_2$ and on $\mathcal{L}$. Since $\kappa_0 > 0$, the above expansion shows that, for large r, the perturbed spheres S_{φ_r} become smaller than S as the perturbation parameter r decreases. With regard to the order $r^{-N-\alpha}$, the shrinking process is uniform on S, whereas non-uniform deformations of the spheres may appear at the order $r^{-N-\alpha-2}$. In the case $M = N$, it is not known whether φ_r is not constant on S. In fact, we have the following more explicit form of φ_r in the case of square lattices. Assume that $\mathcal{L}$ is a square lattice. Then

$$\varphi_r(\theta) = r^{-N-\alpha}\left(-\kappa_0 + r^{-2}\left\{\tilde{\kappa}_1 \sum_{j=1}^{M} \theta_j^2 - \kappa_2\right\} + o(r^{-2})\right) \quad \text{for } \theta \in S \text{ as } r \to +\infty,$$

where $\tilde{\kappa}_1 = \dfrac{\kappa_1}{M} \displaystyle\sum_{p\in\mathcal{L}_*} \frac{1}{|p|^{N+\alpha+2}}$. In particular, if $M = N$, then

$$\varphi_r(\theta) = r^{-N-\alpha}\left(-\kappa_0 + r^{-2}\left(\tilde{\kappa}_1 - \kappa_2\right) + o(r^{-2})\right) \qquad \text{as } r \to \infty.$$

Hence the deformation of the lattice $S_{\varphi_r} + r\mathcal{L}$ is uniform up to the order $r^{-N-\alpha-2}$. It is conjectured in Cabré, M. M. Fall, and Weth [ibid.] that that $H_\alpha(\mathcal{S}_0^r;\cdot)$ is non-constant for any N-dimensional lattice $\mathcal{L}$, as long as $\mathcal{S}_0^r$ is a hypersurface of class $C^{1,\beta}$, $\beta \in (\alpha, 1)$.
The proof of Theorem 3.11, is based on the application of the implicit function theorem, to solve the r-parameter dependence fractional quasilinear elliptic type problem on the sphere,

$$H_\alpha(\mathcal{S}_\varphi^r;\cdot) = H_\alpha(S_\varphi;\cdot) + r^{-N-\alpha}G(r,\varphi;\cdot) = H_\alpha(S_0;\cdot).$$

The term containing G in the above identity is a lower order term. In fact the function G together with all its derivatives in the φ variables are bounded, for r large enough. On the other hand the leading quasilinear operator given by $H_\alpha(S_\varphi;\cdot)$ is the NMC operator acting on graphs on the sphere. Its expression, derived from (3.1), is explicitly given in the following

Proposition 3.12. *Let $\psi \in C^{1,\beta}(S^{N-1})$ with $\psi > 0$. Then the NMC of the hypersurface*

$$\Sigma_\psi := \left\{\sigma\psi(\sigma) \,:\, \sigma \in S^{N-1}\right\}$$

at the point $q = \theta\psi(\theta)$ is given by

$$\begin{aligned} h_\alpha(\psi)(\theta) &:= \frac{\alpha}{2} H_\alpha(\Sigma_\psi; q) \\ &= -\psi(\theta)\int_S \frac{\psi(\theta) - \psi(\sigma) - (\theta-\sigma)\cdot\nabla\psi(\sigma)}{|\theta-\sigma|^{N+\alpha}}\psi^{N-2}(\sigma)\,\mathcal{K}(\psi,\sigma,\theta)\,dV(\sigma) \\ &\qquad + \int_S \frac{(\psi(\theta)-\psi(\sigma))^2}{|\theta-\sigma|^{N+\alpha}}\psi^{N-2}(\sigma)\,\mathcal{K}(\psi,\sigma,\theta)\,dV(\sigma) \\ &\qquad + \frac{\psi(\theta)}{2}\int_S \frac{\psi^{N-1}(\sigma)}{|\theta-\sigma|^{N+\alpha-2}}\,\mathcal{K}(\psi,\sigma,\theta)\,dV(\sigma), \end{aligned}$$

where $\mathcal{K}(\psi,\sigma,\theta) := \dfrac{1}{\left(\frac{(\psi(\theta)-\psi(\sigma))^2}{|\theta-\sigma|^2} + \psi(\sigma)\psi(\theta)\right)^{(N+\alpha)/2}}$. Moreover, all integrals above converge absolutely.

Establishing the regularity of the NMC operator h_α, appearing in Proposition 3.12, as a map from open subsets of $C^{1,\beta}(S^{N-1})$ and taking values in $C^{0,\beta-\alpha}(S^{N-1})$, turns out to be an involved task. The inconveniences in the above expression is the presence of θ in the singular kernel $|\theta-\sigma|^{-N-\alpha}$ and the fact that it involves only euclidean distance instead of geodesic distance on the sphere.

4 Serrin's overdetermined problems

We consider the problem of finding domains (not necessarily bounded) and functions $u \in C^2(\overline{\Omega})$ such that

$$(4.1) \qquad -\Delta_g u = 1 \quad \text{in } \Omega, \qquad u = 0, \quad \partial_\nu u = const. \quad \text{on } \partial\Omega,$$

A domain Ω is called a *Serrin domain* if it is of class C^2 and if (4.1) admits a solution. System (4.1) was considered in eulidean space by J. Serrin in 1971 in his seminal paper Serrin [1971].

Theorem 4.1 (Serrin [ibid.], Weinberger [1971])**.** *Bounded Serrin domains in Euclidean space are balls.*

Serrin's argument relies on Alexandrov's moving plane method, with refined comparison principle. The proof of Weinberger [ibid.] uses the P-function method, Pohozaev identity and maximum principles. Serrin's result can be also derived from Alexandrov's rigidity result. Namely if (4.1) has a solution then $\partial\Omega$ has CMC and thus is a ball, see Farina and Kawohl [2008] and Choulli and Henrot [1998].
The additional Neumann boundary condition in (4.1) arises in many applications as a shape optimization problem for the underlying domain Ω. For a detailed discussion of some applications e.g. in fluid dynamics and the linear theory of torsion, see Serrin [1971] and Sirakov [2002]. Moreover, as observed in Minlend [2017], Serrin domains arise also in the context of Cheeger sets in a Riemannian framework. To explain this connection more precisely, we recall that the Cheeger constant of a Lipschitz subdomain $\Omega \subset \mathfrak{M}$ is given by

$$h(\Omega) := \inf_{A\subset\Omega} \frac{P(A)}{|A|}.$$

Here the infimum is taken over measurable subsets $A \subset \Omega$, with finite perimeter $P(A)$ and where $|A|$ denotes the volume of A (both with respect to the metric g). If $h(\Omega)$ is uniquely attained by Ω itself, then Ω is called uniquely self-Cheeger. By means of the Weinberger's P-function method, it is shown in Minlend [ibid.] that every bounded Serrin domain in a compact Riemannian manifold $\mathfrak{M}$, with Ricci curvature bounded below by some constant, is uniquely self-Cheeger. Cheeger constants play an important role in eigenvalue estimates on Riemannian manifolds (see Chavel [1984]), whereas in the classical Euclidean case $(\mathfrak{M}, g) = (\mathbb{R}^N, g_{eucl})$ these notions have applications in the denoising problem in image processing, see e.g. Parini [2011] and Leonardi [2015].
The above Serrin's classification result was extended in Kumaresan and Prajapat [1998] to subdomains Ω of the round hemis-sphere $\mathfrak{M} = S^N$ or $\mathfrak{M}$ is a space forms of constant negative sectional curvatures. More precisely, it is proved in Kumaresan and Prajapat [ibid.] that any smooth Serrin domain Ω contained in a hemisphere of S^N is a geodesic ball while the geodesic ball is the only Serrin domain in space forms of constant negative sectional curvatures. It is an interesting and widely open problem to construct and classify Serrin domains.

Minlend and the author in M. M. Fall and Minlend [2015], found that on any compact Riemaniann manifold, there exists a Serrin domain, which is a perturbation of a geodesic balls. On the hand the symmetry group of the ambient manifold might be used to find nontrivial Serrin domains with different geometry. Indeed, very recently, Minlend, Weth and the author in M. M. Fall, Minlend, and Weth [2018, 2017] considered the cases $\mathfrak{M} = S^N$ the unit sphere and the case $\mathfrak{M} = \mathbb{R}^n \times \mathbb{R}^m / 2\pi\mathbb{Z}^m$, endowed with the flat metric, proving existences of Serrin domains with nonconstant principal curvatures of the boundary.

From an analytic point of view, the question of finding Serrin domains share similar features to the one of finding CNMC hypersurfaces. Indeed, an admissible class of parametrizations φ of the unknown domain boundary is considered. Then the solvability condition is formulated as an operator equation of the form $H(\varphi) = const.$, where H is a nonlinear Dirichlet-to-Neumann operator, thus sharing same traits with a quasilinear nonlocal operators. In fact, the nonlocal character of this operator become apparent when computing its linearization along a trivial branch of the problem, see e.g. M. M. Fall, Minlend, and Weth [2018, 2017]. Here we end up with a pseudo-differential operator of order 1 that shares many features with the square-root of the negative of the Laplace operator, see Guillen, Kitagawa, and Schwab [2017].

From a geometric point of view, the results in M. M. Fall, Minlend, and Weth [2018, 2017], Cabré, M. M. Fall, and Weth [2018], and Cabré, Fall, Moustapha, Solà-Morales, and Weth [n.d.] support naturally the perception that the structure of the set of Serrin domains in a manifold $(\mathfrak{M}, g)$ share similarites to those of CNMC hypersurfaces, which both, up to a *dimesion shift*, share some structure to CMC hypersurfaces in $\mathfrak{M}$. Indeed, as in the theory CNMC hypersurfaces, there exists in $\mathbb{R}^2$, non-straight periodic Serrin domains, see M. M. Fall, Minlend, and Weth [2017]. We also emphasize from M. M. Fall, Minlend, and Weth [2018] that there exist Serrin domains in S^2 which are not bounded by geodesic circles, whereas CMC-hypersurfaces in S^2 are obviously trivial, i.e., they are geodesic circles.

Next, we recall that Alexandrov also proved in Alexandrov [1962] that any closed embedded CMC hypersurface contained in a hemisphere of S^N is a geodesic sphere. An explicit family of embedded CMC hypersurfaces in S^N with nonconstant principal curvatures was found in Perdomo [2010] in the case $N \geq 3$. These hypersurfaces seem somewhat related to the Serrin domains in M. M. Fall, Minlend, and Weth [2018], although they bound a tubular neighborhood of S^1 and not of S^{N-1}.

Overdetermined boundary value problems involving different elliptic equations has intensively studied recently. Starting with the pioneering papers of Pacard and Sicbaldi [2009], Sicbaldi [2010] and Hauswirth, Hélein, and Pacard [2011], the construction of nontrivial domains giving rise to solutions of overdetermined problems has been performed in many specific settings, see e.g. Del Pino, Pacard, and Wei [2015], Schlenk and Sicbaldi [2012], Ros, Ruiz, and Sicbaldi [2016], and Morabito and Sicbaldi [2016]. Moreover, the rigidity results for these domains were derived in Farina and Valdinoci [2010a,b], Ros and Sicbaldi [2013], Ros, Ruiz, and Sicbaldi [2017], and Traizet [2014].

References

Nicola Abatangelo and Enrico Valdinoci (2014). "A notion of nonlocal curvature". *Numer. Funct. Anal. Optim.* 35.7-9, pp. 793–815. MR: 3230079 (cit. on p. 1637).

A. D. Aleksandrov (1958). "Uniqueness theorems for surfaces in the large. V". *Vestnik Leningrad. Univ.* 13.19, pp. 5–8. MR: 0102114 (cit. on p. 1633).

A. D. Alexandrov (1962). "A characteristic property of spheres". *Ann. Mat. Pura Appl. (4)* 58, pp. 303–315. MR: 0143162 (cit. on p. 1652).

Luigi Ambrosio, Guido De Philippis, and Luca Martinazzi (2011). "Gamma-convergence of nonlocal perimeter functionals". *Manuscripta Math.* 134.3-4, pp. 377–403. MR: 2765717 (cit. on p. 1638).

Begoña Barrios, Alessio Figalli, and Enrico Valdinoci (2014). "Bootstrap regularity for integro-differential operators and its application to nonlocal minimal surfaces". *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* 13.3, pp. 609–639. MR: 3331523 (cit. on p. 1639).

Claudia Bucur and Enrico Valdinoci (2016). *Nonlocal diffusion and applications*. Vol. 20. Lecture Notes of the Unione Matematica Italiana. Springer, [Cham]; Unione Matematica Italiana, Bologna, pp. xii+155. MR: 3469920 (cit. on pp. 1639, 1640).

Xavier Cabre, Eleonora Cinti, and Joaquim Serra (Oct. 2017). "Stable s-minimal cones in $\mathbb{R}^3$ are flat for $s \sim 1$". arXiv: 1710.08722 (cit. on p. 1639).

Xavier Cabre and Matteo Cozzi (Nov. 2017). "A gradient estimate for nonlocal minimal graphs". arXiv: 1711.08232 (cit. on pp. 1639, 1640).

Xavier Cabré, Mouhamed Moustapha Fall, and Tobias Weth (2017). "Near-sphere lattices with constant nonlocal mean curvature". *Math. Ann.* (Cit. on pp. 1647–1649).

– (2018). "Delaunay hypersurfaces with constant nonlocal mean curvature". *J. Math. Pures Appl. (9)* 110, pp. 32–70. MR: 3744919 (cit. on pp. 1641, 1643, 1646, 1652).

Xavier Cabré, Fall, Mouhamed Moustapha, J. Solà-Morales, and Tobias Weth (n.d.). "Curves and surfaces with constant nonlocal mean curvature: meeting Alexandrov and Delaunay". To appear in J. Reine Angew. Math. (cit. on pp. 1644–1646, 1652).

L. Caffarelli, J.-M. Roquejoffre, and O. Savin (2010). "Nonlocal minimal surfaces". *Comm. Pure Appl. Math.* 63.9, pp. 1111–1144. MR: 2675483 (cit. on pp. 1638, 1639, 1646).

Luis A. Caffarelli and Panagiotis E. Souganidis (2010). "Convergence of nonlocal threshold dynamics approximations to front propagation". *Arch. Ration. Mech. Anal.* 195.1, pp. 1–23. MR: 2564467 (cit. on p. 1631).

Luis Caffarelli and Enrico Valdinoci (2011). "Uniform estimates and limiting arguments for nonlocal minimal surfaces". *Calc. Var. Partial Differential Equations* 41.1-2, pp. 203–240. MR: 2782803 (cit. on p. 1639).

– (2013). "Regularity properties of nonlocal minimal surfaces via limiting arguments". *Adv. Math.* 248, pp. 843–871. arXiv: 1105.115. MR: 3107529 (cit. on p. 1639).

Isaac Chavel (1984). *Eigenvalues in Riemannian geometry*. Vol. 115. Pure and Applied Mathematics. Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk. Academic Press, Inc., Orlando, FL, pp. xiv+362. MR: 768584 (cit. on p. 1651).

Mourad Choulli and Antoine Henrot (1998). "Use of the domain derivative to prove symmetry results in partial differential equations". *Math. Nachr.* 192, pp. 91–103. MR: 1626395 (cit. on p. 1651).

Eleonora Cinti, Juan Davila, and Manuel Del Pino (2016). "Solutions of the fractional Allen-Cahn equation which are invariant under screw motion". *J. Lond. Math. Soc. (2)* 94.1, pp. 295–313. MR: 3532174 (cit. on p. 1639).

Eleonora Cinti, Joaquim Serra, and Enrico Valdinoci (Feb. 2016). "Quantitative flatness results and BV-estimates for stable nonlocal minimal surfaces". arXiv: 1602.00540 (cit. on p. 1639).

Eleonora Cinti, Carlo Sinestrari, and Enrico Valdinoci (July 2016). "Neckpinch singularities in fractional mean curvature flows". arXiv: 1607.08032 (cit. on p. 1640).

Giulio Ciraolo, Alessio Figalli, Francesco Maggi, and M. Novaga (n.d.). "Rigidity and sharp stability estimates for hypersurfaces with constant and almost-constant nonlocal mean curvature". To appear in J. Reine Angew. Math. (cit. on pp. 1644, 1645).

Michael G. Crandall and Paul H. Rabinowitz (1971). "Bifurcation from simple eigenvalues". *J. Functional Analysis* 8, pp. 321–340. MR: 0288640 (cit. on pp. 1636, 1647).

J. Dávila, M. del Pino, and J. Wei (n.d.). "Nonlocal s-minimal surfaces and Lawson cones". To appear in *J. Diff. Geom.* (cit. on pp. 1637–1639).

Juan Dávila, Manuel del Pino, Serena Dipierro, and Enrico Valdinoci (2016). "Nonlocal Delaunay surfaces". *Nonlinear Anal.* 137, pp. 357–380. MR: 3485130 (cit. on pp. 1646, 1647).

Manuel Del Pino, Frank Pacard, and Juncheng Wei (2015). "Serrin's overdetermined problem and constant mean curvature surfaces". *Duke Math. J.* 164.14, pp. 2643–2722. MR: 3417183 (cit. on p. 1652).

Ch. Delaunay (1841). "Sur la surface de révolution dont la courbure moyenne est constante". 6, pp. 309–315 (cit. on pp. 1635, 1636).

Serena Dipierro, Joaquim Serra, and Enrico Valdinoci (Nov. 2016). "Improvement of flatness for nonlocal phase transitions". arXiv: 1611.10105.

Mouhamed Moustapha Fall (Nov. 2017). "Periodic patterns for a model involving short-range and long-range interactions". arXiv: 1711.10825 (cit. on p. 1632).

Mouhamed Moustapha Fall and Sven Jarohs (2015). "Overdetermined problems with fractional Laplacian". *ESAIM Control Optim. Calc. Var.* 21.4, pp. 924–938. MR: 3395749 (cit. on p. 1645).

Mouhamed Moustapha Fall and Ignace Aristide Minlend (2015). "Serrin's overdetermined problem on Riemannian manifolds". *Adv. Calc. Var.* 8.4, pp. 371–400. MR: 3403432 (cit. on p. 1651).

Mouhamed Moustapha Fall, Ignace Aristide Minlend, and Tobias Weth (2017). "Unbounded periodic solutions to Serrin's overdetermined boundary value problem". *Arch. Ration. Mech. Anal.* 223.2, pp. 737–759. MR: 3590664 (cit. on pp. 1651, 1652).

– (2018). "Serrin's overdetermined problem on the sphere". *Calc. Var. Partial Differential Equations* 57.1, Art. 3, 24. MR: 3723376 (cit. on pp. 1651, 1652).

A. Farina and B. Kawohl (2008). "Remarks on an overdetermined boundary value problem". *Calc. Var. Partial Differential Equations* 31.3, pp. 351–357. MR: 2366129 (cit. on p. 1651).

Alberto Farina, Luciano Mari, and Enrico Valdinoci (2013). "Splitting theorems, symmetry results and overdetermined problems for Riemannian manifolds". *Comm. Partial Differential Equations* 38.10, pp. 1818–1862. MR: 3169764.

Alberto Farina and Enrico Valdinoci (2010a). "Flattening results for elliptic PDEs in unbounded domains with applications to overdetermined problems". *Arch. Ration. Mech. Anal.* 195.3, pp. 1025–1058. MR: 2591980 (cit. on p. 1652).

– (2010b). "Overdetermined problems in unbounded domains with Lipschitz singularities". *Rev. Mat. Iberoam.* 26.3, pp. 965–974. MR: 2789372 (cit. on p. 1652).

A. Figalli, N. Fusco, F. Maggi, V. Millot, and M. Morini (2015). "Isoperimetry and stability properties of balls with respect to nonlocal energies". *Comm. Math. Phys.* 336.1, pp. 441–507. MR: 3322379 (cit. on pp. 1638, 1639).

Alessio Figalli and Enrico Valdinoci (2017). "Regularity and Bernstein-type results for nonlocal minimal surfaces". *J. Reine Angew. Math.* 729, pp. 263–273. MR: 3680376 (cit. on p. 1639).

Robert Finn (1986). *Equilibrium capillary surfaces*. Vol. 284. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, New York, pp. xvi+245. MR: 816345 (cit. on p. 1631).

Rupert L. Frank and Robert Seiringer (2008). "Non-linear ground state representations and sharp Hardy inequalities". *J. Funct. Anal.* 255.12, pp. 3407–3430. MR: 2469027 (cit. on p. 1638).

Nicola Fusco, Vincent Millot, and Massimiliano Morini (2011). "A quantitative isoperimetric inequality for fractional perimeters". *J. Funct. Anal.* 261.3, pp. 697–715. MR: 2799577 (cit. on p. 1638).

A. M. Garsia and E. Rodemich (1974). "Monotonicity of certain functionals under rearrangement". *Ann. Inst. Fourier (Grenoble)* 24.2. Colloque International sur les Processus Gaussiens et les Distributions Aléatoires (Colloque Internat. du CNRS, No. 222, Strasbourg, 1973), pp. vi, 67–116. MR: 0414802 (cit. on p. 1638).

David Gilbarg and Neil S. Trudinger (1983). *Elliptic partial differential equations of second order*. Second. Vol. 224. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, pp. xiii+513. MR: 737190 (cit. on p. 1634).

Nestor Guillen, Jun Kitagawa, and Russell W. Schwab (Oct. 2017). "Estimates for Dirichlet-to-Neumann maps as integro-differential operators". arXiv: 1710.03152 (cit. on p. 1652).

Laurent Hauswirth, Frédéric Hélein, and Frank Pacard (2011). "On an overdetermined elliptic problem". *Pacific J. Math.* 250.2, pp. 319–334. MR: 2794602 (cit. on p. 1652).

Heinz Hopf (1989). *Differential geometry in the large*. Second. Vol. 1000. Lecture Notes in Mathematics. Notes taken by Peter Lax and John W. Gray, With a preface by S. S. Chern, With a preface by K. Voss. Springer-Verlag, Berlin, pp. viii+184. MR: 1013786 (cit. on p. 1633).

Wu-yi Hsiang (1982). "Generalized rotational hypersurfaces of constant mean curvature in the Euclidean spaces. I". *J. Differential Geom.* 17.2, pp. 337–356. MR: 664499 (cit. on p. 1633).

Wu-yi Hsiang and Wen Ci Yu (1981). "A generalization of a theorem of Delaunay". *J. Differential Geom.* 16.2, pp. 161–177. MR: 638783 (cit. on p. 1636).

Cyril Imbert (2009). "Level set approach for fractional mean curvature flows". *Interfaces Free Bound.* 11.1, pp. 153–176. MR: 2487027 (cit. on p. 1640).

J. H. Jellet (1853). "Sur la Surface dont la Courbure Moyenne est Constante". *J. Math. Pures Appl.* 18, pp. 163–167 (cit. on p. 1633).

Nicolaos Kapouleas (1991). "Compact constant mean curvature surfaces in Euclidean three-space". *J. Differential Geom.* 33.3, pp. 683–715. MR: 1100207 (cit. on p. 1633).

Katsuei Kenmotsu (1979). "Weierstrass formula for surfaces of prescribed mean curvature". *Math. Ann.* 245.2, pp. 89–99. MR: 552581 (cit. on p. 1635).

Hansjörg Kielhöfer (2004). *Bifurcation theory*. Vol. 156. Applied Mathematical Sciences. An introduction with applications to PDEs. Springer-Verlag, New York, pp. viii+346. MR: 2004250.

S. Kumaresan and Jyotshana Prajapat (1998). "Serrin's result for hyperbolic space and sphere". *Duke Math. J.* 91.1, pp. 17–28. MR: 1487977 (cit. on p. 1651).

Gian Paolo Leonardi (2015). "An overview on the Cheeger problem". In: *New trends in shape optimization*. Vol. 166. Internat. Ser. Numer. Math. Birkhäuser/Springer, Cham, pp. 117–139. MR: 3467379 (cit. on p. 1651).

Francesco Maggi and Enrico Valdinoci (2017). "Capillarity problems with nonlocal surface tension energies". *Comm. Partial Differential Equations* 42.9, pp. 1403–1446. arXiv: 1606.08610. MR: 3717439 (cit. on p. 1640).

William H. Meeks III, Joaquín Pérez, and Giuseppe Tinaglia (2016). "Constant mean curvature surfaces". In: *Surveys in differential geometry 2016. Advances in geometry and mathematical physics*. Vol. 21. Surv. Differ. Geom. Int. Press, Somerville, MA, pp. 179–287. arXiv: 1605.02512. MR: 3525098 (cit. on p. 1633).

Cornelia Mihaila (Oct. 2017). "Axial symmetry for fractional capillarity droplets". arXiv: 1710.03421 (cit. on p. 1640).

Ignace Aristide Minlend (2017). "Existence of self-Cheeger sets on Riemannian manifolds". *Arch. Math. (Basel)* 109.4, pp. 393–400. MR: 3698326 (cit. on p. 1651).

Filippo Morabito and Pieralberto Sicbaldi (2016). "Delaunay type domains for an overdetermined elliptic problem in $\mathbb{S}^n \times \mathbb{R}$ and $\mathbb{H}^n \times \mathbb{R}$". *ESAIM Control Optim. Calc. Var.* 22.1, pp. 1–28. MR: 3489374 (cit. on p. 1652).

Frank Pacard and Pieralberto Sicbaldi (2009). "Extremal domains for the first eigenvalue of the Laplace-Beltrami operator". *Ann. Inst. Fourier (Grenoble)* 59.2, pp. 515–542. MR: 2521426 (cit. on p. 1652).

Enea Parini (2011). "An introduction to the Cheeger problem". *Surv. Math. Appl.* 6, pp. 9–21. MR: 2832554 (cit. on p. 1651).

Oscar M. Perdomo (2010). "Embedded constant mean curvature hypersurfaces on spheres". *Asian J. Math.* 14.1, pp. 73–108. MR: 2726595 (cit. on p. 1652).

Antonio Ros, David Ruiz, and Pieralberto Sicbaldi (Sept. 2016). "Solutions to overdetermined elliptic problems in nontrivial exterior domains". arXiv: 1609.03739 (cit. on p. 1652).

– (2017). "A rigidity result for overdetermined elliptic problems in the plane". *Comm. Pure Appl. Math.* 70.7, pp. 1223–1252. MR: 3666566 (cit. on p. 1652).

Antonio Ros and Pieralberto Sicbaldi (2013). "Geometry and topology of some overdetermined elliptic problems". *J. Differential Equations* 255.5, pp. 951–977. MR: 3062759 (cit. on p. 1652).

Mariel Sáez and Enrico Valdinoci (Nov. 2015). "On the evolution by fractional mean curvature". arXiv: 1511.06944 (cit. on p. 1640).

Ovidiu Savin and Enrico Valdinoci (2012). "Γ-convergence for nonlocal phase transitions". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 29.4, pp. 479–500. MR: 2948285 (cit. on p. 1632).

– (2013). "Regularity of nonlocal minimal cones in dimension 2". *Calc. Var. Partial Differential Equations* 48.1-2, pp. 33–39. MR: 3090533 (cit. on p. 1639).

Felix Schlenk and Pieralberto Sicbaldi (2012). "Bifurcating extremal domains for the first eigenvalue of the Laplacian". *Adv. Math.* 229.1, pp. 602–632. MR: 2854185 (cit. on pp. 1636, 1652).

James Serrin (1971). "A symmetry problem in potential theory". *Arch. Rational Mech. Anal.* 43, pp. 304–318. MR: 0333220 (cit. on pp. 1650, 1651).

Pieralberto Sicbaldi (2010). "New extremal domains for the first eigenvalue of the Laplacian in flat tori". *Calc. Var. Partial Differential Equations* 37.3-4, pp. 329–344. MR: 2592974 (cit. on p. 1652).

B. Sirakov (2002). "Overdetermined Elliptic Problems in Physics". In: *Nonlinear PDE's in Condensed Matter and Reactive Flows*. Ed. by Henri Berestycki and Yves Pomeau. Springer (cit. on p. 1651).

Joel Smoller (1994). *Shock waves and reaction-diffusion equations*. Second. Vol. 258. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, New York, pp. xxiv+632. MR: 1301779.

Michael Spivak (1979). *A comprehensive introduction to differential geometry. Vol. V*. Second. Publish or Perish, Inc., Wilmington, Del., pp. viii+661. MR: 532834 (cit. on p. 1633).

Martin Traizet (2014). "Classification of the solutions to an overdetermined elliptic problem in the plane". *Geom. Funct. Anal.* 24.2, pp. 690–720. MR: 3192039 (cit. on p. 1652).

H. F. Weinberger (1971). "Remark on the preceding paper of Serrin". *Arch. Rational Mech. Anal.* 43, pp. 319–320. MR: 0333221 (cit. on pp. 1650, 1651).

Henry C. Wente (1986). "Counterexample to a conjecture of H. Hopf". *Pacific J. Math.* 121.1, pp. 193–243. MR: 815044 (cit. on p. 1633).

Received 2017-11-30.

Mouhamed Moustapha Fall
African Institute for Mathematical Sciences of Senegal
KM 2, Route de Joal
B.P. 14 18. Mbour
Sénégal
mouhamed.m.fall@aims-senegal.org

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1657–1690)

RIGIDITY FOR VON NEUMANN ALGEBRAS

Adrian Ioana

Abstract

We survey some of the progress made recently in the classification of von Neumann algebras arising from countable groups and their measure preserving actions on probability spaces. We emphasize results which provide classes of (W*-superrigid) actions that can be completely recovered from their von Neumann algebras and II$_1$ factors that have a unique Cartan subalgebra. We also present cocycle superrigidity theorems and some of their applications to orbit equivalence. Finally, we discuss several recent rigidity results for von Neumann algebras associated to groups.

1 Introduction

A *von Neumann algebra* is an algebra of bounded linear operators on a Hilbert space which is closed under the adjoint operation and in the weak operator topology. Von Neumann algebras arise naturally from countable groups and their actions on probability spaces, via two seminal constructions of Murray and von Neumann [1936, 1943]. Given a countable group Γ, the left regular representation of Γ on $\ell^2\Gamma$ generates the *group von Neumann algebra* $L(\Gamma)$. Equivalently, $L(\Gamma)$ is the weak operator closure of the complex group algebra $\mathbb{C}\Gamma$ acting on $\ell^2\Gamma$ by left convolution. Every nonsingular action $\Gamma \curvearrowright (X,\mu)$ of a countable group Γ on a probability space (X,μ) gives rise to the *group measure space von Neumann algebra* $L^\infty(X) \rtimes \Gamma$.

A central theme in the theory of von Neumann algebras is the classification of $L(\Gamma)$ in terms of the group Γ and of $L^\infty(X) \rtimes \Gamma$ in terms of the group action $\Gamma \curvearrowright (X,\mu)$. These problems are typically studied when Γ has infinite non-trivial conjugacy classes (icc) and when $\Gamma \curvearrowright (X,\mu)$ is free ergodic and measure preserving, respectively. These assumptions guarantee that the corresponding algebras are *II$_1$ factors*: indecomposable infinite dimensional von Neumann algebras which admit a trace. Moreover, it follows

Supported in part by NSF Career Grant DMS #1253402 and a Sloan Foundation Fellowship.
MSC2010: primary 46L36; secondary 37A20, 28D15.
Keywords: von Neumann algebra, II$_1$ factor, measure preserving action, group measure space construction, Cartan subalgebra, Bernoulli action, W*-superrigidity.

that $L^\infty(X)$ is a *Cartan subalgebra* of $L^\infty(X) \rtimes \Gamma$, that is, a maximal abelian subalgebra whose normalizer generates $L^\infty(X) \rtimes \Gamma$.

The classification of II_1 factors is a rich subject, with deep connections to several areas of mathematics. Over the years, it has repeatedly provided fertile ground for the development of new, exciting theories: Jones' subfactor theory, Voiculescu's free probability theory, and Popa's deformation/rigidity theory. The subject has been connected to group theory and ergodic theory from its very beginning, via the group measure space construction Murray and von Neumann [1936]. Later on, Singer made the observation that the isomorphism class of $L^\infty(X) \rtimes \Gamma$ only depends on the equivalence relation given by the orbits of $\Gamma \curvearrowright (X, \mu)$ Singer [1955]. This soon led to a new branch of ergodic theory, which studies group actions up to orbit equivalence Dye [1959]. Orbit equivalence theory, further developed in the 1980s (see Ornstein and Weiss [1980], Connes, Feldman, and Weiss [1981], and Zimmer [1984]), has seen an explosion of activity in the last twenty years (see Shalom [2005], Furman [2011], and Gaboriau [2010]). This progress has been in part triggered by the success of the deformation/rigidity approach to the classification of II_1 factors (see Popa [2007b], Vaes [2010], and Ioana [2013]). More broadly, this approach has generated a wide range of applications to ergodic theory and descriptive set theory, including:

- the existence of non-orbit equivalent actions of non-amenable groups Gaboriau and Popa [2005], Ioana [2011c], and Epstein [n.d.].

- cocycle superrigidity theorems for Bernoulli actions in Popa [2007a, 2008], leading to examples of non-Bernoullian factors of Bernoulli actions for a class of countable groups Popa [2006b] and Popa and Sasyk [2007], and the solution of some open problems in descriptive set theory Thomas [2009]

- cocycle superrigidity theorems for profinite actions Ioana [2011b], Furman [2011], and Gaboriau, Ioana, and Tucker-Drob [n.d.], leading to an explicit uncountable family of Borel incomparable treeable equivalence relations Ioana [2016].

- solid ergodicity of Bernoulli actions Chifan and Ioana [2010] (see also Ozawa [2006]).

The classification of II_1 factors is governed by a strong amenable/non-amenable dichotomy. Early work in this area culminated with Connes' celebrated theorem from the mid 1970s: all II_1 factors arising from infinite amenable groups and their actions are isomorphic to the hyperfinite II_1 factor Connes [1976]. Amenable groups thus manifest a striking absence of rigidity: any property of the group or action, other than the amenability of the group, is lost in the passage to von Neumann algebras.

In contrast, it gradually became clear that in the non-amenable case various aspects of groups and actions are remembered by their von Neumann algebras. Thus, non-amenable groups were used in McDuff [1969] and Connes [1975] to construct large families of non-isomorphic II_1 factors. Rigidity phenomena for von Neumann algebras first emerged in the work of Connes from the early 1980s Connes [1980]. He showed that II_1 factors arising from property (T) groups have countable symmetry (fundamental and outer automorphism) groups. Representation theoretic properties of groups (Kazhdan's property (T), Haagerup's property, weak amenability) were then used to prove unexpected non-embeddability results for II_1 factors associated to certain lattices in Lie groups Connes and Jones [1985] and Cowling and Haagerup [1989]. But while these results showcased the richness of the theory, the classification problem for non-amenable II_1 factors remained by and large intractable.

A major breakthrough in the classification of II_1 factors was made by Popa with his invention of deformation/rigidity theory Popa [2006a,d,e] (see the surveys Popa [2007b] and Vaes [2007]). The studied II_1 factors, M, admit a distinguished subalgebra A (e.g., $L^\infty(X)$ or $L(\Gamma)$ when $M = L^\infty(X) \rtimes \Gamma$) such that the inclusion $A \subset M$ satisfies both a deformation and a rigidity property. Popa discovered that the combination of these properties can be used to detect the position of A inside M, or even recover the underlying structure of M (e.g., the group Γ and action $\Gamma \curvearrowright X$ when $M = L^\infty(X) \rtimes \Gamma$). He also developed a series of powerful technical tools to exploit this principle.

Popa first implemented this idea and techniques to provide a class of II_1 factors, M, which admit a unique Cartan subalgebra, A, with the relative property (T) Popa [2006a]. The uniqueness of A implies that any invariant of the inclusion $A \subset M$ is in fact an invariant of M. When applied to $M = L^\infty(\mathbb{T}^2) \rtimes \mathrm{SL}_2(\mathbb{Z})$, it follows that the fundamental group of M is equal to that of the equivalence relation of the action $\mathrm{SL}_2(\mathbb{Z}) \curvearrowright \mathbb{T}^2$. Since the latter is trivial by Gaboriau's work Gaboriau [2000, 2002], this makes M the first example of a II_1 factor with trivial fundamental group Popa [2006a], thereby solving a longstanding problem.

In Popa [2006d,e], Popa greatly broadened the scope of deformation/rigidity theory by obtaining the first strong rigidity theorem for group measure space factors. To make this precise, suppose that $\Gamma \curvearrowright (X, \mu)$ is a Bernoulli action of an icc group Γ and $\Lambda \curvearrowright (Y, \nu)$ is a free ergodic probability measure preserving action of a property (T) group Λ (e.g., $\Lambda = \mathrm{SL}_{n\geq 3}(\mathbb{Z})$). Under this assumptions, it is shown in Popa [2006d,e] that if the group measure space factors $L^\infty(X) \rtimes \Gamma$ and $L^\infty(Y) \rtimes \Lambda$ are isomorphic, then the groups Γ and Λ are isomorphic and their actions are conjugate.

The goal of this survey is to present some of the progress achieved recently in the classification of II_1 factors. We focus on advances made since 2010, and refer the reader to Popa [2007b] and Vaes [2010] for earlier developments. A topic covered there but omitted here is the calculation of symmetry groups of II_1 factors; see Popa [2006a,d], Ioana, Peterson, and Popa [2008], and Popa and Vaes [2010a] for several key results in this direction. There is some overlap with Ioana [2013], but overall the selection of topics and presentation are quite different.

We start with a section of preliminaries (Section 2) and continue with a discussion of the main ideas from Popa's deformation/rigidity theory (Section 3). Before giving an overview of Sections 4-6, we first recall some terminology and background. Two free ergodic p.m.p. (probability measure preserving) actions $\Gamma \curvearrowright (X,\mu)$ and $\Lambda \curvearrowright (Y,\nu)$ are called:

1. **conjugate** if there exist an isomorphism of probability spaces $\alpha : (X,\mu) \to (Y,\nu)$ and an isomorphism of groups $\delta : \Gamma \to \Lambda$ such that $\alpha(g \cdot x) = \delta(g) \cdot \alpha(x)$, for all $g \in \Gamma$ and almost every $x \in X$.

2. **orbit equivalent (OE)** if there exists an isomorphism of probability spaces $\alpha : (X,\mu) \to (Y,\nu)$ such that $\alpha(\Gamma \cdot x) = \Lambda \cdot \alpha(x)$, for almost every $x \in X$.

3. **W*-equivalent** if $L^\infty(X) \rtimes \Gamma$ is isomorphic to $L^\infty(Y) \rtimes \Lambda$.

Singer showed that OE amounts to the existence of an isomorphism $L^\infty(X) \rtimes \Gamma \cong L^\infty(Y) \rtimes \Lambda$ which identifies the Cartan subalgebras $L^\infty(X)$ and $L^\infty(Y)$ Singer [1955]. Thus, OE implies W*-equivalence. Since conjugacy clearly implies orbit equivalence, putting these together we have:

$$\text{conjugacy} \implies \text{orbit equivalence} \implies \text{W}^*\text{-equivalence}$$

Rigidity usually refers to a situation in which a weak equivalence between two objects can be used to show that the objects are equivalent in a much stronger sense or even isomorphic. In the present context, rigidity occurs whenever some of the above implications can be reversed for all actions $\Gamma \curvearrowright (X,\mu)$ and $\Lambda \curvearrowright (Y,\nu)$ belonging to two classes of actions. The most extreme form of rigidity, called superrigidity, happens when this can be achieved without any restrictions on the second class of actions. Thus, an action $\Gamma \curvearrowright (X,\mu)$ is *W*-superrigid* (respectively, *OE-superrigid*) if any free ergodic p.m.p. action $\Lambda \curvearrowright (Y,\nu)$ which is W*-equivalent (respectively, OE) to $\Gamma \curvearrowright (X,\mu)$ must be conjugate to it. In other words, the conjugacy class of the action can be entirely reconstructed from the isomorphism class of its von Neumann algebra (respectively, its orbit equivalence class).

The seminal results from Popa [2006a,d,e] suggested two rigidity conjectures which have guided much of the work in the area in the ensuing years. First, Popa [2006a] provided a class of II_1 factors $L^\infty(X) \rtimes \mathbb{F}_n$ associated to actions of the free groups $\mathbb{F}_n$, with $n \geq 2$, for which $L^\infty(X)$ is the unique Cartan subalgebra satisfying the relative property (T). This led to the conjecture that the same might be true for arbitrary Cartan subalgebras of arbitrary free group measure space factors: **(A)** $L^\infty(X) \rtimes \mathbb{F}_n$ *has a unique Cartan* *subalgebra, up to unitary conjugacy, for any free ergodic p.m.p. action* $\mathbb{F}_n \curvearrowright (X, \mu)$ *of* $\mathbb{F}_n$ *with* $n \geq 2$. Second, Popa [2006d,e] showed that within the class of Bernoulli actions of icc property (T) groups, W*-equivalence implies conjugacy. Moreover, it was proved in Popa [2007a] that such actions are OE-superrigid. These results naturally led to the following conjecture: **(B)** *Bernoulli actions* $\Gamma \curvearrowright (X, \mu)$ *of icc property (T) groups* Γ *are W***-superrigid.*

In Section 4, we discuss the positive resolutions of the above conjectures. These were the culmination of a period of intense activity which has generated a series of striking unique Cartan decomposition and W*-superrigidity results. The first breakthrough was made by Ozawa and Popa who confirmed conjecture **(A)** in the case of profinite actions Ozawa and Popa [2010a]. The class of groups whose profinite actions give rise to II$_1$ factors with a unique group measure space decomposition was then shown to be much larger in Ozawa and Popa [2010a,b] and Peterson [n.d.].

But, since none of these actions was known to be OE-superrigid, W*-superrigidity could not be deduced. Indeed, proving that an action $\Gamma \curvearrowright (X, \mu)$ is W*-superrigid amounts to showing that the action is OE-superrigid *and* that $L^\infty(X) \rtimes \Gamma$ has a unique *group measure space Cartan subalgebra*[1]. Nevertheless, Peterson was able to show the existence of "virtually" W*-superrigid actions Peterson [n.d.]. Soon after, Popa and Vaes discovered a large class of amalgamated free product groups Γ whose every free ergodic p.m.p. action $\Gamma \curvearrowright (X, \mu)$ gives rise to a II$_1$ factor with a unique group measure space Cartan subalgebra, up to unitary conjugacy Popa and Vaes [2010b]. Applying OE-superrigidity results from Popa [2007a, 2008] and Kida [2011] enabled them to provide the first concrete classes of W*-superrigid actions. However, these results do not apply to actions of property (T) groups and so despite all of this progress, conjecture **(B)** remained open until it was eventually settled by the author in Ioana [2011a].

New uniqueness theorems for group measure space Cartan subalgebras were then obtained in Chifan and Peterson [2013] and in Ioana [2012b], while Chifan and Sinclair extended the results of Ozawa and Popa [2010a] from free groups to hyperbolic groups in Chifan and Sinclair [2013]. However, the available uniqueness results for Cartan subalgebras required either a rigidity property of the group (excluding the free groups), or the action to be in a specific class. Thus, the situation for arbitrary actions of the free groups remained unclear until conjecture **(A)** was resolved by Popa and Vaes in their breakthrough work Popa and Vaes [2014a]. Subsequently, several additional families of groups were shown to satisfy conjecture **(A)**, including non-elementary hyperbolic groups Popa and

[1]This terminology is used to distinguish the Cartan subalgebras coming from the group measure space construction from general Cartan subalgebras, see Section 2.6.

Vaes [2014b], free product groups Ioana [2013], and central quotients of braid groups Chifan, Ioana, and Kida [2015], while conjecture **(B)** was established for the more general class of mixing Gaussian actions in Boutonnet [2013].

Section 5 is devoted to rigidity results in orbit equivalence. After recalling the pioneering rigidity and superrigidity phenomena discovered by Zimmer [1980, 1984] and Furman [1999], we discuss several recent cocycle superrigidity results. Our starting point is Popa's remarkable cocycle superrigidity theorem: any cocycle for a Bernoulli action of a property (T) group into a countable group is cohomologous to a homomorphism Popa [2007a]. The property (T) assumption was removed in Popa [2008], where the same was shown to hold for Bernoulli actions of products of non-amenable groups. A cocycle superrigidity theorem for a different class of actions of property (T) groups, the profinite actions, was then obtained in Ioana [2011b]. Answering a question motivated by the analogy with Bernoulli actions, this theorem was recently extended to product groups in Gaboriau, Ioana, and Tucker-Drob [n.d.].

These and many other results provide large classes of "rigid" groups (including property (T) groups, product groups, and by Kida's work Kida [2010], most mapping class groups) which admit OE-superrigid actions. In contrast, other non-amenable groups, notably the free groups $\mathbb{F}_n$, posses no OE-superigid actions. Nevertheless, as we explain in the second part of Section 4, a general OE-rigidity theorem for profinite actions was discovered in Ioana [2016]. This result imposes no assumptions on the acting groups and so it applies, in novel fashion, to actions of groups such as $\mathbb{F}_n$ and $\mathrm{SL}_2(\mathbb{Z})$. As an application, it led to a continuum of mutually non-OE and Borel incomparable actions of $\mathrm{SL}_2(\mathbb{Z})$, confirming a conjecture from Thomas [2003, 2006]. It also motivated a "local spectral gap" theorem for dense subgroups of simple Lie groups in Boutonnet, Ioana, and Golsefidy [2017].

In Section 6, we discuss recent rigidity results for group von Neumann algebras. These give instances when certain algebraic properties of groups, such as the absence or presence of a direct product decomposition, are remembered by their von Neumann algebras. A remarkable result of Ozawa shows that for any icc hyperbolic group Γ, the II_1 factor $L(\Gamma)$ is prime, i.e. it cannot be decomposed as a tensor product of two II_1 factors Ozawa [2004]. In particular, this recovered the primeness of $L(\mathbb{F}_n)$, for $n \geq 2$, which was first proved in Ge [1998] using Voiculescu's free probability techniques. Later on, several other large classes of icc groups Γ were shown to give rise to prime II_1 factors (see e.g. Ozawa [2006], Peterson [2009], Popa [2007c], and Chifan and Houdayer [2010]). However, all such groups, Γ, satisfy various properties which relate them closely to lattices in rank one simple Lie groups. On the other hand, it is an open problem whether II_1 factors associated to icc irreducible lattices Γ in higher rank simple or semisimple Lie groups G are prime. We present in Section 6 a result from Drimbe, D. Hoff, and Ioana [n.d.] which answers this positively in the case when G is a product of simple Lie groups of rank one.

We continue with the recent finding in Chifan, de Santiago, and Sinclair [2016] and Chifan and Ioana [2018] of large classes of product and amalgamated free product groups whose product (respectively, amalgam) structure can be recognized from their von Neumann algebras. Finally, we turn our attention to the strongest type of rigidity for group II_1 factors $L(\Gamma)$, called W*-superrigidity. This occurs when the isomorphism class of Γ can be reconstructed from the isomorphism class of $L(\Gamma)$. We conclude the section with a discussion of the first examples of W*-superrigid groups discovered in our joint work Ioana, Popa, and Vaes [2013], and the subsequent examples exhibited in Berbec and Vaes [2014] and Chifan and Ioana [2018].

Let us mention a few exciting topics related to the classification of II_1 factors which have received a lot of attention recently but are not covered here, due to limitations of space. First, we point out the impressive work of Houdayer and his co-authors (including Houdayer and Vaes [2013], Boutonnet, Houdayer, and Raum [2014], Houdayer and Isono [2017], and Boutonnet, Houdayer, and Vaes [n.d.]) where the deformation/rigidity framework is adapted to study von Neumann algebras of type III. A notable advance in this direction is the classification theorem for free Araki-Woods factors obtained by Houdayer, Shlyakhtenko, and Vaes [n.d.]. We also highlight Peterson's remarkable work Peterson [2015] (see also Creutz and Peterson [n.d.]) which shows that lattices in higher rank simple Lie groups admit a unique II_1 factor representation, the regular representation, thus solving a conjecture of Connes from the 1980s. Finally, we mention the model theory for II_1 factors which was introduced in Farah, Hart, and Sherman [2013, 2014a,b]. Subsequently, our joint work with Boutonnet, Chifan, and Ioana [2017] settled a basic question in the theory. More precisely, we showed the existence of uncountably many different elementary classes of II_1 factors (equivalently, of uncountably many II_1 factors with pairwise non-isomorphic ultrapowers).

Acknowledgments. It is my pleasure to thank Rémi Boutonnet, Cyril Houdayer, and Sorin Popa for many comments that helped improve the exposition.

2 Preliminaries

2.1 Tracial von Neumann algebras. A von Neumann algebra M is called *tracial* if it admits a linear functional $\tau : M \to \mathbb{C}$, called a *trace*, which is

1. *positive*: $\tau(x^*x) \geq 0$, for all $x \in M$.

2. *faithful*: $\tau(x^*x) = 0$, for some $x \in M$, implies that $x = 0$.

3. *normal*: $\tau(\sum_{i\in I} p_i) = \sum_{i\in I} \tau(p_i)$, for any family $\{p_i\}_{i\in I}$ of mutually orthogonal projections.

4. *tracial*: $\tau(xy) = \tau(yx)$, for all $x, y \in M$.

A von Neumann algebra with trivial center is called a *factor*. An infinite dimensional tracial factor is called a *II_1 factor*. Note that any II_1 factor M admits a unique trace τ such that $\tau(1) = 1$.

Any tracial von Neumann algebra (M, τ) admits a canonical (or *standard*) representation on a Hilbert space. Denote by $L^2(M)$ the completion of M with respect to the 2-norm $\|x\|_2 := \sqrt{\tau(x^*x)}$. Then the left and right multiplication actions of M on itself give rise to representations of M and its opposite algebra M^{op} on $L^2(M)$. This makes $L^2(M)$ a *Hilbert M-bimodule*, i.e. a Hilbert space $\mathcal{H}$ endowed with commuting normal representations $M \subset \mathbb{B}(\mathcal{H})$ and $M^{\mathrm{op}} \subset \mathbb{B}(\mathcal{H})$.

Let $P \subset M$ be a von Neumann subalgebra and denote by $E_P : M \to P$ the unique τ-preserving conditional expectation onto P. Specifically, E_P is determined by the identity $\tau(E_P(x)y) = \tau(xy)$, for all $x \in M$ and $y \in P$. By completing the algebraic tensor product $M \otimes_{\mathrm{alg}} M$ with respect to the scalar product $\langle x_1 \otimes x_2, y_1 \otimes y_2 \rangle = \tau(y_2^* E_P(x_2^* x_1) y_1)$ we obtain the Hilbert M-bimodule $L^2(M) \bar{\otimes}_P L^2(M)$. Alternatively, this bimodule can be realized as the L^2-space of Jones' basic construction $\langle M, P\rangle$. In the case $P = M$ and $P = \mathbb{C}1$, we recover the so-called *trivial* and *coarse* bimodules, $L^2(M)$ and $L^2(M) \bar{\otimes} L^2(M)$, respectively.

2.2 Group von Neumann algebras. Let Γ be a countable group and denote by $\{\delta_h\}_{h\in\Gamma}$ the usual orthonormal basis of $\ell^2\Gamma$. The *left regular representation* $u : \Gamma \to \mathcal{U}(\ell^2\Gamma)$ is given by $u_g(\delta_h) = \delta_{gh}$. The *group von Neumann algebra* $L(\Gamma)$ is defined as the weak operator closure of the linear span of $\{u_g\}_{g\in\Gamma}$. It is a tracial von Neumann algebra with a trace $\tau : L(\Gamma) \to \mathbb{C}$ given by $\tau(x) = \langle x\delta_e, \delta_e\rangle$. Equivalently, τ is the unique trace on $L(\Gamma)$ satisfying $\tau(u_g) = \delta_{g,e}$, for all $g \in \Gamma$.

Note that $L(\Gamma)$ is a II_1 factor if and only if Γ is *icc*: $\{ghg^{-1} | g \in \Gamma\}$ is infinite, for all $h \in \Gamma \setminus \{e\}$.

2.3 Group measure space von Neumann algebras. Let $\Gamma \curvearrowright (X, \mu)$ be a p.m.p. action of a countable group Γ on a probability space (X, μ). For $g \in \Gamma$ and $c \in L^2(X)$, let $\sigma_g(c) \in L^2(X)$ be defined by $\sigma_g(c)(x) = c(g^{-1}\cdot x)$. The elements of both Γ and $L^\infty(X)$ can be represented as operators on the Hilbert space $L^2(X)\bar{\otimes}\ell^2\Gamma$ through the formulae $u_g(c \otimes \delta_h) = \sigma_g(c) \otimes \delta_{gh}$ and $b(c \otimes \delta_h) = bc \otimes \delta_h$. Then u_g is a unitary operator and $u_g b u_g^* = \sigma_g(b)$, for all $g \in \Gamma$ and $b \in L^\infty(X)$. As a consequence, the linear span of $\{bu_g | b \in L^\infty(X), g \in \Gamma\}$ is a $*$-algebra.

The *group measure space von Neumann algebra* $L^\infty(X) \rtimes \Gamma$ is defined as the weak operator closure of the linear span of $\{bu_g | b \in L^\infty(X), g \in \Gamma\}$. It is a tracial von Neumann algebra with a trace $\tau : L^\infty(X) \rtimes \Gamma \to \mathbb{C}$ given by $\tau(a) = \langle a(1 \otimes \delta_e), 1 \otimes \delta_e\rangle$.

Equivalently, τ is the unique trace on $L^\infty(X) \rtimes \Gamma$ satisfying $\tau(bu_g) = \delta_{g,e} \int_X b \, d\mu$, for all $g \in \Gamma$ and $b \in L^\infty(X)$.

Any element $a \in L^\infty(X) \rtimes \Gamma$ admits a *Fourier decomposition* $a = \sum_{g\in\Gamma} a_g u_g$, with $a_g \in L^\infty(X)$ and the series converging in $\|.\|_2$. These so-called *Fourier coefficients* $\{a_g\}_{g\in\Gamma}$ of a are given by $a_g = E_{L^\infty(X)}(au_g^*)$ and satisfy

$$\sum_{g\in\Gamma} \|a_g\|_2^2 = \|a\|_2^2.$$

If the action $\Gamma \curvearrowright (X, \mu)$ is (essentially) free and ergodic, then $L^\infty(X) \rtimes \Gamma$ is a II$_1$ factor. Recall that the action $\Gamma \curvearrowright (X, \mu)$ is called *free* if $\{x | \; g \cdot x = x\}$ is a null set, for all $g \in \Gamma \setminus \{e\}$, and *ergodic* if any Γ-invariant measurable subset $A \subset X$ must satisfy $\mu(A) \in \{0, 1\}$.

2.4 Examples of free ergodic p.m.p. actions.

1. Let Γ be a countable group and (X_0, μ_0) be a non-trivial probability space. Then Γ acts on the space X_0^Γ of sequences $x = (x_h)_{h\in\Gamma}$ by shifting the indices: $g \cdot x = (x_{g^{-1}h})_{h\in\Gamma}$. This action, called the *Bernoulli action* with base (X_0, μ_0), preserves the product probability measure μ_0^Γ, and is free and ergodic.

2. Let Γ be a countable group together with a dense embedding into a compact group G. The *left translation action* of Γ on G given by $g \cdot x = gx$ preserves the Haar measure $\mathbf{m}_G$ of G, and is free and ergodic. For instance, assume that Γ is residually finite, and let $G = \varprojlim \Gamma/\Gamma_n$ be its *profinite competition* with respect to a chain $\Gamma = \Gamma_0 > \Gamma_1 > \dots > \Gamma_n > \dots$ of finite index normal subgroups with trivial intersection. Then G is a profinite hence totally disconnected compact group, and the map $g \mapsto (g\Gamma_n)_n$ gives a dense embedding of Γ into G. At the opposite end, we have the case when G is a connected compact group (e.g. $G = SO(n)$) and Γ is a countable dense subgroup of G.

3. Generalizing example (2), a p.m.p. action $\Gamma \curvearrowright (X, \mu)$ is called *compact* if the closure of Γ in $\mathrm{Aut}(X, \mu)$ is compact, and *profinite* if it is an inverse limit of actions $\Gamma \curvearrowright (X_n, \mu_n)$, with X_n a finite set, for all n. Any profinite p.m.p. action is compact. Any ergodic compact p.m.p. action is isomorphic to a left translation action of the form $\Gamma \curvearrowright (G/K, \mathbf{m}_{G/K})$, where G is a compact group containing Γ densely, $K < G$ is a closed subgroup and $\mathbf{m}_{G/K}$ is the unique G-invariant Borel probability measure on G/K. Any ergodic profinite p.m.p. action is of this form, with G a profinite group.

4. Finally, the standard action of $\mathrm{SL}_n(\mathbb{Z})$ on the n-torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ preserves the Lebesgue measure, and is free and ergodic. The left multiplication action of $\mathrm{PSL}_n(\mathbb{Z})$ on $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$ preserves the unique $\mathrm{SL}_n(\mathbb{R})$-invariant probability measure, and is free and ergodic.

2.5 Equivalence relations. An equivalence relation $\mathcal{R}$ on a standard probability space (X,μ) is called *countable p.m.p.* if $\mathcal{R}$ has countable classes, $\mathcal{R}$ is a Borel subset of $X \times X$, and any Borel automorphism of X whose graph is contained in $\mathcal{R}$ preserves μ Feldman and Moore [1977]. If $\Gamma \curvearrowright (X,\mu)$ is a p.m.p. action of a countable group Γ, its *orbit equivalence relation* $\mathcal{R}_{\Gamma\curvearrowright X} := \{(x,y) \in X^2 \mid \Gamma\cdot x = \Gamma\cdot y\}$ is a countable p.m.p. equivalence relation. It is now clear that two p.m.p. actions of countable groups $\Gamma \curvearrowright (X,\mu)$ and $\Lambda \curvearrowright (Y,\nu)$ are orbit equivalent precisely when their OE relations are isomorphic: $(\alpha\times\alpha)(\mathcal{R}_{\Gamma\curvearrowright X}) = \mathcal{R}_{\Lambda\curvearrowright Y}$, for some isomorphism of probability spaces $\alpha : (X,\mu) \to (Y,\nu)$.

2.6 Cartan subalgebras. Let M be a II_1 factor. The *normalizer* of a subalgebra $A \subset M$, denoted by $\mathcal{N}_M(A)$, is the group of unitaries $u \in M$ satisfying $uAu^* = A$. An abelian von Neumann subalgebra $A \subset M$ is called a *Cartan subalgebra* if it is maximal abelian and its normalizer generates M. For instance, if $\Gamma \curvearrowright (X,\mu)$ is a free ergodic p.m.p. action, then $A := L^\infty(X)$ is a Cartan subalgebra of the group measure space II_1 factor $M := L^\infty(X) \rtimes \Gamma$. To distinguish such Cartan subalgebras from arbitrary ones we call them of *group measure space* type.

In general, any Cartan subalgebra inclusion $A \subset M$ can be identified with an inclusion of the form $L^\infty(X) \subset L(\mathcal{R}, w)$, where (X,μ) is a probability space and $L(\mathcal{R}, w)$ is the von Neumann algebra associated to a countable p.m.p. equivalence relation $\mathcal{R}$ on X and a 2-cocycle $w \in \mathrm{H}^2(\mathcal{R}, \mathbb{T})$ Feldman and Moore [ibid.]. By Feldman and Moore [ibid.], $\mathcal{R}$ arises as the OE relation of a p.m.p action $\Gamma \curvearrowright (X,\mu)$. If the action is free and the cocycle is trivial, then we canonically have $M = L^\infty(X) \rtimes \Gamma$. However, the action cannot always be chosen to be free Furman [1999], and thus not all Cartan subalgebras are of group measure space type.

The next proposition makes clear the importance of Cartan subalgebras in the study of group measure space factors.

Proposition 2.1 (Singer [1955])**.** *If $\Gamma \curvearrowright (X,\mu)$ and $\Lambda \curvearrowright (Y,\nu)$ are free ergodic p.m.p. actions, then the following conditions are equivalent*

1. *the actions $\Gamma \curvearrowright (X,\mu)$ and $\Lambda \curvearrowright (Y,\nu)$ are orbit equivalent.*
2. *there exists a $*$-isomorphism $\theta : L^\infty(X)\rtimes\Gamma \to L^\infty(Y)\rtimes\Lambda$ such that $\theta(L^\infty(X)) = L^\infty(Y)$.*

Proposition 2.1 is extremely useful in two ways. First, the implication (1) $\Rightarrow$ (2) provides an approach to the study of orbit equivalence of actions using von Neumann algebras. This has been instrumental in several developments, including the finding of non-OE actions of non-amenable groups (see Gaboriau and Popa [2005] and Ioana [2009, 2011c]) and of new OE-superrigidity results (see Popa [2007a, 2008] and Ioana [2011b]). On the other hand, the implication (2) $\Rightarrow$ (1) allows one to reduce the classification of group measure space factors to the classification of the corresponding actions up to orbit equivalence, whenever uniqueness of group measure space Cartan subalgebras can be established. This has been used for instance to exhibit the first families of W*-superrigid actions in Peterson [n.d.], Popa and Vaes [2010b], and Ioana [2011a].

2.7 Amenability and property (T). In the early 1980s, Connes discovered that Hilbert bimodules provide an appropriate representation theory for tracial von Neumann algebras, paralleling the theory of unitary representations for groups (see Connes [1982] and Popa [1986]).

To illustrate this point, assume that $M = L(\Gamma)$, for a countable group Γ. Given a unitary representation $\pi : \Gamma \to \mathcal{U}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$, the Hilbert space $\mathcal{H}\bar{\otimes}\ell^2\Gamma$ carries a natural Hilbert M-bimodule structure: $u_g(\xi \otimes \delta_h)u_k = \pi(g)(\xi) \otimes \delta_{ghk}$. Moreover, the map $\xi \mapsto \xi \otimes \delta_e$ turns sequences of Γ-almost invariant unit vectors into sequences of M-almost central tracial vectors. Here, for a Hilbert M-bimodule $\mathcal{K}$ and a subalgebra $P \subset M$, we say that a vector $\xi \in \mathcal{K}$ is *tracial* if $\langle x\xi, \xi\rangle = \langle \xi x, \xi\rangle = \tau(x)$, for all $x \in M$, and *P-central* if $y\xi = \xi y$, for all $y \in P$. A net of vectors $\xi_n \in \mathcal{K}$ is called *P-almost central* if $\|y\xi_n - \xi_n y\| \to 0$, for all $y \in P$.

The analogy between representations and bimodules led to von Neumann algebraic analogues of various representation theoretic properties of groups, including amenability and property (T):

Definition 2.2. Let (M, τ) be a tracial von Neumann algebra, and $P, Q \subset M$ be subalgebras.

1. We say that M is *amenable* if there exists a net $\xi_n \in L^2(M)\bar{\otimes}L^2(M)$ of tracial, M-almost central vectors Popa [1986]. We say that *Q is amenable relative to P* if there exists a net $\xi_n \in L^2(M)\bar{\otimes}_P L^2(M)$ of tracial, Q-almost central vectors Ozawa and Popa [2010a].

2. We say that M has *property (T)* if any Hilbert M-bimodule without M-central vectors does not admit a net of M-almost central unit vectors Connes and Jones [1985]. We say that $P \subset M$ has the *relative property (T)* if any Hilbert M-bimodule without P-central vectors does not admit a net of M-almost central, tracial vectors Popa [2006a].

2.8 Popa's intertwining-by-bimodules. In Popa [2006a,d], Popa developed a powerful technique for showing unitary conjugacy of subalgebras of a tracial von Neumann algebra.

Theorem 2.3 (Popa [2006a,d]). *If P, Q are von Neumann subalgebras of a separable tracial von Neumann algebra (M,τ), then the following are equivalent:*

1. *There is no sequence of unitaries $u_n \in P$ satisfying $\|E_Q(au_nb)\|_2 \to 0$, for all $a,b \in M$.*
2. *There are non-zero projections $p \in P, q \in Q$, a $*$-homomorphism $\theta : pPp \to qQq$, and a non-zero partial isometry $v \in qMp$ such that $\theta(x)v = vx$, for all $x \in pPp$.*

If these conditions hold, we say that a corner of P embeds into Q.

Moreover, if P and Q are Cartan subalgebras of M and a corner of P embeds into a corner of Q, then there is a unitary $u \in M$ such that $P = uQu^*$.

3 Popa's deformation/rigidity theory

3.1 Deformations. Since its introduction in the early 2000's, Popa's deformation/rigidity theory has had a transformative impact on the theory of von Neumann algebras. The theory builds on Popa's innovative idea of using the deformations of a II_1 factor to locate its rigid subalgebras. Before illustrating this principle with several examples, let us make precise the notion of a deformation.

Definition 3.1. A *deformation* of the identity of a tracial von Neumann algebra (M,τ) is a sequence of unital, trace preserving, completely positive maps $\phi_n : M \to M$ satisfying

$$\|\phi_n(x) - x\|_2 \to 0, \quad \text{for all } x \in M.$$

A linear map $\phi : M \to M$ is called *completely positive* if the map $\phi^{(m)} : \mathbb{M}_m(M) \to \mathbb{M}_m(M)$ given by $\phi^{(m)}([x_{i,j}]) = [\phi(x_{i,j})]$ is positive, for all $m \geq 1$. Note that any unital, trace preserving, completely positive map $\phi : M \to M$ extends to a contraction $\phi : L^2(M) \to L^2(M)$.

Remark 3.2. Deformations arise naturally from continuous families of automorphisms of larger von Neumann algebras. To be precise, let $(\tilde{M}, \tilde{\tau})$ be a tracial von Neumann algebra containing M such that $\tilde{\tau}_{|M} = \tau$. Assume that $(\theta_t)_{t\in\mathbb{R}}$ is a pointwise $\|.\|_2$-continuous family of trace preserving automorphisms of $\tilde{M}$ with $\theta_0 = \mathrm{id}$. Then $\phi_n := E_M \circ \theta_{t_n} : M \to M$ defines a deformation of M, for any sequence $t_n \to 0$. Abusing notation, such pairs $(\tilde{M}, (\theta_t)_{t\in\mathbb{R}})$ are also called deformations of M.

Next, we proceed to give three examples of deformations. For a comprehensive list of examples, we refer the reader to Ioana [2015, Section 3].

Example 3.3. First, let Γ be a countable group and $\varphi_n : \Gamma \to \mathbb{C}$ be a sequence of positive definite functions such that $\varphi_n(e) = 1$, for all n, and $\varphi_n(g) \to 1$, for all $g \in \Gamma$. Then $\phi_n(u_g) = \varphi_n(g)u_g$ defines a deformation of the group algebra $L(\Gamma)$ and $\phi_n(au_g) = \varphi_n(g)au_g$ defines a deformation of any group measure space algebra $L^\infty(X) \rtimes \Gamma$.

If Γ has Haagerup's property Haagerup [1978/79], then there is such a sequence $\varphi_n : \Gamma \to \mathbb{C}$ satisfying $\varphi_n \in c_0(\Gamma)$, for all n. When applied to $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, the above procedure gives a deformation of the II_1 factor $M = L^\infty(\mathbb{T}^2) \rtimes \mathrm{SL}_2(\mathbb{Z})$ which is compact relative to $L^\infty(\mathbb{T}^2)$. This fact was a crucial ingredient in Popa's proof that M has a trivial fundamental group Popa [2006a].

Example 3.4. Second, let $\Gamma \curvearrowright (X, \mu) := ([0,1]^\Gamma, \mathbf{Leb}^\Gamma)$ be the Bernoulli action of a countable group Γ. In Popa [2006c,d], Popa discovered that Bernoulli actions have a remarkable deformation property, called *malleability*: there is a continuous family of automorphisms $(\alpha_t)_{t\in\mathbb{R}}$ of the product space $X \times X$ which commute with diagonal action of Γ and satisfy $\alpha_0 = \mathrm{id}$ and $\alpha_1(x, y) = (y, x)$.

To see this, first construct a continuous family of automorphisms $(\alpha_t^0)_{t\in\mathbb{R}}$ of the probability space $[0,1] \times [0,1]$ such that $\alpha_0^0 = \mathrm{id}$ and $\alpha_1^0(x, y) = (y, x)$. For example, we can take

$$\alpha_t^0(x,y) = \begin{cases} (x,y), & \text{if } |x-y| \geq t \\ (y,x), & \text{if } |x-y| < t. \end{cases}$$

Then identify $X \times X = ([0,1] \times [0,1])^\Gamma$ and define $\alpha_t((x_g)_{g\in\Gamma}) = (\alpha_t^0(x_g))_{g\in\Gamma}$.

The automorphisms $(\alpha_t)_{t\in\mathbb{R}}$ of $X \times X$ give rise to a deformation of $M := L^\infty(X) \rtimes \Gamma$, as follows. Since α_t commutes with the diagonal action of Γ on $X \times X$, the formula $\theta_t(au_g) = (a \circ \alpha_t^{-1})u_g$ defines a trace preserving automorphism of $\tilde{M} := L^\infty(X \rtimes X) \rtimes \Gamma$. Thus, $(\theta_t)_{t\in\mathbb{R}}$ is a continuous family of automorphisms of $\tilde{M}$ such that $\theta_0 = \mathrm{id}$ and $\theta_1(a \otimes b) = b \otimes a$, for all $a, b \in L^\infty(X)$. Since M embeds into $\tilde{M}$ via the map $au_g \mapsto (a \otimes 1)u_g$, one obtains a deformation of M (see Remark 3.2).

Example 3.5. Finally, we recall from Popa [1986, 2007c] the construction of a malleable deformation for the free group factors, $L(\mathbb{F}_n)$. For simplicity, we consider the case $n = 2$ and put $M = L(\mathbb{F}_n)$. Denote by a_1, a_2, b_1, b_2 the generators of $\mathbb{F}_4$, and view $\mathbb{F}_2$ as the subgroup of $\mathbb{F}_4$ generated by a_1, a_2. This gives an embedding of M into $\tilde{M} = L(\mathbb{F}_4)$. If we see b_1 and b_2 as unitary elements of $\tilde{M}$, then we can find self-adjoint operators h_1 and h_2 such that $b_1 = \exp(ih_1)$ and $b_2 = \exp(ih_2)$. One can now define a 1-parameter group of automorphism $(\theta_t)_{t\in\mathbb{R}}$ of $\tilde{M}$ as follows:

$$\theta_t(a_1) = \exp(ith_1)a_1, \quad \theta_t(a_2) = \exp(ith_2)a_2, \quad \theta_t(b_1) = b_1, \quad \text{and} \quad \theta_t(b_2) = b_2.$$

3.2 Deformation vs. rigidity. We will now explain briefly and informally how Popa used these deformations to prove structural results for subalgebras satisfying various rigidity properties.

A main source of rigidity is provided by the relative property (T). Indeed, assume that $P \subset M$ is an inclusion with the relative property (T). The correspondence between completely positive maps and bimodules (see, e.g., Popa [2007b, Section 2.1]) implies that any deformation $\phi_n : M \to M$ must converge uniformly to the identity in $\|.\|_2$ on the unit ball of P Popa [2006a]. In particular, for any large enough $n_0 \geq 1$ one has that

$$\text{(i)} \quad \|\phi_{n_0}(u) - u\|_2 \leq 1/2 \quad \text{for all unitaries } u \in P.$$

In Popa [2006a,d], this analytical condition is combined with the intertwining-by-bimodule technique to deduce P can be unitarily conjugate into a distinguished subalgebra $Q \subset M$.

First, suppose that $M = L^\infty(\mathbb{T}^2) \rtimes \Gamma$, where $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, and denote $Q = L^\infty(\mathbb{T}^2)$. Since Γ has Haagerup's property, we can find positive definite functions $\varphi_n \in c_0(\Gamma)$ such that $\varphi_n \to 1$ pointwise. As in Example 3.3, we obtain a deformation $\phi_n : M \to M$ given by $\phi_n(au_g) = \varphi_n(g)au_g$. Consider the Fourier decomposition $u = \sum_{g\in\Gamma} E_Q(uu_g^*)u_g$ of a unitary u belonging to the subalgebra $P \subset M$ with the relative property (T). The specific formula of ϕ_{n_0} allows one to rewrite (i) as

$$\text{(ii)} \quad \sum_{g\in\Gamma} |\varphi_{n_0}(g) - 1|^2 \, \|E_Q(uu_g^*)\|_2^2 \leq 1/4 \quad \text{for all unitaries } u \in P.$$

Since $\varphi_{n_0} \in c_0(\Gamma)$, we have that $|\varphi_{n_0}(g) - 1|^2 \geq 1/2$, for all $g \in \Gamma$ outside a finite subset F. Taking into account that $\sum_{g\in\Gamma} \|E_Q(uu_g^*)\|_2^2 = \tau(u^*u) = 1$ and using (ii), one concludes that

$$\text{(iii)} \quad \sum_{g\in F} \|E_Q(uu_g^*)\|_2^2 \geq 1/2 \quad \text{for all unitaries } u \in P.$$

As a corollary, P does not admit a sequence a unitaries satisfying condition (1) of Theorem 2.3. In other words, a corner of P embeds into Q. If P is a Cartan subalgebra of M, then the moreover part of Theorem 2.3 implies that P must be unitarily conjugate to Q Popa [2006a].

Second, suppose that $M = L^\infty(X) \rtimes \Gamma$, where $\Gamma \curvearrowright (X,\mu) = ([0,1]^\Gamma, \mathbf{Leb}^\Gamma)$ is the Bernoulli action. Denote $Q = L(\Gamma)$ and let $(\tilde{M}, (\theta_t)_{t\in\mathbb{R}})$ the deformation introduced in Example 3.4. Then the deformation $\phi_n := E_M \circ \theta_{1/2^n}$ converges uniformly to the identity in $\|.\|_2$ on the unit ball of any subalgebra $P \subset M$ with the property (T). It is immediate that the same must be true for $\theta_{1/2^n}$. In particular, for any large enough $n_0 \geq 1$ one has that

$$\text{(iv)} \quad \|\theta_{1/2^{n_0}}(u) - u\|_2 \leq 1/2 \quad \text{for all unitaries } u \in P.$$

This implies the existence of a non-zero element $v \in \tilde{M}$ satisfying $\theta_{1/2^{n_0}}(u)v = vu$ for all unitaries $u \in P$. By employing a certain symmetry of the deformation $(\theta_t)_{t\in\mathbb{R}}$, Popa proved that the same holds for θ_1. Using the formula of the restriction θ_1 to M, it follows that a corner of P embeds into Q. Assuming that Γ is icc, Popa concludes that P can be unitarily conjugate into Q Popa [2006d].

The rigidity considered in Popa [2006a,d] is due to (relative) property (T) assumptions. In Popa [2008, 2007c], Popa discovered a less restrictive from of rigidity, arising from the presence of subalgebras with non-amenable relative commutant. To illustrate this, assume the context from Example 3.5: $M = L(\mathbb{F}_2) \subset \tilde{M} = L(\mathbb{F}_4)$ and $(\theta_t)_{t\in\mathbb{R}}$ is the 1-parameter group of automorphisms of $\tilde{M}$ defined therein. The deformation $(\tilde{M}, (\theta_t)_{t\in\mathbb{R}})$ has two crucial properties:

(a) the M-bimodule $L^2(\tilde{M}) \ominus L^2(M)$ is isomorphic to a multiple of $L^2(M)\bar{\otimes}L^2(M)$.

(b) the contraction $E_M \circ \theta_t : L^2(M) \to L^2(M)$ is a compact operator, for any $t > 0$.

Recovering Ozawa's solidity theorem Ozawa [2004] in the case of the free group factors, Popa proved in Popa [2007c] that the relative commutant $P' \cap M$ is amenable, for any diffuse subalgebra $P \subset M$. Let us explain how (a) and (b) are combined in Popa [ibid.] to deduce the following weaker statement: there is no diffuse subalgebra $P \subset M$ which commutes with a non-amenable II_1 subfactor $N \subset M$.

Assume that we can find such commuting subalgebras $P, N \subset M$. The non-amenability of N leads to a spectral gap condition for its coarse bimodule Connes [1976]: there is a finite set $F \subset N$ such that

$$\text{(v)} \quad \|\xi\| \leq \sum_{x\in F} \|x\cdot\xi - \xi\cdot x\| \quad \text{for all vectors } \xi \in L^2(N)\bar{\otimes}L^2(N).$$

The spectral gap condition is then used to establish the following rigidity property for P: the deformation $E_M \circ \theta_t : M \to M$ converges uniformly on the unit ball of P. One first notes that (a) implies that (v) holds for every $\xi \in \tilde{M}$ with $E_M(\xi) = 0$. Thus, we may take $\xi = \theta_t(u) - E_M(\theta_t(u))$, for any unitary $u \in P$ and $t > 0$. Using that u commutes with every $x \in F$, one derives that

$$\begin{aligned} \text{(vi)} \quad \|x\cdot\xi - \xi\cdot x\| &\leq \|x\theta_t(u) - \theta_t(u)x\|_2 \\ &= \|\theta_{-t}(x)u - u\theta_{-t}(x)\|_2 \\ &= \|(\theta_{-t}(x) - x)u - u(\theta_{-t}(x) - x)\|_2 \\ &\leq 2\|\theta_{-t}(x) - x\|_2. \end{aligned}$$

Combining (v) and (vi) gives that if $t > 0$ is chosen small enough then $\|\theta_t(u) - E_M(\theta_t(u))\|_2 \leq 1/2$, for all unitaries $u \in P$. Since P is assumed diffuse it contains a sequence of unitaries

u_n converging weakly to 0. The compactness of $E_M \circ \theta_t$ gives that $\|E_M(\theta_t(u_n))\|_2 \to 0$, leading to a contradiction.

4 Uniqueness of Cartan subalgebras and W*-superrigidity of Bernoulli actions

4.1 Uniqueness of Cartan subalgebras. The first result showing uniqueness, up to unitary conjugacy, of arbitrary Cartan subalgebras, was proved by Ozawa and Popa:

Theorem 4.1 (Ozawa and Popa [2010a]). *Let $\mathbb{F}_n \curvearrowright (X, \mu)$ be a free ergodic profinite p.m.p. action of $\mathbb{F}_n$, for $n \geq 2$. Then $M := L^\infty(X) \rtimes \mathbb{F}_n$ has a unique Cartan subalgebra, up to unitary conjugacy: if $P \subset M$ is any Cartan subalgebra, then $P = uL^\infty(X)u^*$, for some unitary element $u \in M$.*

The proof of Theorem 4.1 relies on the *complete metric approximation property* (CMAP) of $\mathbb{F}_n$, used as a weak form of a deformation of the II_1 factor $L(\mathbb{F}_n)$. Recall that a countable group Γ has the CMAP Haagerup [1978/79] if there exists a sequence of finitely supported functions $\varphi_k : \Gamma \to \mathbb{C}$ such that $\varphi_k(g) \to 1$, for all $g \in \Gamma$, and the linear maps $\phi_k : L(\Gamma) \to L(\Gamma)$ given by $\phi_k(u_g) = \varphi_k(g)u_g$, for all $g \in \Gamma$, satisfy $\limsup_k \|\phi_k\|_{\mathrm{cb}} = 1$. If the last condition is weakened by assuming instead that $\limsup_k \|\phi_k\|_{\mathrm{cb}} < \infty$, then Γ is called *weakly amenable* Cowling and Haagerup [1989]. Here, $\|\phi_k\|_{\mathrm{cb}}$ denotes the completely bounded norm of ϕ_k.

Since $\mathbb{F}_n$ has the CMAP Haagerup [1978/79] and the action $\mathbb{F}_n \curvearrowright X$ is profinite, the II_1 factor M also has the CMAP: there exists a sequence of finite rank completely bounded maps $\phi_k : M \to M$ such that $\|\phi_k(x) - x\|_2 \to 0$, for all $x \in M$, and $\limsup_k \|\phi_k\|_{\mathrm{cb}} = 1$.

Let $P \subset M$ be an arbitrary diffuse amenable subalgebra and denote by $\mathcal{G}$ its normalizer in M. If the conjugation action $\mathcal{G} \curvearrowright P$ happens to be compact, i.e. the closure of $\mathcal{G}$ inside $\mathrm{Aut}(P)$ is compact, then the unitary representation $\mathcal{G} \curvearrowright L^2(P)$ is a direct sum of finite dimensional representations. This provides many vectors $\xi \in L^2(P)\bar{\otimes}L^2(P)$ that are invariant under the diagonal action of $\mathcal{G}$. Indeed, if $\eta_1, \ldots, \eta_d$ is an orthonormal basis of any finite dimensional $\mathcal{G}$-invariant subspace of $L^2(P)$, then $\xi = \sum_{i=1}^d \eta_i \otimes \eta_i^*$ has this property.

Ozawa and Popa made the fundamental discovery that since M has the CMAP, the action $\mathcal{G} \curvearrowright P$ is *weakly compact* (although it is typically not compact). More precisely, they showed the existence of a net of vectors $\xi_k \in L^2(P)\bar{\otimes}L^2(P)$ which are almost invariant under the diagonal action of $\mathcal{G}$. In the second part of the proof, they combined the weak compactness property, with a malleable deformation of M analogous to the one from Example 3.5 and a spectral gap rigidity argument. Thus, they concluded that either a corner of P embeds into $L^\infty(X)$ or the von Neumann algebra generated by $\mathcal{G}$ is amenable.

If $P \subset M$ is a Cartan subalgebra, then the first condition must hold since M is not amenable. By Theorem 2.3, this forces that P is unitarily conjugate to $L^\infty(X)$.

Theorem 4.1 is restricted both by the class of groups and the family of actions it applies to. Ozawa [2012] showed that the weak compactness property established in the proof of Theorem 4.1 more generally holds for profinite actions of weakly amenable groups. Building on this and the weak amenability of hyperbolic groups Ozawa [2008], Chifan and Sinclair extended Theorem 4.1 to all non-elementary hyperbolic groups Γ Chifan and Sinclair [2013]. A conceptual novelty of their approach was the usage of quasi-cocyles (rather than cocycles Peterson and Sinclair [2012] and Sinclair [2011]) to build deformations.

Soon after, Popa and Vaes obtained uniqueness results of unprecedented generality in Popa and Vaes [2014a,b]. Generalizing Ozawa and Popa [2010a] and Chifan and Sinclair [2013], they showed that Theorem 4.1 holds for arbitrary actions of free groups and hyperbolic groups:

Theorem 4.2 (Popa and Vaes [2014a,b]). *Let $\Gamma \curvearrowright (X,\mu)$ be a free ergodic p.m.p. action of a countable group Γ. Assume either that Γ is weakly amenable and admits an unbounded cocycle into a non-amenable mixing orthogonal representation, or Γ is non-elementary hyperbolic. Then $L^\infty(X) \rtimes \Gamma$ has a unique Cartan subalgebra, up to unitary conjugacy.*

This result covers any weakly amenable group Γ with a positive first ℓ^2-Betti number, $\beta_1^{(2)}(\Gamma) > 0$. Indeed, the latter holds if and only if Γ is non-amenable and admits an unbounded cocycle into its left regular representation Bekka and Valette [1997] and Peterson and Thom [2011].

Theorem 4.2 led to the resolution of the group measure space analogue of the famous, still unsolved, *free group factor problem* which asks whether $L(\mathbb{F}_n)$ and $L(\mathbb{F}_m)$ are isomorphic or not, for $n \neq m$. More precisely, Popa and Vaes showed in Popa and Vaes [2014a] that if $2 \leq n, m \leq \infty$ and $n \neq m$, then for any free ergodic p.m.p. actions $\mathbb{F}_n \curvearrowright (X,\mu)$ and $\mathbb{F}_m \curvearrowright (Y,\nu)$ one has:

$$L^\infty(X) \rtimes \mathbb{F}_n \not\cong L^\infty(Y) \rtimes \mathbb{F}_m.$$

If these factors were isomorphic, then Theorem 4.2 would imply that the actions $\mathbb{F}_n \curvearrowright X$, $\mathbb{F}_m \curvearrowright Y$ are orbit equivalent. However, it was shown by Gaboriau [2000, 2002] that free groups of different ranks do not admit orbit equivalent free actions.

In the setting of Theorem 4.2, let $P \subset L^\infty(X) \rtimes \Gamma$ be a diffuse amenable subalgebra and denote by $\mathcal{G}$ its normalizer. The weak amenability of Γ is used to show that, roughly speaking, the action $\mathcal{G} \curvearrowright P$ is weakly compact relative to $L^\infty(X)$. When P is a Cartan subalgebra, in combination with the deformations obtained from cocycles or quasi-cocycles of Γ, one concludes that P is unitarily conjugate to $L^\infty(X)$.

Now, assume that Γ is a non-abelian free group or, more generally, a non-elementary hyperbolic group. In this context, Popa and Vaes proved a deep, important dichotomy for arbitrary diffuse amenable subalgebras P of arbitrary tracial crossed products $B \rtimes \Gamma$: either a corner of P embeds into B, or otherwise the von Neumann algebra generated by the normalizer of P is amenable relative to B Popa and Vaes [2014a,b]. This property of such groups Γ, called *relative strong solidity*, has since found a number of impressive applications. In particular, it was used in Ioana [2013] to provide the first class of non-weakly amenable groups satisfying Theorem 4.2:

Theorem 4.3. Ioana [ibid.] *Let $\Gamma = \Gamma_1 *_\Sigma \Gamma_2$ be an amalgamated free product group. Assume that $[\Gamma_1 : \Sigma] \geq 2$, $[\Gamma_2 : \Sigma] \geq 3$, and $\cap_{i=1}^n g_i \Sigma g_i^{-1} = \{e\}$, for some elements $g_1, ..., g_n \in \Gamma$. Let $\Gamma \curvearrowright (X, \mu)$ be a free ergodic p.m.p. action. Then $L^\infty(X) \rtimes \Gamma$ has a unique Cartan subalgebra, up to unitary conjugacy.*

Theorem 4.3 in particular applies to any free product group $\Gamma = \Gamma_1 * \Gamma_2$ with $|\Gamma_1| \geq 2$, $|\Gamma_2| \geq 3$. Since such groups have a positive first ℓ^2-Betti number, Theorem 4.2 and Theorem 4.3 both provide supporting evidence for the following general conjecture:

Problem I. *Let Γ be a countable group with $\beta_1^{(2)}(\Gamma) > 0$. Prove that $L^\infty(X) \rtimes \Gamma$ has a unique Cartan subalgebra, up to unitary conjugacy, for any free ergodic p.m.p. action $\Gamma \curvearrowright (X, \mu)$.*

A weaker version of Problem I asks to prove that $L^\infty(X) \rtimes \Gamma$ has a unique group measure space Cartan subalgebra. This has been confirmed by Chifan and Peterson [2013] under the additional assumption that Γ has a non-amenable subgroup with the relative property (T) (see also Vaes [2013]). A positive answer was also obtained in Ioana [2012b,a] when the action $\Gamma \curvearrowright (X, \mu)$ is rigid or profinite.

If Problem I or its weaker version were solved, then as ℓ^2-Betti numbers of groups are invariant under orbit equivalence Gaboriau [2002], it would follow that $\beta_1^{(2)}(\Gamma)$ is an isomorphism invariant of $L^\infty(X) \rtimes \Gamma$. This would provide a computable invariant for the class of group measure space II_1 factors.

4.2 Non-uniqueness of Cartan subalgebras. If Γ is any group as in Theorem 4.2 and Theorem 4.3, then $L^\infty(X) \rtimes \Gamma$ has a unique Cartan subalgebras, for any free ergodic p.m.p. action of Γ, while $L(\Gamma)$ and $L(\Gamma)\bar{\otimes}N$ do not have Cartan subalgebras, for any II_1 factor N. This provides several large families of II_1 factors with at most one Cartan subalgebra. However, in the non-uniqueness regime, little is known about the possible cardinality of the set of Cartan subalgebras of a II_1 factor.

Although the hyperfinite II_1 factor R admits uncountably many Cartan subalgebras up to unitary conjugacy Packer [1985], any two Cartan subalgebras are conjugated by an automorphism of R Connes, Feldman, and Weiss [1981]. The first class of examples of

II_1 factors admitting two Cartan subalgebras that are not conjugated by an automorphism was given by Connes and Jones [1982]. A second class of examples of such II_1 factors where the two Cartan subalgebras are explicit was found by Ozawa and Popa [2010b] (see also Popa and Vaes [2010b]). A class of II_1 factors M whose Cartan subalgebras cannot be concretely classified up to unitary conjugacy or up to conjugation by an automorphism of M was then introduced in Speelman and Vaes [2012]. Most recently, a family of II_1 factors whose all group measure space Cartan subalgebras can be described explicitly was constructed in Krogager and Vaes [2017]. In particular, the authors give examples of II_1 factors having exactly two group measure space Cartan subalgebras, up to unitary conjugacy, and a prescribed number of group measure space Cartan subalgebras, up to conjugacy with an automorphism. However, there are currently no known examples of II_1 having precisely $n \geq 2$ arbitrary Cartan subalgebras:

Problem II. *Given an integer $n \geq 2$, find II_1 factors M which have exactly n Cartan subalgebras, up to unitary conjugacy (or up to conjugacy with an automorphism).*

4.3 W*-superrigidity of Bernoulli actions.

Popa's strong rigidity theorem Popa [2006e] led to the natural conjecture that Bernoulli actions of icc property (T) groups are W*-superrigid. In this section we discuss the solution of this conjecture.

Theorem 4.4 (Ioana [2011a]). *Let Γ be an icc property (T) group and (X_0, μ_0) be a non-trivial probability space. Then the Bernoulli action $\Gamma \curvearrowright (X, \mu) = (X_0^\Gamma, \mu_0^\Gamma)$ is W*-superrigid.*

The proof of Theorem 4.4 relies on a general strategy for analyzing group measure space decompositions of II_1 factors Ioana [ibid.]. Let $M = L^\infty(X) \rtimes \Gamma$ be the II_1 factor arising from a "known" free ergodic p.m.p. action $\Gamma \curvearrowright (X, \mu)$. Assume that we can also decompose $M = L^\infty(Y) \rtimes \Lambda$, for some "mysterious" free ergodic p.m.p. action $\Lambda \curvearrowright (Y, \nu)$.

The new group measure space decomposition of M gives rise to an embedding $\Delta : M \to M \bar{\otimes} M$ defined by $\Delta(bu_h) = u_h \otimes u_h$, for all $b \in L^\infty(X)$ and $h \in \Lambda$. This embedding has been introduced in Popa and Vaes [2010b] were it was used to transfer rigidity properties through W*-equivalence. The strategy of Ioana [2011a] is to first prove a classification of all embeddings of M into $M \bar{\otimes} M$, and then apply it to Δ. This provides a relationship between the actions $\Gamma \curvearrowright X, \Lambda \curvearrowright Y$ which is typically stronger than the original W*-equivalence and which, ideally, can be exploited to show that the actions are orbit equivalent or even conjugate.

In the case $\Gamma \curvearrowright (X, \mu)$ is a Bernoulli action of a property (T) group, a classification of all possible embeddings $\theta : M \to M \bar{\otimes} M$ was obtained in Ioana [ibid.]. This classification is precise enough so that when combined with the above strategy it implies that

$L^\infty(X)$ is the unique group measure space Cartan subalgebra of M, up to unitary conjugacy. Since the action $\Gamma \curvearrowright (X,\mu)$ is OE-superrigid by a theorem of Popa [2007a], it follows that the action is also W*-superrigid.

The main novelty of Ioana [2011a] is a structural result for abelian subalgebras D on M that are normalized by a sequence of unitary elements $u_n \in L(\Gamma)$ converging weakly to 0. Under this assumption, it is shown that D and its relative commutant $D' \cap M$ can be essentially unitarily conjugated into either $L(\Gamma)$ or $L^\infty(X)$. An analogous dichotomy is proven in Ioana [ibid.] for abelian subalgebras $D \subset M \bar{\otimes} M$ which are normalized by "many" unitary elements from $L(\Gamma)\bar{\otimes}L(\Gamma)$. This is applied to study embeddings $\theta : M \to M\bar{\otimes}M$ as follows. Since $\theta(L(\Gamma))$ is a property (T) subalgebra of $M\bar{\otimes}M$, by adapting the arguments described in Section 3.2, we may assume that $\theta(L(\Gamma)) \subset L(\Gamma)\bar{\otimes}L(\Gamma)$. As a consequence, $D = \theta(L^\infty(X))$ is normalized by the group $\theta(\Gamma) \subset L(\Gamma)\bar{\otimes}L(\Gamma)$, and thus the dichotomy can be applied to D.

By Theorem 4.4, II_1 factors arising from Bernoulli actions of icc property (T) groups have a unique group measure space Cartan subalgebra. The following problem proposed by Popa asks to prove that the same holds for arbitrary non-amenable groups and general Cartan subalgebras:

Problem III. *Let $\Gamma \curvearrowright (X,\mu) = (X_0^\Gamma, \mu_0^\Gamma)$ be a Bernoulli action of a non-amenable group Γ. Then $L^\infty(X) \rtimes \Gamma$ has a unique Cartan subalgebra, up to unitary conjugacy.*

A positive answer to this problem would imply that if two Bernoulli actions of non-amenable groups are W*-equivalent, then they are orbit equivalent. By Popa [2007a,b], for Bernoulli actions of groups in a large class (containing all infinite property (T) groups), orbit equivalence implies conjugacy. However, this does not hold for arbitrary non-amenable groups. Indeed, if $n \geq 2$, then the Bernoulli actions $\mathbb{F}_n \curvearrowright (X_0^{\mathbb{F}_n}, \mu_0^{\mathbb{F}_n})$ of $\mathbb{F}_n$ are completely classified up to conjugacy by the entropy of base space (X_0, μ_0) Bowen [2010]; on the other hand, all Bernoulli actions of $\mathbb{F}_n$ are orbit equivalent Bowen [2011].

5 Orbit equivalence rigidity

Pioneering OE rigidity results were obtained by Zimmer for actions of higher rank semisimple Lie groups and their lattices by using his influential cocycle superrigidity theorem Zimmer [1980, 1984]. Deducing OE rigidity results from cocycle superrigidity theorems has since become a paradigm in the area. An illustration of this is Furman's proof that "generic" free ergodic p.m.p. actions of higher rank lattices, including the actions $\mathrm{SL}_n(\mathbb{Z}) \curvearrowright (\mathbb{T}^n, \mathbf{Leb})$ for $n \geq 3$, are virtually OE-superrigid. Since then, numerous striking OE superrigidity results have been discovered in Popa [2007a, 2008], Kida [2010], Ioana [2011b], Popa and Vaes [2011], Furman [2011], Kida [2011], Peterson and Sinclair

[2012], Tucker-Drob [n.d.], Ioana [2017], Chifan and Kida [2015], Drimbe [n.d.], and Gaboriau, Ioana, and Tucker-Drob [n.d.].

5.1 Cocycle superrigidity. Many of these results have been obtained by applying techniques and ideas from Popa's deformation/rigidity theory. In all of these cases, one proves that much more than being OE superrigid, the actions in question are cocycle superrigid Popa [2007a, 2008], Ioana [2011b], Popa and Vaes [2011], Furman [2011], Peterson and Sinclair [2012], Tucker-Drob [n.d.], Ioana [2017], Drimbe [n.d.], and Gaboriau, Ioana, and Tucker-Drob [n.d.]. These developments were trigerred by Popa's discovery of a striking new cocycle superrigidity phenomenon:

Theorem 5.1 (Popa [2007a, 2008]). *Assume that Γ is either an infinite property (T) group or the product $\Gamma_1 \times \Gamma_2$ of an infinite group and a non-amenable group. Let $\Gamma \curvearrowright (X, \mu) := (X_0^\Gamma, \mu_0^\Gamma)$ be the Bernoulli action, where (X_0, μ_0) is a non-trivial standard probability space. Let Λ be a countable group. Then any cocycle $w : \Gamma \times X_0^\Gamma \to \Lambda$ is cohomologous to a group homomorphism $\delta : \Gamma \to \Lambda$.*

Theorem 5.1 more generally applies to cocycles with values into $\mathcal{U}_{\text{fin}}$ groups Λ, i.e., isomorphic copies of closed subgroups of the unitary group of a separable II_1 factor.

Recall that if Λ is a Polish group, then a measurable map $w : \Gamma \times X \to \Lambda$ is called a *cocycle* if it satisfies the identity $w(g_1 g_2, x) = w(g_1, g_2 \cdot x) w(g_2, x)$ for all $g_1, g_2 \in \Gamma$ and almost every $x \in X$. Two cocycles $w_1, w_2 : \Gamma \times X \to \Lambda$ are called *cohomologous* if there exists a measurable map $\varphi : X \to \Lambda$ such that $w_2(g, x) = \varphi(g \cdot x) w_1(g, x) \varphi(x)^{-1}$, for all $g \in \Gamma$ and almost every $x \in X$. Any group homomorphism $\delta : \Gamma \to \Lambda$ gives rise to a constant cocycle, $w(g, x) := \delta(g)$.

Remark 5.2. If $\alpha : (X, \mu) \to (Y, \nu)$ is an OE between $\Gamma \curvearrowright (X, \mu)$ and a free p.m.p. action $\Lambda \curvearrowright (Y, \nu)$, then the map $w : \Gamma \times X \to \Lambda$ uniquely determined by the formula $\alpha(gx) = w(g, x) \cdot \alpha(x)$ is a cocycle, called the *Zimmer cocycle*. Assume that w is cohomologous to a homomorphism, i.e. $w(g, x) = \varphi(g \cdot x)\delta(g)\varphi(x)^{-1}$. Then the map $\tilde{\alpha} : X \to Y$ given by $\tilde{\alpha}(x) := \varphi(x)^{-1} \cdot \alpha(x)$ satisfies $\tilde{\alpha}(g \cdot x) = \delta(g) \cdot \tilde{\alpha}(x)$. This can be often used to conclude that the actions are conjugate, e.g., if Γ is icc and its action is free and weakly mixing Popa [2007a]. Consequently, the actions from Theorem 5.1 are OE-superrigid whenever Γ is icc.

The proof of Theorem 5.1 relies on the malleability of Bernoulli actions. Assume for simplicity that $(X_0, \mu_0) = ([0, 1], \textbf{Leb})$. Let $(\alpha_t)_{t \in \mathbb{R}}$ be a continuous family of automorphisms of $X \times X$ commuting with the diagonal action of Γ and satisfying $\alpha_0 = \text{id}$ and $\alpha_1(x, y) = (y, x)$, as in Example 3.4. Then the formula $w_t(g, (x, y)) = w(g, \alpha_t(x, y))$ defines a family $(w_t)_{t \in \mathbb{R}}$ of cocycles for the product action $\Gamma \curvearrowright X \times X$. Note that $w_0(g, (x, y)) = w(g, x)$, while $w_1(g, (x, y)) = w(g, y)$.

In Popa [2007a], Popa uses property (T) to deduce that w_t is cohomologous to w_0, for $t > 0$ small enough. The same conclusion is derived in Popa [2008] via a spectral gap rigidity argument (see Section 3.2). In both papers, this is then shown to imply that w is cohomologous to w_1. Finally, using the weak mixing property of Bernoulli actions, it is concluded that w is cohomologous to a homomorphism. For a presentation of the proof of Theorem 5.1 in the property (T) case, see also Furman [2007] and Vaes [2007].

In addition to property (T) and product groups, Theorem 5.1 has been shown to hold for groups Γ such that $L(\Gamma)$ is L^2-rigid in the sense of Peterson [2009] (see Peterson and Sinclair [2012]) and for inner amenable non-amenable groups Γ (see Tucker-Drob [n.d.]). The following problem due to Popa remains however open:

Problem IV. *Characterize the class of groups Γ whose Bernoulli actions are cocycle superrigid (for arbitrary countable or $\mathcal{U}_{fin}$ "target" groups Λ), in the sense of Theorem 5.1.*

This class is conjecturally characterised by the vanishing of the first L^2-Betti number. Indeed, all groups Γ known to belong to the class satisfy $\beta_1^{(2)}(\Gamma) = 0$. On the other hand, if $\beta_1^{(2)}(\Gamma) > 0$, then the Bernoulli actions of Γ are not cocycle superrigid with $\Lambda = \mathbb{T}$ as the target group Peterson and Sinclair [2012].

In Ioana [2011b], the author established a cocycle superrigidity theorem for translation actions $\Gamma \curvearrowright (G, \mathbf{m}_G)$ of property (T) groups Γ on their profinite completions G, see part (1) of Theorem 5.3. Note that these actions are in some sense the farthest from being weakly mixing or Bernoulli: the unitary representation $\Gamma \curvearrowright L^2(G)$ is a sum of finite dimensional representations, by the Peter-Weyl theorem.

Motivated by the analogy with Theorem 5.1, it was asked in Ioana [ibid.] whether a version of the cocycle superrigidity theorem obtained therein holds for product groups, such as $\Gamma = \mathbb{F}_2 \times \mathbb{F}_2$. The interest in this question was especially high at the time, since a positive answer combined with the work Ozawa and Popa [2010a] would have lead to the (then) first examples of virtually W*-superrigid actions. This question was recently settled in Gaboriau, Ioana, and Tucker-Drob [n.d.], see part (2) of Theorem 5.3.

Theorem 5.3. *Let Γ and Δ be countable dense subgroups of a compact profinite group G. Consider the left translation action $\Gamma \curvearrowright (G, \mathbf{m}_G)$ and the left-right translation action $\Gamma \times \Delta \curvearrowright (G, \mathbf{m}_G)$.*
Let Λ be a countable group. Let $w : \Gamma \times G \to \Lambda$ and $v : (\Gamma \times \Delta) \times G \to \Lambda$ be any cocycles.

1. Ioana [2011b] *Assume that Γ has property (T). Then we can find an open subgroup $G_0 < G$ such that the restriction of w to $(\Gamma \cap G_0) \times G_0$ is cohomologous to a homomorphism $\delta : \Gamma \cap G_0 \to \Lambda$.*

2. Gaboriau, Ioana, and Tucker-Drob [n.d.] *Assume that $\Gamma \curvearrowright (G, \mathbf{m}_G)$ has spectral gap, and Γ, Λ are finitely generated. Then we can find an open subgroup $G_0 < G$*

such that the restriction of v to $[(\Gamma \cap G_0) \times (\Delta \cap G_0)] \times G_0$ *is cohomologous to a homomorphism* $\delta : (\Gamma \cap G_0) \times (\Delta \cap G_0) \to \Lambda$.

Recall that the left-right translation action $\Gamma \times \Delta \curvearrowright (G, \mathbf{m}_G)$ is given by $(g,h) \cdot x = gxh^{-1}$. A p.m.p. action $\Gamma \curvearrowright (X, \mu)$ is said to have *spectral gap* if the representation $\Gamma \curvearrowright L^2(X) \ominus \mathbb{C}\mathbf{1}$ does not have almost invariant vectors. A well-known result of Selberg implies that the left translation action $\mathrm{SL}_2(\mathbb{Z}) \curvearrowright \mathrm{SL}_2(\mathbb{Z}_p)$ has spectral gap, for every prime p. This was recently generalized in Bourgain and Varjú [2012] to arbitrary non-amenable subgroups $\Gamma < \mathrm{SL}_2(\mathbb{Z})$: the left translation of Γ onto its closure $\bar{\Gamma} < \mathrm{SL}_2(\mathbb{Z}_p)$ has spectral gap.

Note that in Furman [2011], Furman provided an alternative approach to part (1) of Theorem 5.3 which applies to the wider class of compact actions.

The proof of Theorem 5.3 relies on the following criterion for untwisting cocycles $w : \Gamma \times G \to \Lambda$. First, endow the space of such cocycles with the "uniform" metric

$$d(w, w') := \sup_{g \in \Gamma} \mathbf{m}_G(\{x \in G \mid w(g,x) \neq w'(g,x)\}).$$

Second, since the left and right translation actions of G on itself commute, $w_t(g,x) = w(g,xt)$ defines a family of cocycles $(w_t)_{t \in G}$. It is then shown in Ioana [2011b] and Furman [2011] that if w verifies the uniformity condition $d(w_t, w) \to 0$, as $t \to 1_G$, then w untwists, in the sense of part (1) of Theorem 5.3.

5.2 OE rigidity for actions of non-rigid groups. In this section, we present a rigidity result for translation actions with spectral gap which describes precisely when two such actions are OE.

Moreover, the result also applies to the notion of Borel reducibility from descriptive set theory (see e.g. the survey Thomas [2006]). If $\mathcal{R}, \mathcal{S}$ are equivalence relations on standard Borel spaces X, Y, we say that $\mathcal{R}$ is *Borel reducible* to $\mathcal{S}$ whenever there exists a Borel map $\alpha : X \to Y$ such that $(x,y) \in \mathcal{R} \Leftrightarrow (\alpha(x), \alpha(y)) \in \mathcal{S}$. This encodes that the classification problem associated to $\mathcal{R}$ is no more complicated than the classification problem associated to $\mathcal{S}$.

Theorem 5.4. Ioana [2016] *Let Γ and Λ be countable dense subgroups of profinite groups G and H. Assume that the left translation action $\Gamma \curvearrowright (G, \mathbf{m}_G)$ has spectral gap.*

1. *$\Gamma \curvearrowright (G, \mathbf{m}_G)$ is OE to $\Lambda \curvearrowright (H, \mathbf{m}_H)$ iff there exist open subgroup $G_0 < G$, $H_0 < H$ and a topological isomorphism $\delta : G_0 \to H_0$ such that $\delta(\Gamma \cap G_0) = \Lambda \cap H_0$ and $[G : G_0] = [H : H_0]$.*

2. *$\mathcal{R}_{\Gamma \curvearrowright G}$ is Borel reducible to $\mathcal{R}_{\Lambda \curvearrowright H}$ iff we can find an open subgroup $G_0 < G$, a closed subgroup $H_0 < H$ and a topological isomorphism $\delta : G_0 \to H_0$ such that $\delta(\Gamma \cap G_0) = \Lambda \cap H_0$.*

The main novelty of this theorem lies in that there are no assumptions on the groups, but instead, all the assumptions are imposed on their actions. Thus, the theorem applies to many natural families of actions of $\mathrm{SL}_2(\mathbb{Z})$ and the free groups $\mathbb{F}_n$, leading to the following:

Corollary 5.5. Ioana [2016] *If S and T are distinct non-empty sets of primes, then the actions*

$$SL_2(\mathbb{Z}) \curvearrowright \prod_{p \in S} SL_2(\mathbb{Z}_p) \quad \textit{and} \quad SL_2(\mathbb{Z}) \curvearrowright \prod_{p \in T} SL_2(\mathbb{Z}_p)$$

are not orbit equivalent, and their equivalence relations are not Borel reducible one to another.

This result settles a conjecture of Thomas (see Thomas [2003, Conj. 5.7] and Thomas [2006, Conj. 2.14]). In particular, it provides a continuum of treeable countable Borel equivalence relations that are pairwise incomparable with respect to Borel reducibility. Note that the existence of uncountably many such equivalence relations has been first established by Hjorth [2012]. However, prior to Corollary 5.5, not a single example of a pair of treeable countable Borel equivalence relations such that neither is Borel reducible to the other was known.

Theorem 5.4 also leads to natural concrete uncountable families of pairwise non-OE free ergodic p.m.p. actions of the non-abelian free groups, $\mathbb{F}_n$ (the existence of such families was proved in Gaboriau and Popa [2005], while the first explicit families were found in Ioana [2011c]). In contrast, a non-rigidity result of Bowen shows that any two Bernoulli actions of $\mathbb{F}_n$ are OE Bowen [2011].

Theorem 5.4 admits a version which applies to connected (rather than profinite) compact groups. More precisely, let $\Gamma < G$ and $\Lambda < H$ be countable dense subgroups of compact connected Lie groups with trivial centers. Assuming that $\Gamma \curvearrowright (G, \mathbf{m}_G)$ has spectral gap, it is shown in Ioana [2016] that the actions $\Gamma \curvearrowright G$ and $\Lambda \curvearrowright H$ are orbit equivalent iff they are conjugate. Subsequently, this has been generalized in Ioana [2017] to the case when G and H are arbitrary, not necessarily compact, connected Lie groups with trivial centers. The only difference is that, in the locally compact setting, the spectral gap assumption no longer makes sense and has to be replaced with the assumption that the action $\Gamma \curvearrowright (G, \mathbf{m}_G)$ is *strongly ergodic*. After choosing a Borel probability measure μ on G equivalent to $\mathbf{m}_G$, the latter requires that any sequence of measurable sets $A_n \subset G$ satisfying $\mu(gA_n \Delta A_n) \to 0$, for all $g \in \Gamma$, must be asymptotically trivial, in the sense that $\mu(A_n)\mu(A_n^c) \to 0$.

If G is a connected compact simple Lie group (e.g., if $G = \mathrm{SO}(n)$, for $n \geq 3$), the translation action $\Gamma \curvearrowright (G, \mathbf{m}_G)$ has spectral gap and thus is strongly ergodic, whenever Γ is generated by matrices with algebraic entries. This result is due to Bourgain and Gamburd for $G = \mathrm{SO}(3)$ Bourgain and Gamburd [2008], and to Benoist and de Saxcé in general Benoist and de Saxcé [2016]. In joint work with Boutonnet and Salehi-Golsefidy, we have shown that the same holds if G is an arbitrary connected simple Lie group Boutonnet, Ioana, and Golsefidy [2017]. In particular, the translation action $\Gamma \curvearrowright (G, \mathbf{m}_G)$ is strongly ergodic, whenever $\Gamma < G := \mathrm{SL}_n(\mathbb{R})$ is a dense subgroup whose elements are matrices with algebraic entries.

6 Structure and rigidity for group von Neumann algebras

In this section, we survey some recent developments in the classification of von Neumann algebras arising from countable groups. Our presentation follows three directions.

6.1 Structural results. We first present results which provide classes of group II_1 factors $L(\Gamma)$ with various indecomposability properties, such as primeness and the lack of Cartan subalgebras.

In Popa [1983], Popa proved that II_1 factors arising from free groups with uncountably many generators are prime and do not have Cartan subalgebras. The first examples of separable such II_1 factors were obtained in the mid 1990s as an application of free probability theory. Thus, Voiculescu showed that the free group factors $L(\mathbb{F}_n)$, with $n \geq 2$, do not admit Cartan subalgebras Voiculescu [1996]. Subsequently, Ge used the techniques from Voiculescu [ibid.] to prove that the free group factors are also prime Ge [1998].

These results have been since generalized and strengthened in several ways. Using subtle C*-algebras techniques, Ozawa remarkably proved that II_1 factors associated to icc hyperbolic groups Γ are *solid*: the relative commutant $A' \cap L(\Gamma)$ of any diffuse von Neumann subalgebra $A \subset L(\Gamma)$ is amenable Ozawa [2004]. In particular, $L(\Gamma)$ and all of its non-amenable subfactors are prime. By developing a novel technique based on closable derivations, Peterson showed that II_1 factors arising from icc groups with positive first ℓ^2-Betti number are prime Peterson [2009]. A new proof of solidity of $L(\mathbb{F}_n)$ was found by Popa [2007c] (see Example 3.5), while II_1 factors coming from icc groups Γ admitting a proper cocycle into $\ell^2(\Gamma)$ were shown to be solid in Peterson [2009]. For additional examples of prime and solid II_1 factors, see Ozawa [2006], Popa [2008], Chifan and Ioana [2010], Chifan and Houdayer [2010], Vaes [2013], Boutonnet [2013], Houdayer and Vaes [2013], Dabrowski and Ioana [2016], Chifan, Kida, and Pant [2016], and D. J. Hoff [2016].

In Ozawa and Popa [2010a], Ozawa and Popa discovered that the free group factors enjoy a striking structural property, called *strong solidity*, which strengthens both primeness and absence of Cartan subalgebras: the normalizer of any diffuse amenable subalgebra $A \subset L(\mathbb{F}_n)$ is amenable. Generalizing this result, Chifan and Sinclair showed that in fact the von Neumann algebra of any icc hyperbolic group is strongly solid Chifan and Sinclair [2013]. For further examples of strongly solid factors, see Ozawa and Popa [2010b], Houdayer [2010], Houdayer and Shlyakhtenko [2011], and Sinclair [2011].

The groups Γ for which $L(\Gamma)$ was shown to be prime have certain properties, such as hyperbolicity or the existence of unbounded quasi-cocycles, relating them to lattices in rank one Lie groups. On the other hand, the primeness question for the higher rank lattices such as $\mathrm{PSL}_m(\mathbb{Z})$, $m \geq 2$, remains a major open problem. Moreover, little is known about the structure of II_1 factors associated to lattices in higher rank semisimple Lie groups. We formulate a general question in this direction:

Problem V. *Let Γ be an icc irreducible lattice in a direct product $G = G_1 \times ... \times G_n$ of connected non-compact simple real Lie groups with finite center. Prove that $L(\Gamma)$ is prime and does not have a Cartan subalgebra. Moreover, if $G_1, ..., G_n$ are of rank one, prove that $L(\Gamma)$ is strongly solid.*

A subgroup $\Gamma < G$ is called a *lattice* if it is discrete and has finite co-volume, $\mathbf{m}_G(G/\Gamma) < \infty$. A lattice $\Gamma < G$ is called *irreducible* if its projection onto $\times_{j \neq i} G_j$ is dense, for all $1 \leq i \leq n$.

When $n = 1$, Problem V has been resolved in Ozawa [2004] and Chifan and Sinclair [2013] (see also Ozawa and Popa [2010b] and Sinclair [2011]) whenever $G = G_1$ has rank one, but is open if G has higher rank.

In the rest of the section, we record recent progress in the case $n \geq 2$. Note that the irreducibility assumption on Γ is needed in order to exclude product lattices $\Gamma = \Gamma_1 \times ... \times \Gamma_n$, whose II_1 factors are obviously non-prime. However, it seems plausible that $L(\Gamma)$ does not have a Cartan subalgebra, for an arbitrary lattice Γ of G, irreducible or not. Indeed, this was established by Popa and Vaes [2014b] if $G_1, \ldots, G_n$ are all of rank one. Complementing their work, the first examples of prime II_1 factors arising from lattices in higher rank semisimple Lie groups were only recently obtained in Drimbe, D. Hoff, and Ioana [n.d.]:

Theorem 6.1. Drimbe, D. Hoff, and Ioana [ibid.] *If Γ is an icc irreducible lattice in a product $G = G_1 \times ... \times G_n$ of $n \geq 2$ connected non-compact rank one simple real Lie groups with finite center, then $L(\Gamma)$ is prime.*

Theorem 6.1 implies that the II_1 factor associated to $\mathrm{PSL}_2(\mathbb{Z}[\sqrt{2}])$, which can be realized as an irreducible lattice in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$, is prime. Theorem 6.1 also applies when each G_i is a rank one non-compact simple algebraic group over a local field. This implies that the II_1 factor arising from $\mathrm{PSL}_2(\mathbb{Z}[\frac{1}{p}])$, which is an irreducible lattice in $\mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{Q}_p)$, is prime for any prime p. It remains however an open problem whether $\mathrm{PSL}_2(\mathbb{Z}[\sqrt{2}])$, $\mathrm{PSL}_2(\mathbb{Z}[\frac{1}{p}])$, or any other group covered by Theorem 6.1 gives rise to a strongly solid II_1 factor.

To overview the proof of Theorem 6.1, put $M = L(\Gamma)$ and assume the existence of a decomposition $M = M_1 \bar{\otimes} M_2$ into a tensor product of factors. Let $\Delta_\Gamma : M \to M \bar{\otimes} M$ be the embedding given by

$$\Delta_\Gamma(u_g) = u_g \otimes u_g, \quad \text{for all } g \in \Gamma.$$

Since G_i has rank one, it admits a non-elementary hyperbolic lattice $\Lambda_i < G_i$. As a consequence, Γ is measure equivalent (in the sense of Gromov) to a product $\Lambda = \Lambda_1 \times \ldots \times \Lambda_n$ of hyperbolic groups. In combination with the relative strong solidity of hyperbolic groups Popa and Vaes [2014b], this allows one to conclude that a corner of $\Delta_\Gamma(M_j)$ embeds into $M \bar{\otimes} M_j$, for all $j \in \{1, 2\}$. By making crucial use of an ultrapower technique from Ioana [2012b], we derive that Γ admits commuting non-amenable subgroups. Indeed, by Ioana [ibid.] the existence of commuting non-amenable subgroups can be deduced whenever we can find II_1 subfactors $A, B \subset M$ such that a corner of $\Delta_\Gamma(A)$ embeds into $M \bar{\otimes}(B' \cap M)$. While more work is needed in general, this easily gives a contradiction for $\Gamma = \mathrm{PSL}_2(\mathbb{Z}[\sqrt{2}])$ or $\mathrm{PSL}_2(\mathbb{Z}[\frac{1}{p}])$.

6.2 Algebraic rigidity.

The primeness results discussed in Section 6.1 provide classes of groups for which the absence of a direct product decomposition is inherited by their von Neumann algebras. In a complementary direction, it was recently shown in Chifan, de Santiago, and Sinclair [2016] that for a wide family of product groups Γ, the von Neumann algebra $L(\Gamma)$ completely remembers the product structure of Γ.

Theorem 6.2. Chifan, de Santiago, and Sinclair [ibid.] *Let $\Gamma = \Gamma_1 \times \ldots \times \Gamma_n$, where $\Gamma_1, \ldots, \Gamma_n$ are $n \geq 2$ icc hyperbolic groups. Then any countable group Λ such that $L(\Gamma) \cong L(\Lambda)$ is a product of n icc groups, $\Lambda = \Lambda_1 \times \ldots \times \Lambda_n$.*

By applying the unique prime factorization theorem from Ozawa and Popa [2004], Theorem 6.2 can be strengthened to moreover show that $L(\Lambda_i)$ is stably isomorphic to $L(\Gamma_i)$, for all i.

To outline the proof of Theorem 6.2, assume that $n = 2$ and put $M = L(\Gamma)$ and $M_i = L(\Gamma_i)$. Consider the embedding $\Delta_\Lambda : M \to M \bar{\otimes} M$ associated to an arbitrary group von Neumann algebra decomposition $M = L(\Lambda)$. First, solidity properties of hyperbolic groups Ozawa [2004] and Brown and Ozawa [2008] are used to derive the existence of $i, j \in \{1, 2\}$ such that a corner of $\Delta_\Lambda(M_i)$ embeds into $M \bar{\otimes} M_j$. The ultrapower technique of Ioana [2012b] then allows to transfer the presence of non-amenable commuting subgroups from Γ to the mysterious group Λ. Using a series of ingenious combinatorial lemmas and strong solidity results from Chifan and Sinclair [2013] and Popa and Vaes [2014b], the authors conclude that Λ is indeed a product of icc groups.

Motivated by Theorem 6.2, it seems natural to investigate what other constructions in group theory can be detected at the level of the associated von Neumann algebra. Very recently, progress on this problem has been made in Chifan and Ioana [2018] by providing a class of amalgamated free product groups $\Gamma = \Gamma_1 *_\Sigma \Gamma_2$ whose von Neumann algebra $L(\Gamma)$ entirely recognizes the amalgam structure of Γ. This class includes all groups of the form $(A * B \times A * B) *_{A \times A} (A * B \times A * B)$, where A is an icc amenable group and B is a non-trivial hyperbolic group. For such groups Γ it is shown in Chifan and Ioana [ibid.] that any group Λ satisfying $L(\Gamma) \cong L(\Lambda)$ admits a amalgamated free product decomposition $\Lambda = \Lambda_1 *_\Omega \Lambda_2$ such that the inclusions $L(\Sigma) \subset L(\Gamma_i)$ and $L(\Omega) \subset L(\Lambda_i)$ are isomorphic, for all $i \in \{1, 2\}$.

6.3 W*-superrigidity. The rigidity results presented in Section 6.1 and Section 6.2 give instances when various algebraic aspects of groups can be recovered from their von Neumann algebras. We now discuss the most extreme type of rigidity for group von Neumann algebras. This occurs when the von Neumann algebra $L(\Gamma)$ completely remembers the group Γ. Thus, we say that a countable group Γ is *W*-superrigid* if any group Λ satisfying $L(\Gamma) \cong L(\Lambda)$ must be isomorphic to Γ.

The first class of W*-superrigid groups was discovered in our joint work with Ioana, Popa, and Vaes [2013]:

Theorem 6.3. Ioana, Popa, and Vaes [ibid.] *Let G_0 be any non-amenable group and S be any infinite amenable group. Define the wreath product group $G = G_0^{(S)} \rtimes S$, and consider the left multiplication action of G on $I = G/S$. Then the generalized wreath product group $\Gamma = (\mathbb{Z}/2\mathbb{Z})^{(I)} \rtimes G$ is W*-superrigid.*

The conclusion of Theorem 6.3 does not hold for *plain* wreath product groups $\Gamma = (\mathbb{Z}/2\mathbb{Z})^{(G)} \rtimes G$. In fact, for any non-trivial torsion free group G, there is a torsion free group Λ with $L(\Gamma) \cong L(\Lambda)$, while Γ and Λ are not isomorphic. Nevertheless, for a large family of groups G, including all icc property (T) groups, we are able to conclude that any group Λ satisfying $L(\Gamma) \cong L(\Lambda)$ decomposes as a semi-direct product $\Gamma = \Sigma \rtimes G$, for some abelian group Σ. This is in particular recovers a seminal result of Popa [2006e] showing that if $\Lambda = (\mathbb{Z}/2\mathbb{Z})^{(H)} \rtimes H$ is also a plain wreath product, then $L(\Gamma) \cong L(\Lambda)$ entails $G \cong H$.

To describe the strategy of the proof of Theorem 6.3, let $M = L(\Gamma)$ and assume that $M = L(\Lambda)$, for a countable group Λ. Then we have two embeddings $\Delta_\Gamma, \Delta_\Lambda : M \to M \bar{\otimes} M$. Notice that M can also be realized as the group measure space factor of the generalized Bernoulli action $G \curvearrowright \{0, 1\}^I$. By extending the methods of Ioana [2011a] from plain to generalized Bernoulli actions, we give a classification of all possible embeddings $\Delta : M \to M \bar{\otimes} M$. This enables us to deduce the existence of a unitary element $u \in M \bar{\otimes} M$ such that $\Delta_\Lambda(x) = u \Delta_\Gamma(x) u^*$, for all $x \in M$. A principal novelty of Ioana, Popa, and Vaes [2013] is being able to conclude that the groups Γ and Λ are isomorphic (in fact, unitarily conjugate modulo scalars inside M) from the mere existence of u.

Following Ioana, Popa, and Vaes [ibid.], several other classes of W*-superrigid groups were found in Berbec and Vaes [2014] and Chifan and Ioana [2018]. Thus, it was shown in Berbec and Vaes [2014] that the left-right wreath product $\Gamma = (\mathbb{Z}/2\mathbb{Z})^{(\Gamma)} \rtimes (\Gamma \times \Gamma)$ is W*-superrigid, for any icc hyperbolic group Γ. Very recently, a class of W*-superrigid amalgamated free product groups was found in Chifan and Ioana [2018]. These groups are also $\mathrm{C}^*_{\mathrm{red}}$-superrigid since, unlike the examples from Ioana, Popa, and Vaes [2013] and Berbec and Vaes [2014], they do not admit non-trivial normal amenable subgroups.

While the above results provide several large families of W*-superrigid groups, the superrigidity question remains open for many natural classes of groups including the lattices $\mathrm{PSL}_m(\mathbb{Z})$, $m \geq 3$. In fact, a well-known *rigidity conjecture* of Connes [1982] asks if $L(\Gamma) \cong L(\Lambda)$ for icc property (T) groups Γ and Λ implies $\Gamma \cong \Lambda$. Since property (T) is a von Neumann algebra invariant Connes and Jones [1985], Connes' conjecture is equivalent to asking whether icc property (T) groups Γ are W*-superrigid.

References

Mohammed E. B. Bekka and Alain Valette (1997). "Group cohomology, harmonic functions and the first L^2-Betti number". *Potential Anal.* 6.4, pp. 313–326. MR: 1452785 (cit. on p. 1673).

Yves Benoist and Nicolas de Saxcé (2016). "A spectral gap theorem in simple Lie groups". *Invent. Math.* 205.2, pp. 337–361. MR: 3529116 (cit. on p. 1681).

Mihaita Berbec and Stefaan Vaes (2014). "W*-superrigidity for group von Neumann algebras of left-right wreath products". *Proc. Lond. Math. Soc. (3)* 108.5, pp. 1116–1152. MR: 3214676 (cit. on pp. 1663, 1685).

Jean Bourgain and Alex Gamburd (2008). "On the spectral gap for finitely-generated subgroups of SU(2)". *Invent. Math.* 171.1, pp. 83–121. MR: 2358056 (cit. on p. 1681).

Jean Bourgain and Péter P. Varjú (2012). "Expansion in $SL_d(\mathbf{Z}/q\mathbf{Z})$, q arbitrary". *Invent. Math.* 188.1, pp. 151–173. MR: 2897695 (cit. on p. 1679).

Rémi Boutonnet (2013). "W*-superrigidity of mixing Gaussian actions of rigid groups". *Adv. Math.* 244, pp. 69–90. MR: 3077866 (cit. on pp. 1662, 1681).

Rémi Boutonnet, Ionuţ Chifan, and Adrian Ioana (2017). "II_1 factors with nonisomorphic ultrapowers". *Duke Math. J.* 166.11, pp. 2023–2051. MR: 3694564 (cit. on p. 1663).

Rémi Boutonnet, Cyril Houdayer, and Sven Raum (2014). "Amalgamated free product type III factors with at most one Cartan subalgebra". *Compos. Math.* 150.1, pp. 143–174. MR: 3164361 (cit. on p. 1663).

Rémi Boutonnet, Cyril Houdayer, and Stefaan Vaes (n.d.). "Strong solidity of free Araki-Woods factors". Preprint, to appear in Amer. J. Math. arXiv: 1512.04820 (cit. on p. 1663).

Rémi Boutonnet, Adrian Ioana, and Alireza Salehi Golsefidy (2017). "Local spectral gap in simple Lie groups and applications". *Invent. Math.* 208.3, pp. 715–802. MR: 3648974 (cit. on pp. 1662, 1681).

Lewis Phylip Bowen (2010). "A measure-conjugacy invariant for free group actions". *Ann. of Math. (2)* 171.2, pp. 1387–1400. MR: 2630067 (cit. on p. 1676).

– (2011). "Orbit equivalence, coinduced actions and free products". *Groups Geom. Dyn.* 5.1, pp. 1–15. MR: 2763777 (cit. on pp. 1676, 1680).

Nathanial P. Brown and Narutaka Ozawa (2008). *C^*-algebras and finite-dimensional approximations*. Vol. 88. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xvi+509. MR: 2391387 (cit. on p. 1683).

Ionuţ Chifan and Cyril Houdayer (2010). "Bass-Serre rigidity results in von Neumann algebras". *Duke Math. J.* 153.1, pp. 23–54. MR: 2641939 (cit. on pp. 1662, 1681).

Ionuţ Chifan and Adrian Ioana (2010). "Ergodic subequivalence relations induced by a Bernoulli action". *Geom. Funct. Anal.* 20.1, pp. 53–67. MR: 2647134 (cit. on pp. 1658, 1681).

– (2018). "Amalgamated Free Product Rigidity for Group von Neumann Algebras". *Adv. Math.* 329, pp. 819–850. arXiv: 1705.07350 (cit. on pp. 1663, 1684, 1685).

Ionuţ Chifan, Adrian Ioana, and Yoshikata Kida (2015). "W^*-superrigidity for arbitrary actions of central quotients of braid groups". *Math. Ann.* 361.3-4, pp. 563–582. MR: 3319541 (cit. on p. 1662).

Ionuţ Chifan and Yoshikata Kida (2015). "OE and W^* superrigidity results for actions by surface braid groups". *Proc. Lond. Math. Soc. (3)* 111.6, pp. 1431–1470. MR: 3447799 (cit. on pp. 1676, 1677).

Ionuţ Chifan, Yoshikata Kida, and Sujan Pant (2016). "Primeness results for von Neumann algebras associated with surface braid groups". *Int. Math. Res. Not. IMRN* 16, pp. 4807–4848. MR: 3556425 (cit. on p. 1681).

Ionuţ Chifan and Jesse Peterson (2013). "Some unique group-measure space decomposition results". *Duke Math. J.* 162.11, pp. 1923–1966. MR: 3090781 (cit. on pp. 1661, 1674).

Ionuţ Chifan, Rolando de Santiago, and Thomas Sinclair (2016). "W^*-rigidity for the von Neumann algebras of products of hyperbolic groups". *Geom. Funct. Anal.* 26.1, pp. 136–159. MR: 3494487 (cit. on pp. 1663, 1683).

Ionuţ Chifan and Thomas Sinclair (2013). "On the structural theory of II_1 factors of negatively curved groups". *Ann. Sci. Éc. Norm. Supér. (4)* 46.1, 1–33 (2013). MR: 3087388 (cit. on pp. 1661, 1673, 1682, 1683).

Alain Connes (1975). "Sur la classification des facteurs de type II". *C. R. Acad. Sci. Paris Sér. A-B* 281.1, Aii, A13–A15. MR: 0377534 (cit. on p. 1659).

– (1976). "Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$". *Ann. of Math. (2)* 104.1, pp. 73–115. MR: 0454659 (cit. on pp. 1658, 1671).

– (1980). "A factor of type II_1 with countable fundamental group". *J. Operator Theory* 4.1, pp. 151–153. MR: 587372 (cit. on p. 1659).

– (1982). "Classification des facteurs". In: *Operator algebras and applications, Part 2 (Kingston, Ont., 1980)*. Vol. 38. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, R.I., pp. 43–109. MR: 679497 (cit. on pp. 1667, 1685).

Alain Connes, Jacob Feldman, and B. Weiss (1981). "An amenable equivalence relation is generated by a single transformation". *Ergodic Theory Dynamical Systems* 1.4, 431–450 (1982). MR: 662736 (cit. on pp. 1658, 1674).

Alain Connes and Vaughan F. R. Jones (1982). "A II_1 factor with two nonconjugate Cartan subalgebras". *Bull. Amer. Math. Soc. (N.S.)* 6.2, pp. 211–212. MR: 640947 (cit. on p. 1675).

– (1985). "Property T for von Neumann algebras". *Bull. London Math. Soc.* 17.1, pp. 57–62. MR: 766450 (cit. on pp. 1659, 1667, 1685).

Michael Cowling and Uffe Haagerup (1989). "Completely bounded multipliers of the Fourier algebra of a simple Lie group of real rank one". *Invent. Math.* 96.3, pp. 507–549. MR: 996553 (cit. on pp. 1659, 1672).

Darren Creutz and Jesse Peterson (n.d.). "Character rigidity for lattices and commensurators". Preprint. arXiv: 1311.4513 (cit. on p. 1663).

Yoann Dabrowski and Adrian Ioana (2016). "Unbounded derivations, free dilations, and indecomposability results for II_1 factors". *Trans. Amer. Math. Soc.* 368.7, pp. 4525–4560. MR: 3456153 (cit. on p. 1681).

Daniel Drimbe (n.d.). "Cocycle superrigidity for coinduced actions". Preprint. To appear Erg. Th. Dyn. Sys. arXiv: 1512.00093 (cit. on pp. 1676, 1677).

Daniel Drimbe, Daniel Hoff, and Adrian Ioana (n.d.). "Prime II_1 factors arising from irreducible lattices in products of rank one simple Lie groups". To appear in J. Reine Angew. Math. arXiv: 1611.02209 (cit. on pp. 1662, 1682).

H. A. Dye (1959). "On groups of measure preserving transformation. I". *Amer. J. Math.* 81, pp. 119–159. MR: 0131516 (cit. on p. 1658).

Inessa Epstein (n.d.). "Orbit inequivalent actions of non-amenable groups". Preprint. arXiv: 0707.4215v2 (cit. on p. 1658).

Ilijas Farah, Bradd Hart, and David Sherman (2013). "Model theory of operator algebras I: stability". *Bull. Lond. Math. Soc.* 45.4, pp. 825–838. MR: 3081550 (cit. on p. 1663).

– (2014a). "Model theory of operator algebras II: model theory". *Israel J. Math.* 201.1, pp. 477–505. MR: 3265292 (cit. on p. 1663).

Ilijas Farah, Bradd Hart, and David Sherman (2014b). "Model theory of operator algebras III: elementary equivalence and II_1 factors". *Bull. Lond. Math. Soc.* 46.3, pp. 609–628. MR: 3210717 (cit. on p. 1663).

Jacob Feldman and Calvin C. Moore (1977). "Ergodic equivalence relations, cohomology, and von Neumann algebras. II". *Trans. Amer. Math. Soc.* 234.2, pp. 325–359. MR: 0578730 (cit. on p. 1666).

Alex Furman (1999). "Orbit equivalence rigidity". *Ann. of Math. (2)* 150.3, pp. 1083–1108. MR: 1740985 (cit. on pp. 1662, 1666).

– (2007). "On Popa's cocycle superrigidity theorem". *Int. Math. Res. Not. IMRN* 19, Art. ID rnm073, 46. MR: 2359545 (cit. on p. 1678).

– (2011). "A survey of measured group theory". In: *Geometry, rigidity, and group actions*. Chicago Lectures in Math. Univ. Chicago Press, Chicago, IL, pp. 296–374. MR: 2807836 (cit. on pp. 1658, 1676, 1677, 1679).

Damien Gaboriau (2000). "Coût des relations d'équivalence et des groupes". *Invent. Math.* 139.1, pp. 41–98. MR: 1728876 (cit. on pp. 1659, 1673).

– (2002). "Invariants l^2 de relations d'équivalence et de groupes". *Publ. Math. Inst. Hautes Études Sci.* 95, pp. 93–150. MR: 1953191 (cit. on pp. 1659, 1673, 1674).

– (2010). "Orbit equivalence and measured group theory". In: *Proceedings of the International Congress of Mathematicians. Volume III*. Hindustan Book Agency, New Delhi, pp. 1501–1527. MR: 2827853 (cit. on p. 1658).

– (2011). "Free product actions with relative property (T) and trivial outer automorphism groups". *J. Funct. Anal.* 260.2, pp. 414–427. MR: 2737406.

Damien Gaboriau, Adrian Ioana, and Robin Tucker-Drob (n.d.). "Cocycle superrigidity for translation actions of product groups". Preprint. arXiv: 1603.07616v1 (cit. on pp. 1658, 1662, 1676–1678).

Damien Gaboriau and Russell Lyons (2009). "A measurable-group-theoretic solution to von Neumann's problem". *Invent. Math.* 177.3, pp. 533–540. MR: 2534099.

Damien Gaboriau and Sorin Popa (2005). "An uncountable family of nonorbit equivalent actions of $\mathbb{F}_n$". *J. Amer. Math. Soc.* 18.3, pp. 547–559. MR: 2138136 (cit. on pp. 1658, 1667, 1680).

Liming Ge (1998). "Applications of free entropy to finite von Neumann algebras. II". *Ann. of Math. (2)* 147.1, pp. 143–157. MR: 1609522 (cit. on pp. 1662, 1681).

Uffe Haagerup (1978/79). "An example of a nonnuclear C^*-algebra, which has the metric approximation property". *Invent. Math.* 50.3, pp. 279–293. MR: 520930 (cit. on pp. 1669, 1672).

Greg Hjorth (2012). "Treeable equivalence relations". *J. Math. Log.* 12.1, pp. 1250003, 21. MR: 2950193 (cit. on p. 1680).

Daniel J. Hoff (2016). "Von Neumann algebras of equivalence relations with nontrivial one-cohomology". *J. Funct. Anal.* 270.4, pp. 1501–1536. MR: 3447718 (cit. on p. 1681).

Cyril Houdayer (2010). "Strongly solid group factors which are not interpolated free group factors". *Math. Ann.* 346.4, pp. 969–989. MR: 2587099 (cit. on p. 1682).

Cyril Houdayer and Yusuke Isono (2017). "Unique prime factorization and bicentralizer problem for a class of type III factors". *Adv. Math.* 305, pp. 402–455. MR: 3570140 (cit. on p. 1663).

Cyril Houdayer, Sorin Popa, and Stefaan Vaes (2013). "A class of groups for which every action is W*-superrigid". *Groups Geom. Dyn.* 7.3, pp. 577–590. MR: 3095710.

Cyril Houdayer and Dimitri Shlyakhtenko (2011). "Strongly solid II_1 factors with an exotic MASA". *Int. Math. Res. Not. IMRN* 6, pp. 1352–1380. MR: 2806507 (cit. on p. 1682).

Cyril Houdayer, Dimitri Shlyakhtenko, and Stefaan Vaes (n.d.). "Classification of a family of non almost periodic free Araki-Woods factors". Preprint. To appear in J. Eur. Math. Soc. (JEMS). arXiv: 1605.06057 (cit. on p. 1663).

Cyril Houdayer and Stefaan Vaes (2013). "Type III factors with unique Cartan decomposition". *J. Math. Pures Appl. (9)* 100.4, pp. 564–590. MR: 3102166 (cit. on pp. 1663, 1681).

Adrian Ioana (2009). "Non-orbit equivalent actions of $\mathbb{F}_n$". *Ann. Sci. Éc. Norm. Supér. (4)* 42.4, pp. 675–696. MR: 2568879 (cit. on p. 1667).

– (2011a). "W^*-superrigidity for Bernoulli actions of property (T) groups". *J. Amer. Math. Soc.* 24.4, pp. 1175–1226. MR: 2813341 (cit. on pp. 1661, 1667, 1675, 1676, 1684).

– (2011b). "Cocycle superrigidity for profinite actions of property (T) groups". *Duke Math. J.* 157.2, pp. 337–367. MR: 2783933 (cit. on pp. 1658, 1662, 1667, 1676–1679).

– (2011c). "Orbit inequivalent actions for groups containing a copy of $\mathbb{F}_2$". *Invent. Math.* 185.1, pp. 55–73. MR: 2810796 (cit. on pp. 1658, 1667, 1680).

– (2012a). "Compact actions and uniqueness of the group measure space decomposition of II_1 factors". *J. Funct. Anal.* 262.10, pp. 4525–4533. MR: 2900475 (cit. on p. 1674).

– (2012b). "Uniqueness of the group measure space decomposition for Popa's $\mathcal{HT}$ factors". *Geom. Funct. Anal.* 22.3, pp. 699–732. MR: 2972606 (cit. on pp. 1661, 1674, 1683).

– (2013). "Classification and rigidity for von Neumann algebras". In: *European Congress of Mathematics*. Eur. Math. Soc., Zürich, pp. 601–625. MR: 3469148 (cit. on pp. 1658, 1660, 1662, 1674).

– (2015). "Cartan subalgebras of amalgamated free product II_1 factors". *Ann. Sci. Éc. Norm. Supér. (4)* 48.1. With an appendix by Ioana and Stefaan Vaes, pp. 71–130. MR: 3335839 (cit. on p. 1669).

– (2016). "Orbit equivalence and Borel reducibility rigidity for profinite actions with spectral gap". *J. Eur. Math. Soc. (JEMS)* 18.12, pp. 2733–2784. MR: 3576536 (cit. on pp. 1658, 1662, 1679, 1680).

Adrian Ioana (2017). "Strong ergodicity, property (T), and orbit equivalence rigidity for translation actions". *J. Reine Angew. Math.* 733, pp. 203–250. arXiv: 1406.6628. MR: 3731329 (cit. on pp. 1676, 1677, 1680).

Adrian Ioana, Jesse Peterson, and Sorin Popa (2008). "Amalgamated free products of weakly rigid factors and calculation of their symmetry groups". *Acta Math.* 200.1, pp. 85–153. MR: 2386109 (cit. on p. 1659).

Adrian Ioana, Sorin Popa, and Stefaan Vaes (2013). "A class of superrigid group von Neumann algebras". *Ann. of Math. (2)* 178.1, pp. 231–286. MR: 3043581 (cit. on pp. 1663, 1684, 1685).

D. A. Každan (1967). "On the connection of the dual space of a group with the structure of its closed subgroups". *Funkcional. Anal. i Priložen.* 1, pp. 71–74. MR: 0209390.

Yoshikata Kida (2010). "Measure equivalence rigidity of the mapping class group". *Ann. of Math. (2)* 171.3, pp. 1851–1901. MR: 2680399 (cit. on pp. 1662, 1676).

– (2011). "Rigidity of amalgamated free products in measure equivalence". *J. Topol.* 4.3, pp. 687–735. MR: 2832574 (cit. on pp. 1661, 1676).

Anna Sofie Krogager and Stefaan Vaes (2017). "A class of II_1 factors with exactly two group measure space decompositions". *J. Math. Pures Appl. (9)* 108.1, pp. 88–110. arXiv: 1512.06677. MR: 3660770 (cit. on p. 1675).

Dusa McDuff (1969). "Uncountably many II_1 factors". *Ann. of Math. (2)* 90, pp. 372–377. MR: 0259625 (cit. on p. 1659).

F. J. Murray and J. von Neumann (1936). "On rings of operators". *Ann. of Math. (2)* 37.1, pp. 116–229. MR: 1503275 (cit. on pp. 1657, 1658).

– (1943). "On rings of operators. IV". *Ann. of Math. (2)* 44, pp. 716–808. MR: 0009096 (cit. on p. 1657).

Donald S. Ornstein and Benjamin Weiss (1980). "Ergodic theory of amenable group actions. I. The Rohlin lemma". *Bull. Amer. Math. Soc. (N.S.)* 2.1, pp. 161–164. MR: 551753 (cit. on p. 1658).

Narutaka Ozawa (2004). "Solid von Neumann algebras". *Acta Math.* 192.1, pp. 111–117. MR: 2079600 (cit. on pp. 1662, 1671, 1681–1683).

– (2006). "A Kurosh-type theorem for type II_1 factors". *Int. Math. Res. Not.* Art. ID 97560, 21. MR: 2211141 (cit. on pp. 1658, 1662, 1681).

– (2008). "Weak amenability of hyperbolic groups". *Groups Geom. Dyn.* 2.2, pp. 271–280. MR: 2393183 (cit. on p. 1673).

– (2012). "Examples of groups which are not weakly amenable". *Kyoto J. Math.* 52.2, pp. 333–344. MR: 2914879 (cit. on p. 1673).

Narutaka Ozawa and Sorin Popa (2004). "Some prime factorization results for type II_1 factors". *Invent. Math.* 156.2, pp. 223–234. MR: 2052608 (cit. on p. 1683).

– (2010a). "On a class of II_1 factors with at most one Cartan subalgebra". *Ann. of Math. (2)* 172.1, pp. 713–749. MR: 2680430 (cit. on pp. 1661, 1667, 1672, 1673, 1678, 1681).

– (2010b). "On a class of II_1 factors with at most one Cartan subalgebra, II". *Amer. J. Math.* 132.3, pp. 841–866. MR: 2666909 (cit. on pp. 1661, 1675, 1682).

Judith A. Packer (1985). "Point spectrum of ergodic abelian group actions and the corresponding group-measure factors". *Pacific J. Math.* 119.2, pp. 381–405. MR: 803126 (cit. on p. 1674).

Jesse Peterson (n.d.). "Examples of group actions which are virtually W*-superrigid". Preprint. arXiv: 1002.1745v1 (cit. on pp. 1661, 1667).

– (2009). "L^2-rigidity in von Neumann algebras". *Invent. Math.* 175.2, pp. 417–433. MR: 2470111 (cit. on pp. 1662, 1678, 1681).

– (2015). "Character rigidity for lattices in higher-rank groups". Preprint (cit. on p. 1663).

Jesse Peterson and Thomas Sinclair (2012). "On cocycle superrigidity for Gaussian actions". *Ergodic Theory Dynam. Systems* 32.1, pp. 249–272. MR: 2873170 (cit. on pp. 1673, 1676–1678).

Jesse Peterson and Andreas Thom (2011). "Group cocycles and the ring of affiliated operators". *Invent. Math.* 185.3, pp. 561–592. MR: 2827095 (cit. on p. 1673).

Sorin Popa (1983). "Orthogonal pairs of $*$-subalgebras in finite von Neumann algebras". *J. Operator Theory* 9.2, pp. 253–268. MR: 703810 (cit. on p. 1681).

– (1986). "Correspondences". In: 56 (cit. on pp. 1667, 1669).

– (2006a). "On a class of type II_1 factors with Betti numbers invariants". *Ann. of Math. (2)* 163.3, pp. 809–899. MR: 2215135 (cit. on pp. 1659, 1660, 1667–1671).

– (2006b). "Some computations of 1-cohomology groups and construction of non-orbit-equivalent actions". *J. Inst. Math. Jussieu* 5.2, pp. 309–332. MR: 2225044 (cit. on p. 1658).

– (2006c). "Some rigidity results for non-commutative Bernoulli shifts". *J. Funct. Anal.* 230.2, pp. 273–328. MR: 2186215 (cit. on p. 1669).

– (2006d). "Strong rigidity of II_1 factors arising from malleable actions of w-rigid groups. I". *Invent. Math.* 165.2, pp. 369–408. MR: 2231961 (cit. on pp. 1659–1661, 1668–1671).

– (2006e). "Strong rigidity of II_1 factors arising from malleable actions of w-rigid groups. II". *Invent. Math.* 165.2, pp. 409–451. MR: 2231962 (cit. on pp. 1659–1661, 1675, 1684).

– (2007a). "Cocycle and orbit equivalence superrigidity for malleable actions of w-rigid groups". *Invent. Math.* 170.2, pp. 243–295. MR: 2342637 (cit. on pp. 1658, 1661, 1662, 1667, 1676–1678).

– (2007b). "Deformation and rigidity for group actions and von Neumann algebras". In: *International Congress of Mathematicians. Vol. I*. Eur. Math. Soc., Zürich, pp. 445–477. MR: 2334200 (cit. on pp. 1658, 1659, 1670, 1676).

– (2007c). "On Ozawa's property for free group factors". *Int. Math. Res. Not. IMRN* 11, Art. ID rnm036, 10. MR: 2344271 (cit. on pp. 1662, 1669, 1671, 1681).

Sorin Popa (2008). "On the superrigidity of malleable actions with spectral gap". *J. Amer. Math. Soc.* 21.4, pp. 981–1000. MR: 2425177 (cit. on pp. 1658, 1661, 1662, 1667, 1671, 1676–1678, 1681).

Sorin Popa and Roman Sasyk (2007). "On the cohomology of Bernoulli actions". *Ergodic Theory Dynam. Systems* 27.1, pp. 241–251. MR: 2297095 (cit. on p. 1658).

Sorin Popa and Stefaan Vaes (2010a). "Actions of $\mathbb{F}_\infty$ whose II_1 factors and orbit equivalence relations have prescribed fundamental group". *J. Amer. Math. Soc.* 23.2, pp. 383–403. MR: 2601038 (cit. on p. 1659).

– (2010b). "Group measure space decomposition of II_1 factors and W^*-superrigidity". *Invent. Math.* 182.2, pp. 371–417. MR: 2729271 (cit. on pp. 1661, 1667, 1675).

– (2011). "Cocycle and orbit superrigidity for lattices in $\mathrm{SL}(n, \mathbb{R})$ acting on homogeneous spaces". In: *Geometry, rigidity, and group actions*. Chicago Lectures in Math. Univ. Chicago Press, Chicago, IL, pp. 419–451. MR: 2807839 (cit. on pp. 1676, 1677).

– (2014a). "Unique Cartan decomposition for II_1 factors arising from arbitrary actions of free groups". *Acta Math.* 212.1, pp. 141–198. MR: 3179609 (cit. on pp. 1661, 1673, 1674).

– (2014b). "Unique Cartan decomposition for II_1 factors arising from arbitrary actions of hyperbolic groups". *J. Reine Angew. Math.* 694, pp. 215–239. MR: 3259044 (cit. on pp. 1661, 1673, 1674, 1682, 1683).

Yehuda Shalom (2005). "Measurable group theory". In: *European Congress of Mathematics*. Eur. Math. Soc., Zürich, pp. 391–423. MR: 2185757 (cit. on p. 1658).

Thomas Sinclair (2011). "Strong solidity of group factors from lattices in $\mathrm{SO}(n, 1)$ and $\mathrm{SU}(n, 1)$". *J. Funct. Anal.* 260.11, pp. 3209–3221. MR: 2776567 (cit. on pp. 1673, 1682).

Isadore M. Singer (1955). "Automorphisms of finite factors". *Amer. J. Math.* 77, pp. 117–133. MR: 0066567 (cit. on pp. 1658, 1660, 1666).

An Speelman and Stefaan Vaes (2012). "A class of II_1 factors with many non conjugate Cartan subalgebras". *Adv. Math.* 231.3-4, pp. 2224–2251. MR: 2964637 (cit. on p. 1675).

Simon Thomas (2003). "Superrigidity and countable Borel equivalence relations". *Ann. Pure Appl. Logic* 120.1-3, pp. 237–262. MR: 1949709 (cit. on pp. 1662, 1680).

– (2006). "Borel superrigidity and the classification problem for the torsion-free abelian groups of finite rank". In: *International Congress of Mathematicians. Vol. II*. Eur. Math. Soc., Zürich, pp. 93–116. MR: 2275590 (cit. on pp. 1662, 1679, 1680).

– (2009). "Popa superrigidity and countable Borel equivalence relations". *Ann. Pure Appl. Logic* 158.3, pp. 175–189. MR: 2500091 (cit. on p. 1658).

Robin Tucker-Drob (n.d.). "Invariant means and the structure of inner amenable groups". Preprint. arXiv: 1407.7474 (cit. on pp. 1676–1678).

Stefaan Vaes (2007). "Rigidity results for Bernoulli actions and their von Neumann algebras (after Sorin Popa)". *Astérisque* 311. Séminaire Bourbaki. Vol. 2005/2006, Exp. No. 961, viii, 237–294. MR: 2359046 (cit. on pp. 1659, 1678).

– (2010). "Rigidity for von Neumann algebras and their invariants". In: *Proceedings of the International Congress of Mathematicians. Volume III*. Hindustan Book Agency, New Delhi, pp. 1624–1650. MR: 2827858 (cit. on pp. 1658, 1659).

– (2013). "One-cohomology and the uniqueness of the group measure space decomposition of a II_1 factor". *Math. Ann.* 355.2, pp. 661–696. MR: 3010143 (cit. on pp. 1674, 1681).

Dan V. Voiculescu (1996). "The analogues of entropy and of Fisher's information measure in free probability theory. III. The absence of Cartan subalgebras". *Geom. Funct. Anal.* 6.1, pp. 172–199. MR: 1371236 (cit. on p. 1681).

Robert J. Zimmer (1980). "Strong rigidity for ergodic actions of semisimple Lie groups". *Ann. of Math. (2)* 112.3, pp. 511–529. MR: 595205 (cit. on pp. 1662, 1676).

– (1984). *Ergodic theory and semisimple groups*. Vol. 81. Monographs in Mathematics. Birkhäuser Verlag, Basel, pp. x+209. MR: 776417 (cit. on pp. 1658, 1662, 1676).

Received 2017-12-01.

Adrian Ioana
Department of Mathematics
University of California San Diego
9500 Gilman Drive
La Jolla, CA 92093
USA
aioana@ucsd.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1691–1708)

SOME 20+ YEAR OLD PROBLEMS ABOUT BANACH SPACES AND OPERATORS ON THEM

William B. Johnson

Abstract

In the last few years numerous 20+ year old problems in the geometry of Banach spaces were solved. Some are described herein.

1 Introduction

In this note I describe some problems in Banach space theory from the 1970s and 1980s that were solved after they had been opened for 20+ years. The problems are mostly not connected to one another, so each section is independent from the other sections. I use standard Banach space notation and terminolgy, as is contained e.g. in Lindenstrauss and Tzafriri [1977] or Albiac and N. J. Kalton [2006]. In this introduction I just recall some definitions that are used repeatedly. Other possibly unfamilar definitions are introduced in the sections in which they are used.

All spaces are Banach spaces and subspaces are closed linear subspaces. An operator is a bounded linear operator between Banach spaces. An isomorphism is a not necessarly surjective linear homeomorphism. $L(X, Y)$ denotes the space of operators from X to Y. This is abbreviated to $L(X)$ when $X = Y$. B_X denotes the closed unit ball of the space X. An operator T with domain X is compact if TB_X has compact closure and is weakly compact if TB_X has weakly compact closure. An operator T is strictly singular if the restriction of T to any infinite dimensional subspace of its domain is not an isomorphism. If Y is a Banach space and T is an operator, T is said to be Y-singular if the restriction of T to any subspace of its domain that is isomorphic to Y is not an isomorphism. So T is strictly singular if T is Y-singular for every infinite dimensional space Y. The isomorphism constant or Banach-Mazur distance between Banach spaces X_1 and X_2 is defined as

$$d(X_1, X_2) = \inf \|T\| \cdot \|T^{-1}\|$$

Research supported in part by NSF DMS-1565826, Clay Mathematics Institute, and MSRI..
MSC2010: primary 46B03; secondary 46B20, 46B28.

where the infimum is taken over all isomorphisms from X_1 onto X_2. So $d(X_1, X_2) = \infty$ if X_1 is not isomorphic to X_2.

2 The diameter of the isomorphism class of a Banach space

Given a Banach space X, define

$$D(X) = \sup\{d(X_1, X_2) : X_1, X_2 \text{ are isomorphic to } X\}.$$

J. J. Schäffer considered the following question to be well-known when he included it in his 1976 book Schäffer [1976]:

Is $D(X) = \infty$ for all infinite dimensional X?

My first PhD student, E. Odell (who died much too young) and I gave an affirmative answer for separable X Johnson and E. Odell [2005]. A new definition helped: Call a Banach space X K*-elastic* provided every isomorph of X K-embeds into X. Call X *elastic* if X is K-elastic for some $K < \infty$.

Ted and I proved the following, which easily implies that $D(X) = \infty$ if X is separable and infinite dimensional.

Theorem 2.1. *If X is a separable Banach space so that for some K, every isomorph of X is K-elastic, then X is finite dimensional.*

The only "obvious" example of a separable elastic space is $C[0, 1]$. It is 1-elastic because Mazur proved that every separable Banach space is isometrically isomorphic to a subspace of $C[0, 1]$. Odell and I suspected that isomorphs of $C[0, 1]$ are the only elastic separable spaces and remarked that our proof of Theorem 2.1 could be streamlined a lot if this is true. We could not prove this but were able to use Bourgain's ℓ_∞ index theory Bourgain [1980] to prove that a separable elastic space contains a subspace that is isomorphic to c_0 and used that information in the proof of Theorem 2.1. Ten years later my third PhD student, D. Alspach, and B. Sari created a new index that they used to verify that our suspicion was correct. Their proof is rather complicated, but even more recently Beanland and Causey [n.d.] simplified the proof somewhat by using more descriptive set theory. It looks likely that the Alspach-Sari index will be used more down the road.

Schäffer's problem remains open for non separable spaces. In some models of set theory (GCH) there are spaces of every density character that are 1-elastic by virtue of being universal, but some tools that were used in the separable setting are not available when the spaces are non separable. Godefroy [2010] proved that under Martin's Maximum Axiom Schäffer's problem has an affirmative answer for subspaces of ℓ_∞.

3 Commutators

The commutator of two elements A and B in a Banach algebra is given by

$$[A, B] = AB - BA.$$

A natural problem that arises in the study of derivations on a Banach algebra $\mathfrak{A}$ is to classify the commutators in the algebra. Probably the most natural non commutative Banach algebras other than C^* algebras are the spaces $\mathcal{L}(\mathcal{X})$ of bounded linear operators on a Banach space $\mathcal{X}$. When X is $n < \infty$ dimensional, $\mathcal{L}(\mathcal{X})$ can be identified with the n by n matrices of scalars, and it is classical that such a matrix is a commutator if and only if it has trace zero. There is generally no trace on $L(X)$ when X is infinite dimensional, and the only general obstruction to an operator being a commutator is due to Wintner [1947], who proved that the identity in a unital Banach algebra is not a commutator. It follows immediately by passing to the quotient algebra $\mathcal{L}(\mathcal{X})/\mathcal{I}(\mathcal{X})$ that no element of the form $\lambda I + K$, where K belongs to a proper norm closed ideal $\mathcal{I}(\mathcal{X})$ of $\mathcal{L}(\mathcal{X})$ and $\lambda \neq 0$, can be a commutator. With this in mind we call a Banach space $\mathcal{X}$ a *Wintner space* provided the only non commutators in $\mathcal{L}(\mathcal{X})$ are elements of the form $\lambda I + K$ with $\lambda \neq 0$ and K in a proper closed ideal.

Here is Wielandt's elegant proof Wielandt [1949] of Wintner's theorem that I is not a commutator:

If $I = AB - BA$ then by induction

$$\forall n \quad A^n B - BA^n = nA^{n-1}.$$

So A cannot be nilpotent and

$$n\|A^{n-1}\| \leq 2\|A\| \cdot \|B\| \cdot \|A^{n-1}\|.$$

To determine whether a Banach space $\mathcal{X}$ is a Wintner space, the first thing one most know is what elements in $\mathcal{L}(\mathcal{X})$ lie in a proper closed ideal, so one needs to know what are the maximal ideals in $\mathcal{L}(\mathcal{X})$ (maximal ideals in a unital Banach algebra are automatically closed because the invertible elements are open). In certain classical spaces, such as ℓ_p for $1 \leq p < \infty$, and c_0, there is only one proper closed ideal; namely, the ideal of compact operators on $\mathcal{X}$, (Gohberg, Markus, and Feldman [1960], see also Whitley [1964, Theorem 6.2]), so it is not surprising that these spaces received the most attention early on. After a decade of so research on commutators by numerous people, in 1965 Brown and Pearcy [1965]) made a breakthrough by proving that ℓ_2 is a Wintner space. In 1972, Apostol [1972a] verified that ℓ_p for $1 < p <$ is a Wintner space and a year later Apostol [1973] he proved that c_0 is a Wintner space. Apostol obtained information about commutators on ℓ_1 and ℓ_∞, and there was also research done around the same time about commutators on L_p, but it was only 30 years later that another classification theorem was proved. In 2009, my student D. Dosev showed in his dissertation that ℓ_1 is a Wintner space. In D. Dosev and Johnson [2010], he and I codified what is needed for the technology developed by Brown–Pearcy, Apostol, and him in order to prove that an operator is a commutator in spaces that have a *Pełczyński decomposition*. (The space $\mathcal{X}$ is said to have a Pełczyński decomposition if X is isomorphic to $\left(\sum \mathcal{X}\right)_p$ with $1 \le p \le \infty$ or $p = 0$.) Notice that if $\mathcal{X}$ has a Pełczyński decomposition then one can define right and left shifts of infinite multiplicity on $\mathcal{X}$. Such shifts can be used to show that certain operators on $\mathcal{X}$ are commutators. In D. Dosev and Johnson [ibid.] the following theorem was proved (but it was only stated in D. Dosev, Johnson, and Schechtman [2013]).

Theorem 3.1. *Let $\mathcal{X}$ be a Banach space such that $\mathcal{X}$ is isomorphic to $\left(\sum \mathcal{X}\right)_p$, where $1 \le p \le \infty$ or $p = 0$. Let $T \in \mathcal{L}(\mathcal{X})$ be such that there exists a subspace $X \subset \mathcal{X}$ such that $\simeq \mathcal{X}$, $T_{|X}$ is an isomorphism, $X + T(X)$ is complemented in $\mathcal{X}$, and $d(X, T(X)) > 0$. Then T is a commutator.*

In practice, Theorem 3.1 allows one to avoid operator theoretic arguments when trying to check whether a space $\mathcal{X}$ is a Wintner space and concentrate on the geometry of $\mathcal{X}$. This is particularly important when $K(\mathcal{X})$ is not the only closed ideal in $\mathcal{L}(\mathcal{X})$, as is the case in all classical spaces other than ℓ_p, $1 \le p < \infty$ and c_0. In D. Dosev and Johnson [2010] Dosev and I used Theorem 3.1 to prove that ℓ_∞ is a Wintner space and in D. Dosev, Johnson, and Schechtman [2013] together with Schechtman we used it to prove that $L_p := L_p(0,1)$ is a Wintner space. In ℓ_∞ the unique maximal ideal is not too bad–it is the ideal of strictly singular operators. However, in L_p, the unique maximal ideal is horrendously large and hard to deal with–it is the ideal of L_p-singular operators. Theorem 3.1 also was used in Chen, Johnson, and Zheng [2011] and Zheng [2014].
Here is a wild conjecture that was made in D. Dosev and Johnson [2010]:
If $\mathcal{X}$ has a Pełczyński decomposition then $\mathcal{X}$ is a Wintner space.
The most interesting classical spaces not known to be Wintner spaces are the spaces $C(K)$ where K is an infinite compact metric space with $C(K)$ not isomorphic to c_0–all of these have a .Pełczyński decomposition. The best partial results on these spaces is contained in Dosev's paper D. T. Dosev [2015]. There are other recent papers that prove that some simpler spaces are Wintner spaces, including Zheng [2014] and Chen, Johnson, and Zheng [2011].
After D. Dosev and Johnson [2010] was written it was proved by Tarbard [2012] that not every infinite dimensional Banach space is a Wintner space. Building on the work of his advisor, R. Haydon, and S. Argyros that solved a famous 40+ year old problem that they will discuss at their 2018 ICM lecture, Tarbard constructed a Banach space $\mathcal{X}$ such that every operator on $\mathcal{X}$ has the form $\lambda I + \alpha S + K$ with λ and α scalars, K is compact, and S is special non compact operator whose square is compact. The strictly singular operators

form the unique maximal ideal in $L(\mathfrak{X})$ and it is clear that S is not a commutator, so $\mathfrak{X}$ is not a Wintner space.

Two other well-known open problems about commutators are worth mentioning.

Problem 1. If X is infinite dimensional, then is every compact operator on X a commutator?

I suspect that, to the contrary, there is an infinite dimensional space X such that every finite rank commutator on X has zero trace.

The following problem is open for *every* infinite dimensional space.

Problem 2. Is every compact operator the commutator of two compact operators?

4 Counting Ideals in $L(L_p)$

After C^* algebras, probably the most natural non commutative Banach algebras are the spaces of bounded linear operators on such classical Banach spaces as $L_p := L_p(0,1)$. In order to study any Banach algebra one must understand something about the closed ideals in the algebra. For $L(\ell_p)$, the space of bounded linear operators on ℓ_p, $1 \le p < \infty$, the situation is the same as for ℓ_2. The only non trivial closed ideal is the ideal of compact operators (see Gohberg, Markus, and Feldman [1960] and Whitley [1964]). The situation for $L(L_p)$, $1 \le p \ne 2 < \infty$, is much more complicated. Let's call an ideal $\mathfrak{I}$ *small* if $\mathfrak{I}$ is contained in the ideal of strictly singular operators. Call an ideal large if it is not small. The most natural way to construct a large ideal in $L(X)$ is to find a complemented subspace Y of X and consider the closed ideal $\mathfrak{I}_Y$ generated by a bounded linear projection from X onto Y. If, as is usually the case, Y is isomorphic to $Y \oplus Y$, this ideal is the closure of the collection of all operators on X that factor through Y. Then $\mathfrak{I}_Y$ is a proper ideal as long as X is not isomorphic to a complemented subspace of Y. Schechtman [1975] proved that $L(L_p)$, $1 < p \ne 2 < \infty$, has at least $\aleph_0$ ideals by constructing $\aleph_0$ isomorphically different complemented subspaces of L_p. With Bourgain and Rosenthal, he Bourgain, Rosenthal, and Schechtman [1981] improved this to $\aleph_1$ by constructing $\aleph_1$ isomorphically different complemented subspaces of L_p. It is still open whether in ZFC $L(L_p)$ has a continuum of large ideals.

Only recently was it proved that $L(L_p)$, $1 < p \ne 2 < \infty$, has infinitely many closed small ideals. In fact, building on some other recent work, Schlumprecht and Zsák [2018] show that $L(L_p)$ has a continuum of small closed ideals, solving in the process a problem in Pietsch's 1978 book Pietsch [1978]. It remains open whether $L(L_p)$, $1 < p \ne 2 < \infty$, has more than a continuum of closed ideals.

For $L(L_1)$, the situation was stagnant for an even longer time. In 1978 Pietsch [ibid.] recorded the well-known problem whether there are infinitely many closed ideals in $L(L_1)$.

At that time the only non trivial ideals in $L(L_1)$ known were the ideal of compact operators, the ideal of strictly singular operators, the ideal of operators that factor through ℓ_1, and the unique maximal ideal. It is easy to write down candidates for other ideals, but many turn out to be one of these four. For example, if $1 < p \leq \infty$, the closure of the operators on L_1 that factor through L_p is the ideal of weakly compact operators, and on L_1 an operator is weakly compact if and only if it is strictly singular. Just in the past year, Johnson, Pisier, and Schechtman [n.d.] proved that there are other closed ideals. We constructed a continuum of closed small ideals in $L(L_1)$. For $2 < p < \infty$ we take a $\Lambda(p)$ sequence (x_n^p) of characters that has certain extra properties ("$\Lambda(p)$" means that the L_p and L_2 norms are equivalent on the linear span of the set of characters). Let J_p be the bounded linear operator from ℓ_1 into L_1 that maps the nth unit basis vector to x_n^p and let $\mathfrak{I}(p)$ be the closure of the operators on L_1 that factor through J_p. It turns out that $\mathfrak{I}(p) \neq \mathfrak{I}(q)$ when $p \neq q$.
It is open whether $L(L_1)$ has more than two large ideals. This is closely connected to the famous problem whether every infinite dimensional complemented subspace of L_1 is isomorphic either to ℓ_1 or to L_1.

5 Spaces that are uniformly homeomorphic to $\mathcal{L}_1$ spaces

Banach spaces X and Y are said to be uniformly homeomorphic if there is an injective uniformly continuous function from X onto Y whose inverse is uniformly continuous. B. Maurey, G. Schechtman, and I gave an affirmative answer to the 1982 question of Heinrich and Mankiewicz [1982]:

Are the $\mathcal{L}_1$ spaces are preserved under uniform homeomorphisms?

A Banach space X is said to be $\mathcal{L}_1$ if its dual X^* is isomorphic to $C(K)$ for some compact Hausdorff space K. That is really a theorem Lindenstrauss and Rosenthal [1969]. The definition Lindenstrauss and Pełczyński [1968] is that X is the increasing union of finite dimensional subspaces that are uniformly isomorphic to finite dimensional L_1 spaces.
Subsequently N. J. Kalton [2012] proved that this theorem is optimal by constructing two separable $\mathcal{L}_1$ spaces that are uniformly homeomorphic but not isomorphic.
At the heart of the question is a recurring problem:
Suppose a linear mapping $T : X \to Y$ admits a Lipschitz factorization through a Banach space Z; i.e., we have Lipschitz $F_1 : X \to Z$ and $F_2 : Z \to Y$ and $F_2 \circ F_1 = T$. What extra is needed to guarantee that T admits a linear factorization through Z?
Something extra is needed because the identity on $C[0,1]$ Lipschitz factors through c_0 Aharoni [1974], Lindenstrauss [1964].
The main result in Johnson, Maurey, and Schechtman [2009] is

Theorem 5.1. *Let X be a finite dimensional normed space, Y a Banach space with the Radon-Nikodym property (which means that every Lipschitz mapping from the real line into Y is differentiable almost everywhere) and $T : X \to Y$ a linear operator. Let Z be a separable Banach space and assume there are Lipschitz maps $F_1 : X \to Z$ and $F_2 : Z \to Y$ with $F_2 \circ F_1 = T$. Then for every $\lambda > 1$ there are linear maps $T_1 : X \to L_\infty(Z)$ and $T_2 : L_1(Z) \to Y$ with $T_2 \circ i_{\infty,1} \circ T_1 = T$ and $\|T_1\| \cdot \|T_2\| \leq \lambda \mathrm{Lip}\,(F_1)\mathrm{Lip}\,(F_2)$.*

If Z is $\mathfrak{L}_1$ then so is $L_1(Z)$ and hence T linearly factors through a $\mathfrak{L}_1$ space.
This and fairly standard tools in non linear geometric functional analysis give an affirmative answer to the Heinrich–Mankiewicz problem. The proof of Theorem 5.1 is based on a rather simple local-global linearization idea. For the application we need only the case where Y is finite dimensional.

6 Weakly null sequences in L_1

The first weakly null normalized sequences (WNNS) with no unconditional sub-sequence were constructed by Maurey and Rosenthal [1977]. Their technique was incorporated into the famous paper of Gowers and Maurey [1993] that contains an example of an infinite dimensional Banach space that contains NO unconditional sequence, but the examples in Maurey and Rosenthal [1977] are still interesting because the ambient spaces were $C(K)$ with K countable. These $C(K)$ spaces are hereditarily c_0 and so have unconditional sequences all over the place. Every subsequence of the WNNS they constructed reproduces the (conditional) summing basis on blocks.
In 1977 Maurey and Rosenthal [ibid.] asked whether every WNNS sequence in $L_1 := L_1(0,1)$ has an unconditional subsequence. Like the $C(K)$ spaces with K countable, every infinite dimensional subspace of L_1 contains an unconditional sequence. In Johnson, Maurey, and Schechtman [2007] we constructed a WNNS in L_1 such that every subsequence contains a block basis that is $1+\epsilon$–equivalent to the (conditional) Haar basis for L_1, which implies that the WNNS has no unconditional subsequence. In fact, the theorem stated this way extends to rearrangement invariant spaces which (in some appropriate sense) are not to the right of L_2 (e.g. L_p, $1 < p < 2$) and which are not too close to L_∞.

7 Subspaces of spaces that have an unconditional basis

A problem that goes back to the 1970s is to give an intrinsic characterization of Banach spaces that isomorphically embed into a space that has an unconditional basis. It was shown that every space with an unconditional expansion of the identity (in particular, every space with an unconditional finite dimensional decompostion) embeds into a space with unconditional basis Pełczyński and Wojtaszczyk [1971], Lindenstrauss and Tzafriri

[1977]. However for spaces that lack such a strong approximation property the only apparently useful invariant is that in a subspace of a space with unconditional basis, every weakly null normalized sequence (WNNS) has an unconditional subsequence. A quotient of a space with shrinking unconditional basis has this desirable property Johnson [1977], E. Odell [1992]. (A basis is shrinking provided the linear functionals biorthogonal to the basis vectors are a basis for the dual space. Every basis for a reflexive space is shrinking.) But this condition is not suffiencient even for reflexive spaces: in Johnson and Zheng [2008] my student B. Zheng and I used a variation of a construction of E. W. Odell and Schlumprecht [2006] to build a separable reflexive space that does not embed into a space with unconditional basis, yet every WNNS in the space has an unconditional subsequence. On the other hand, Feder [1980] proved that a reflexive quotient X of a space with shrinking unconditional basis embeds into a space with unconditional basis as long as X has the approximation property. Unfortunately, all classical reflexive spaces other than Hilbert spaces have quotients that fail the approximation property.
So there were two problems
1. Give an intrinsic characterization of Banach spaces that embed into a space that has an unconditional basis.
2. Does every quotient of a space with shrinking unconditional basis embed into a space with unconditional basis?
Much research centered around reflexive spaces. For example, in addition to the result of Feder mentioned above, it was proved that every reflexive subspace of a space with unconditional basis embeds into a reflexive space with unconditional basis Davis, Figiel, Johnson, and Pełczyński [1974], Figiel, Johnson, and Tzafriri [1975].
In Johnson and Zheng [2008] Zheng and I answered both problems in the affirmative for reflexive spaces. Our later paper Johnson and Zheng [2011] gives an affirmative answer to (2) in general and to (1) for spaces that have a separable dual. The answer to (1) for spaces with non separable dual must be completely different because of the space ℓ_1, which has an unconditional basis but also has the Schur property–every WNNS converges in norm to zero.
The answers for reflexive spaces follow from the following omnibus theorem, which basically says that every condition that *might* be equivalent to "the reflexive space X embeds into a space with an unconditional basis" actually *is* equivalent to it.

Theorem 7.1. *Let X be a separable reflexive Banach space. Then the following are equivalent.*

(a) X has the UTP.

(b) X is isomorphic to a subspace of a Banach space with an unconditional basis.

(c) X is isomorphic to a subspace of a reflexive space with an unconditional basis.

(d) *X is isomorphic to a quotient of a Banach space with a shrinking unconditional basis.*

(e) *X is isomorphic to a quotient of a reflexive space with an unconditional basis.*

(f) *X is isomorphic to a subspace of a quotient of a reflexive space with an unconditional basis.*

(g) *X is isomorphic to a subspace of a reflexive quotient of a Banach space with a shrinking unconditional basis.*

(h) *X is isomorphic to a quotient of a subspace of a reflexive space with an unconditional basis.*

(i) *X is isomorphic to a quotient of a reflexive subspace of a Banach space with a shrinking unconditional basis.*

(j) *X^* has the UTP.*

The UTP is a strengthening of the property "every WNNS has an unconditional subsequence". The weaker property for a reflexive space does NOT imply embeddability into a space with unconditional basis Johnson and Zheng [2008]. The definition of the UTP is due to E. W. Odell and Schlumprecht [2006]:

Definition 7.2. *A branch of a tree is a maximal linearly ordered subset of the tree under the tree order. We say X has the C-unconditional tree property (C-UTP) if every normalized weakly null infinitely branching tree in X has a C-unconditional branch. X has the UTP if X has the C-UTP for some $C > 0$.*

The proof of the theorem uses some new tricks, blocking methods developed in the 1970s Johnson and Zippin [1972], Johnson and Zippin [1974], Johnson and E. Odell [1974], Johnson and E. Odell [1981], Johnson [1977], and the Odell-Schlumprecht analysis E. W. Odell and Schlumprecht [2006] relating tree properties to embeddability into spaces that have a finite dimensional decomposition with the corresponding skipped blocking property.

For Banach spaces with a separable dual, there is a similar theorem Johnson and Zheng [2011], but the characterization involves the weak* UTP. A Banach space X is said to have the weak* UTP provided every normalized weak* null infinitely branching tree in X^* has a branch that is an unconditional basic sequence. The main new technical feature in Johnson and Zheng [ibid.] is that blocking and "killing the overlap" techniques originally developed for finite dimensional decompositions are adapted to work for blockings of shrinking M-bases (that is, biorthogonal sequences $\{x_n, x_n^*\}$ with span x_n dense in X and span x_n^* dense in X^*). Shrinking M-bases are known to exist in every Banach space that has a separable dual. These technical advances provide some simplifications of the argument in the reflexive case presented in Johnson and Zheng [2008] and likely will be used in the future to study the structure of Banach spaces that lack a good approximation property.

8 Operators on ℓ_∞ with dense range

In http://mathoverflow.net/questions/101253 A. B. Nasseri asked
"Can anyone give me an example of an *(sic)* bounded and linear operator $T : \ell_\infty \to \ell_\infty$ (the space of bounded sequences with the usual sup-norm), such that T has dense range, but is not surjective?"
This question quickly drew two close votes. Nevertheless it took a couple of years for Nasseri, G. Schechtman, T. Tkocz, and me to resolve it Johnson, Nasseri, Schechtman, and Tkocz [2015].
On separable infinite dimensional spaces, there are always dense range compact operators, but compact operators have separable ranges. On a non separable space, even on a dual to a separable space, it can happen that every dense range operator is surjective: Argyros, Arvanitakis, and Tolias [2006] constructed a separable space X so that X^* is non separable, hereditarily indecomposable (HI) in the sense of Gowers–Maurey, and every strictly singular operator on X^* is weakly compact. Since X^* is HI, every operator on X^* is of the form $\lambda I + S$ with S strictly singular Gowers and Maurey [1993]. If $\lambda \neq 0$, then $\lambda I + S$ is Fredholm of index zero by Kato's classical perturbation theory. On the other hand, since every weakly compact subset of the dual to a separable space is norm separable, every strictly singular operator on X^* has separable range.
It turns out that Nasseri's problem is related to Tauberian operators on $L_1 := L_1(0,1)$. An operator $T : X \to Y$ is called *Tauberian* if $T^{**-1}(Y) = X$ N. Kalton and Wilansky [1976]. The book of González and Martínez-Abejón [2010] on Tauberian operators contains:

Theorem 8.1. *Let $T : L_1(0,1) \to Y$. The following are equivalent.*

0. *T is Tauberian.*

1. *For all normalized disjoint sequences $\{x_i\}$, $\liminf_{i\to\infty} \|Tx_i\| > 0$.*

2. *If $\{x_i\}$ is equivalent to the unit vector basis of ℓ_1 then there is an N such that $T_{|[x_i]_{i=N}^\infty}$ is an isomorphism.*

3. *There are $\varepsilon, \delta > 0$ such that $\|Tf\| \geq \varepsilon\|f\|$ for all f with $|\mathrm{supp}(f)| < \delta$.*

What is the connection between Tauberian operators on L_1 and dense range, non surjective operators on ℓ_∞? If T is injective Tauberian, T^{**} is injective. Thus, if T is a Tauberian operator on L_1 that is injective but does not have closed range, then T^* is a dense range operator on L_∞ that is not surjective. Since L_∞ is isomorphic to ℓ_∞, having an injective Tauberian, non closed range operator on L_1 gives a positive answer to Nasseri's question. In fact, we checked that whether there is such an operator on L_1 is *a priori* equivalent to Nasseri's question.

One of the main open problems mentioned in González and Martínez-Abejón [ibid.], raised in 1984 by Weis and Wolff [1984], is whether there is a Tauberian operator T on L_1 whose kernel is infinite dimensional. If T satisfies this condition, then you can play around and get a perturbation S of T that is Tauberian, injective, and has dense, non closed range (so is not surjective). Taking the adjoint of S and replacing L_∞ by its isomorph ℓ_∞, you would have an injective, dense range, non surjective operator on ℓ_∞. (To get S from T, take an injective nuclear operator from the kernel on T that has dense range in L_1, extend it to a nuclear operator on L_1, and add it to T. This does not quite work, but some fiddling produces the desired S.) In fact, without knowing the solution to either problem, one can check that the Weis—Wolff question is equivalent to Nasseri's question. The bottom line is that the question whether there is a dense range non surjective operator on the non separable space ℓ_∞ is really a question about the existence of a Tauberian operator with infinite dimensional kernel on the separable space L_1.

It happened that T satisfying condition (3) in Theorem 8.1 and having an infinite dimensional kernel has a known finite dimensional analogue:

Theorem 8.2. *[CS result] For each n sufficiently large, putting $m = [3n/4]$, there is an operator $T : \ell_1^n \to \ell_1^m$ such that $\frac{1}{4}\|x\|_1 \le \|Tx\|_1 \le \|x\|_1$ for all x with $\sharp\mathrm{supp}(x) \le n/400$.*

This CS result (where "CS" can be interpreted either to mean "Computer Science" or "Compressed Sensing") is a very special case of a theorem due to Berinde, Gilbert, Indyk, Karloff, and Strauss [2008]. The kernel of T_n has dimension at least $n/4$, so if you take the ultraproduct $\tilde{T}$ of the T_n you get an operator with infinite dimensional kernel on some gigantic L_1 space. Let T be the restriction of $\tilde{T}$ to some separable $\tilde{T}$-invariant L_1 subspace that intersects the kernel of $\tilde{T}$ in an infinite dimensional subspace. As long as $\tilde{T}$ is Tauberian, the operator T will be a Tauberian operator with infinite dimensional kernel on L_1, and we will be done. It remains to isolate a condition implying Tauberianism that is possessed by all T_n and is preserved under ultraproducts.

Say an operator $T : X \to Y$ (X an L_1 space) is (r, N)-Tauberian provided whenever $(x_n)_{n=1}^N$ are disjoint unit vectors in X, then $\max_{1 \le n \le N} \|Tx_n\| \ge r$.

Lemma 8.3. *$T : X \to Y$ is Tauberian iff $\exists\, r > 0$ and N such that T is (r, N)-Tauberian.*

Proof: T being (r, N)-Tauberian implies that if (x_n) is a disjoint sequence of unit vectors in X, then $\liminf_n \|Tx_n\| > 0$, so T is González and Martínez-Abejón [2010]. Conversely, suppose there are disjoint collections $(x_k^n)_{k=1}^n$, $n = 1, 2, \ldots$ with $\max_{1 \le k \le n} \|Tx_k^n\| \to 0$ as $n \to \infty$. Then the closed sublattice generated by $\cup_{n=1}^\infty (x_k^n)_{k=1}^n$ is a separable L_1 space, hence is order isometric to $L_1(\mu)$ for some probability measure μ by Kakutani's theorem. Choose $1 \le k(n) \le n$ so that the support of $x_{k(n)}^n$ in $L_1(\mu)$ has measure at most $1/n$. Since T is Tauberian, necessarily $\liminf_n \|Tx_{k(n)}^n\| > 0$ González and Martínez-Abejón [ibid.], a contradiction.
It is not difficult to prove that the property of being (r, N)-Tauberian is stable under ultraproducts of uniformly bounded operators, so it is just a matter of observing that the operators T_n of Berinde, Gilbert, Indyk, Karloff, and Strauss [2008]. are all $(1/4, 400)$–Tauberian. □
Conclusion: There is a non surjective Tauberian operator on L_1 that has dense range. The operator can be chosen either to be injective or to have infinite dimensional kernel. Consequently, there is a dense range, non surjective, injective operator on ℓ_∞.
Conclusion from the proof: Computer science has applications to non separable Banach space theory!

9 Approximation properties

A Banach space X has the approximation property (AP) provided the identity operator is the limit of finite rank operators in the topology of uniform convergence on compact sets. If these operators can be taken to be uniformly bounded, we say that X has the bounded approximation property (BAP) or λ-BAP if the uniform bound can be λ. Grothendieck [1955] proved that a reflexive space that has the AP must have the 1-BAP, but there are non reflexive spaces that have the AP but fail the BAP Figiel and Johnson [1973]. Sometimes these properties come up when considering problems that, on the surface, have nothing to do with approximation. For example, given a family $\mathfrak{F}$ of operators between Banach spaces, it is natural to try to find a single (usually separable) Banach space Z such that all the operators in $\mathfrak{F}$ factor through Z. If $\mathfrak{F}$ is the collection of all operators between separable Banach spaces that have the BAP, there is such a separable Z; namely, the separable universal basis space of Pełczyński [1969], Pełczyński [1971], Kadec [1971]. This space, as well as smaller (even reflexive) spaces Johnson [1971] have the property that every operator that is uniformly approximable by finite rank operators factors through Z. A. Szankowski and I proved that there is not a separable space such that every operator between separable spaces (not even every operator between spaces that have the AP) factors through it Johnson and Szankowski [1976], but this paper left open the question: Is there a separable space such that every *compact* operator factors through it? 23 years later, in part 2 of Johnson and Szankowski [ibid.], we finally managed to proved that no such space exists Johnson and Szankowski [2009].

* * * * * * * *

A Banach space has the *hereditary approximation property* (HAP) provided every subspace has the approximation property. There are non Hilbertian spaces that have the HAP Johnson [1980], Pisier [1988]. All of these examples are *asymptotically Hilbertian*; i.e., for some K and every n, there is a finite codimensional subspace all of whose n-dimensional subspaces are K-isomorphic to ℓ_2^n. An asymptotically Hilbertian space must be superreflexive and cannot have a symmetric basis unless it is isomorphic to a Hilbert space. This led to two problems Johnson [1980]:

1. Can a non reflexive space have the HAP?
2. Does there exist a non Hilbertian space with a symmetric basis that has the HAP?

The HAP is very difficult to work with, partly because it does not have good permanence properties–there are spaces X and Y that have the HAP such that $X \oplus Y$ fails the HAP Casazza, García, and Johnson [2001].

The main result of Johnson and Szankowski [2012] gives an affirmative answer to problem 2 from Johnson [1980]:

Theorem 9.1. *There is a function $f(n) \uparrow \infty$ such that if for infinitely many n we have $D_n(X) \leq f(n)$, then X has the HAP.*

Here $D_n(X) := \sup d(E, \ell_2^n)$, where the sup is over all n-dimensional subspaces of X. The proof combines the ideas in Johnson [ibid.] with the argument in Lindenstrauss and Tzafriri [1976].

You can build Banach spaces with a symmetric basis, even Orlicz sequence spaces, that are not isomorphic to a Hilbert space and yet $D_n(X)$ goes to infinity as slowly as is desired. Hence problem (2) has an affirmative answer.

It turns out that Theorem 9.1 can be used to give a footnote to the famous theorem of J. Lindenstrauss and L. Tzafriri Lindenstrauss and Tzafriri [1971] that Hilbert spaces are the only, up to isomorphism, Banach spaces in which every subspace is complemented. Timur Oikhberg asked us whether there is a non Hilbertian Banach space in which every subspace is isomorphic to a complemented subspace.

Theorem 9.2. *Johnson and Szankowski [2012] There is a separable, infinite dimensional Banach space not isomorphic to ℓ_2 that is complementably universal for all subspaces of all of its quotients.*

Let X be any non Hilbertian separable Banach space such that $D_{4^n}(X) \leq f(n)$ for all n. Let (E_k) be a sequence of finite dimensional spaces that is dense (in the sense of the Banach-Mazur distance) in the collection of all finite dimensional spaces that are contained in some quotient of $\ell_2(X)$ and let Y be the ℓ_2-sum of the E_k. Then $D_n(Y) \leq f(n)$ for all n. If you are old enough to know the right background, you can give a short argument to prove that Y is complementably universal for all subspaces of all of its quotients. Problem (1) remains open.

* * * * * * * * *

It was a privilege for Tadek Figiel and me to be co-authors on A. Pełczyński's last paper Figiel, Johnson, and Pełczyński [2011]. The solution to a (not especially important) problem that had eluded Tadek and me in the early 1970s Figiel and Johnson [1973] just dropped out, so I have an excuse to include a discussion of part of Figiel, Johnson, and Pełczyński [2011] in this note.

Let X be a Banach space, let $Y \subseteq X$ be a subspace, let $\lambda \geq 1$. The pair (X, Y) is said to have the λ-BAP if for each $\lambda' > \lambda$ and each subspace $F \subseteq X$ with $\dim F < \infty$, there is a finite rank operator $u : X \to X$ such that $||u|| < \lambda'$, $u(x) = x$ for $x \in F$ and $u(Y) \subseteq Y$. If (X, Y) has the λ-BAP then X/Y has the λ-BAP. Thus by a theorem due to Szankowski [2009], for $1 \leq p < 2$ there are subspaces Y of ℓ_p that have the BAP and yet (ℓ_p, Y) fails the BAP.

It is open whether (X, Y) has the BAP if X, Y, and X/Y all have the BAP, but I don't believe it.

If Y is a finite dimensional subspace of X and X has the λ-BAP then also (X, Y) has the λ-BAP and hence also X/Y has the λ-BAP. That is, the λ-BAP passes to quotients by finite dimensional subspaces. By duality you get that if X^* the λ-BAP then every finite codimensional subspace of X has the λ-BAP. In particular, every finite codimensional subspace of an L_1 space has the 1-BAP. Easy as this is, I don't think that anyone previously had noticed this.

In fact,

Proposition 9.3. *X^* has the λ-BAP iff (X, Y) has the λ-BAP for every finite codimensional subspace Y.*

The following proposition turned out to be useful.

Proposition 9.4. *Let X be a Banach space and let $Y \subseteq X$ be a closed subspace such that* $\dim X/Y = n < \infty$ *and Y has the $\lambda-BAP$. Then the pair (X, Y) has the $3\lambda-BAP$.*

This gives the corollary

Corollary 9.5. *If X is a Banach space and Y has the $\lambda-BAP$ for every finite codimensional subspace $Y \subseteq X$, then X^* has the $3\lambda-BAP$.*

Consequently, in contradistinction to the case of commutative L_1 spaces, for every λ there are finite codimensional subspaces Y of the non commutative L_1 space S_1 of trace class operators on ℓ_2 that fail the λ-BAP because Szankowski [1981] proved that $L(\ell_2)$ fails the AP and $L(\ell_2)$ is the dual to S_1.

The main result in my 1972 paper with Figiel and Johnson [1973] is that there is a subspace of c_0 that has the AP but fails the BAP. We could not prove the same result for ℓ_1.

Corollary 9.6. *Figiel, Johnson, and Pełczyński [2011] There is a subspace Y of ℓ_1 that has the AP but fails the BAP.*

Proof. Start with a subspace X of ℓ_1 that fails the approximation property Szankowski [1981]. From the existence of such a space it follows Johnson [1972] that if we let Z be the ℓ_1-sum of a dense sequence (X_n) of finite dimensional subspaces of X, then Z^* fails the BAP and yet Z has the BAP. Then Y can be the ℓ_1- sum of a suitable sequence of finite codimensional subspaces of Z because of Corollary 9.5.

References

Israel Aharoni (1974). "Every separable metric space is Lipschitz equivalent to a subset of c_0^+". *Israel J. Math.* 19, pp. 284–291. MR: 0511661 (cit. on p. 1696).

Fernando Albiac and Nigel J. Kalton (2006). *Topics in Banach space theory*. Vol. 233. Graduate Texts in Mathematics. Springer, New York, pp. xii+373. MR: 2192298 (cit. on p. 1691).

Dale E. Alspach and Bünyamin Sarı (2016). "Separable elastic Banach spaces are universal". *J. Funct. Anal.* 270.1, pp. 177–200. MR: 3419759.

Constantin Apostol (1972a). "Commutators on l^p-spaces". *Rev. Roumaine Math. Pures Appl.* 17, pp. 1513–1534. MR: 0336432 (cit. on p. 1693).

– (1972b). "On the left essential spectrum and non-cyclic operators in Banach spaces". *Rev. Roumaine Math. Pures Appl.* 17, pp. 1141–1147. MR: 0320775.

– (1973). "Commutators on c_0-spaces and on l^∞-spaces". *Rev. Roumaine Math. Pures Appl.* 18, pp. 1025–1032. MR: 0336433 (cit. on p. 1693).

Spiros A. Argyros, Alexander D. Arvanitakis, and Andreas G. Tolias (2006). "Saturated extensions, the attractors method and hereditarily James tree spaces". In: *Methods in Banach space theory*. Vol. 337. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, pp. 1–90. MR: 2326379 (cit. on p. 1700).

Spiros A. Argyros and Richard G. Haydon (2011). "A hereditarily indecomposable $\mathcal{L}_\infty$-space that solves the scalar-plus-compact problem". *Acta Math.* 206.1, pp. 1–54. MR: 2784662.

Kevin Beanland and Ryan Causey (n.d.). "On a generalization of Bourgain's tree index". arXiv: 1603.01133 (cit. on p. 1692).

Radu Berinde, Anna C Gilbert, Piotr Indyk, Howard Karloff, and Martin J Strauss (2008). "Combining geometry and combinatorics: A unified approach to sparse signal recovery". In: *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*. IEEE, pp. 798–805 (cit. on pp. 1701, 1702).

J. Bourgain (1980). "On separable Banach spaces, universal for all separable reflexive spaces". *Proc. Amer. Math. Soc.* 79.2, pp. 241–246. MR: 565347 (cit. on p. 1692).

J. Bourgain, H. P. Rosenthal, and Gideon Schechtman (1981). "An ordinal L^p-index for Banach spaces, with application to complemented subspaces of L^p". *Ann. of Math. (2)* 114.2, pp. 193–228. MR: 632839 (cit. on p. 1695).

Arlen Brown and Carl Pearcy (1965). "Structure of commutators of operators". *Ann. of Math. (2)* 82, pp. 112–127. MR: 0178354 (cit. on p. 1693).

P. G. Casazza, C. L. García, and William B. Johnson (2001). "An example of an asymptotically Hilbertian space which fails the approximation property". *Proc. Amer. Math. Soc.* 129.10, pp. 3017–3023. MR: 1840107 (cit. on p. 1703).

Dongyang Chen, William B. Johnson, and Bentuo Zheng (2011). "Commutators on $(\sum \ell_q)_p$". *Studia Math.* 206.2, pp. 175–190. MR: 2860306 (cit. on p. 1694).

W. J. Davis, T. Figiel, William B. Johnson, and A. Pełczyński (1974). "Factoring weakly compact operators". *J. Functional Analysis* 17, pp. 311–327. MR: 0355536 (cit. on p. 1698).

D. Dosev and Wiliam B. Johnson (2010). "Commutators on ℓ_∞". *Bull. Lond. Math. Soc.* 42.1, pp. 155–169. MR: 2586976 (cit. on p. 1694).

Detelin T. Dosev (2009). "Commutators on l_1". *J. Funct. Anal.* 256.11, pp. 3490–3509. MR: 2514050.

– (2015). "On a class of operators on $C(K)$". *Houston J. Math.* 41.2, pp. 595–610. MR: 3403748 (cit. on p. 1694).

Detelin Dosev, William B. Johnson, and Gideon Schechtman (2013). "Commutators on L_p, $1 \le p < \infty$". *J. Amer. Math. Soc.* 26.1, pp. 101–127. MR: 2983007 (cit. on p. 1694).

Moshe Feder (1980). "On subspaces of spaces with an unconditional basis and spaces of operators". *Illinois J. Math.* 24.2, pp. 196–205. MR: 575060 (cit. on p. 1698).

T. Figiel and William B. Johnson (1973). "The approximation property does not imply the bounded approximation property". *Proc. Amer. Math. Soc.* 41, pp. 197–200. MR: 0341032 (cit. on pp. 1702, 1704, 1705).

T. Figiel, William B. Johnson, and L. Tzafriri (1975). "On Banach lattices and spaces having local unconditional structure, with applications to Lorentz function spaces". *J. Approximation Theory* 13. Collection of articles dedicated to G. G. Lorentz on the occasion of his sixty-fifth birthday, IV, pp. 395–412. MR: 0367624 (cit. on p. 1698).

Tadeusz Figiel, William B. Johnson, and Aleksander Pełczyński (2011). "Some approximation properties of Banach spaces and Banach lattices". *Israel J. Math.* 183, pp. 199–231. MR: 2811159 (cit. on pp. 1704, 1705).

Gilles Godefroy (2010). "On the diameter of the Banach-Mazur set". *Czechoslovak Math. J.* 60(135).1, pp. 95–100. MR: 2595073 (cit. on p. 1692).

I. C. Gohberg, A. S. Markus, and I. A. Feldman (1960). "Normally solvable operators and ideals associated with them". *Bul. Akad. Štiince RSS Moldoven.* 1960.10 (76), pp. 51–70. MR: 0218920 (cit. on pp. 1693, 1695).

Manuel González and Antonio Martínez-Abejón (2010). *Tauberian operators*. Vol. 194. Operator Theory: Advances and Applications. Birkhäuser Verlag, Basel, pp. xii+245. MR: 2574170 (cit. on pp. 1700–1702).

W. T. Gowers and B. Maurey (1993). "The unconditional basic sequence problem". *J. Amer. Math. Soc.* 6.4, pp. 851–874. MR: 1201238 (cit. on pp. 1697, 1700).

Alexandre Grothendieck (1955). "Produits tensoriels topologiques et espaces nucléaires". *Mem. Amer. Math. Soc.* No. 16, p. 140. MR: 0075539 (cit. on p. 1702).

S. Heinrich and P. Mankiewicz (1982). "Applications of ultrapowers to the uniform and Lipschitz classification of Banach spaces". *Studia Math.* 73.3, pp. 225–251. MR: 675426 (cit. on p. 1696).

Wiliam B. Johnson and E. Odell (1974). "Subspaces of L_p which embed into l_p". *Compositio Math.* 28, pp. 37–49. MR: 0352938 (cit. on p. 1699).

– (2005). "The diameter of the isomorphism class of a Banach space". *Ann. of Math. (2)* 162.1, pp. 423–437. MR: 2178965 (cit. on p. 1692).

Wiliam B. Johnson and A. Szankowski (1976). "Complementably universal Banach spaces". *Studia Math.* 58.1, pp. 91–97. MR: 0425582 (cit. on pp. 1702, 1703).

Wiliam B. Johnson and M. Zippin (1972). "On subspaces of quotients of $(\sum G_n)_{l_p}$ and $(\sum G_n)_{c_0}$". *Israel J. Math.* 13, 311–316 (1973). MR: 0331023 (cit. on p. 1699).

William B. Johnson (1971). "Factoring compact operators". *Israel J. Math.* 9, pp. 337–345. MR: 0290133 (cit. on p. 1702).

– (1972). "A complementary universal conjugate Banach space and its relation to the approximation problem". *Israel J. Math.* 13, 301–310 (1973). MR: 0326356 (cit. on p. 1705).

– (1977). "On quotients of L_p which are quotients of l_p". *Compositio Math.* 34.1, pp. 69–89. MR: 0454595 (cit. on pp. 1698, 1699).

– (1980). "Banach spaces all of whose subspaces have the approximation property". In: *Special topics of applied mathematics (Proc. Sem., Ges. Math. Datenverarb., Bonn, 1979)*. North-Holland, Amsterdam-New York, pp. 15–26. MR: 585146 (cit. on p. 1703).

William B. Johnson, B. Maurey, and Gideon Schechtman (2009). "Non-linear factorization of linear operators". *Bull. Lond. Math. Soc.* 41.4, pp. 663–668. MR: 2521361 (cit. on p. 1696).

William B. Johnson, B. Maurey, Gideon Schechtman, and L. Tzafriri (1979). "Symmetric structures in Banach spaces". *Mem. Amer. Math. Soc.* 19.217, pp. v+298. MR: 527010.

William B. Johnson, Bernard Maurey, and Gideon Schechtman (2007). "Weakly null sequences in L_1". *J. Amer. Math. Soc.* 20.1, pp. 25–36. MR: 2257395 (cit. on p. 1697).

William B. Johnson, Amir Bahman Nasseri, Gideon Schechtman, and Tomasz Tkocz (2015). "Injective Tauberian operators on L_1 and operators with dense range on ℓ_∞". *Canad. Math. Bull.* 58.2, pp. 276–280. MR: 3334922 (cit. on p. 1700).

William B. Johnson and E. Odell (1981). "Subspaces and quotients of $l_p \oplus l_2$ and X_p". *Acta Math.* 147.1-2, pp. 117–147. MR: 631092 (cit. on p. 1699).

William B. Johnson, G. Pisier, and G. Schechtman (n.d.). "Ideals in $L(L_1)$". In preparation (cit. on p. 1696).

William B. Johnson and A. Szankowski (2009). "Complementably universal Banach spaces. II". *J. Funct. Anal.* 257.11, pp. 3395–3408. MR: 2571432 (cit. on p. 1703).

– (2012). "Hereditary approximation property". *Ann. of Math. (2)* 176.3, pp. 1987–2001. MR: 2979863 (cit. on p. 1703).

William B. Johnson and Bentuo Zheng (2008). "A characterization of subspaces and quotients of reflexive Banach spaces with unconditional bases". *Duke Math. J.* 141.3, pp. 505–518. MR: 2387429 (cit. on pp. 1698–1700).

– (2011). "Subspaces and quotients of Banach spaces with shrinking unconditional bases". *Israel J. Math.* 185, pp. 375–388. MR: 2837142 (cit. on pp. 1698, 1699).

William B. Johnson and M. Zippin (1974). "Subspaces and quotient spaces of $(\sum G_n)_{l_p}$ and $(\sum G_n)_{c_0}$". *Israel J. Math.* 17, pp. 50–55. MR: 0358296 (cit. on p. 1699).

M. Ĭ. Kadec (1971). "On complementably universal Banach spaces". *Studia Math.* 40, pp. 85–89. MR: 0313764 (cit. on p. 1702).

N. J. Kalton (2012). "The uniform structure of Banach spaces". *Math. Ann.* 354.4, pp. 1247–1288. MR: 2992997 (cit. on p. 1696).

Nigel Kalton and Albert Wilansky (1976). "Tauberian operators on Banach spaces". *Proc. Amer. Math. Soc.* 57.2, pp. 251–255. MR: 0473896 (cit. on p. 1700).

J. Lindenstrauss and A. Pełczyński (1968). "Absolutely summing operators in L_p-spaces and their applications". *Studia Math.* 29, pp. 275–326. MR: 0231188 (cit. on p. 1696).

J. Lindenstrauss and H. P. Rosenthal (1969). "The $\mathcal{L}_p$ spaces". *Israel J. Math.* 7, pp. 325–349. MR: 0270119 (cit. on p. 1696).

J. Lindenstrauss and L. Tzafriri (1971). "On the complemented subspaces problem". *Israel J. Math.* 9, pp. 263–269. MR: 0276734 (cit. on p. 1703).

– (1976). "The uniform approximation property in Orlicz spaces". *Israel J. Math.* 23.2, pp. 142–155. MR: 0399806 (cit. on p. 1703).

Joram Lindenstrauss (1964). "On nonlinear projections in Banach spaces". *Michigan Math. J.* 11, pp. 263–287. MR: 0167821 (cit. on p. 1696).

Joram Lindenstrauss and Lior Tzafriri (1977). *Classical Banach spaces. I.* Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92. Springer-Verlag, Berlin-New York, pp. xiii+188. MR: 0500056 (cit. on pp. 1691, 1697).

B. Maurey and H. P. Rosenthal (1977). "Normalized weakly null sequence with no unconditional subsequence". *Studia Math.* 61.1, pp. 77–98. MR: 0438091 (cit. on p. 1697).

E. Odell (1992). "On quotients of Banach spaces having shrinking unconditional bases". *Illinois J. Math.* 36.4, pp. 681–695. MR: 1215801 (cit. on p. 1698).

Edward W. Odell and Thomas Schlumprecht (2006). "Embedding into Banach spaces with finite dimensional decompositions". *RACSAM. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat.* 100.1-2, pp. 295–323. MR: 2267413 (cit. on pp. 1698, 1699).

A. Pełczyński (1969). "Universal bases". *Studia Math.* 32, pp. 247–268. MR: 0241954 (cit. on p. 1702).

– (1971). "Any separable Banach space with the bounded approximation property is a complemented subspace of a Banach space with a basis". *Studia Math.* 40, pp. 239–243. MR: 0308753 (cit. on p. 1702).

A. Pełczyński and P. Wojtaszczyk (1971). "Banach spaces with finite dimensional expansions of identity and universal bases of finite dimensional subspaces". *Studia Math.* 40, pp. 91–108. MR: 0313765 (cit. on p. 1697).

Albrecht Pietsch (1978). *Operator ideals*. Vol. 16. Mathematische Monographien [Mathematical Monographs]. VEB Deutscher Verlag der Wissenschaften, Berlin, p. 451. MR: 519680 (cit. on p. 1695).

Gilles Pisier (1988). "Weak Hilbert spaces". *Proc. London Math. Soc. (3)* 56.3, pp. 547–579. MR: 931514 (cit. on p. 1703).

Juan Jorge Schäffer (1976). *Geometry of spheres in normed spaces*. Lecture Notes in Pure and Applied Mathematics, No. 20. Marcel Dekker, Inc., New York-Basel, pp. vi+228. MR: 0467256 (cit. on p. 1692).

Gideon Schechtman (1975). "Examples of $\mathcal{L}_p$ spaces $(1 < p \neq 2 < \infty)$". *Israel J. Math.* 22.2, pp. 138–147. MR: 0390722 (cit. on p. 1695).

Thomas Schlumprecht and András Zsák (2018). "The algebra of bounded linear operators on $\ell_p \oplus \ell_q$ has infinitely many closed ideals". *J. Reine Angew. Math.* 735, pp. 225–247. MR: 3757476 (cit. on p. 1695).

A. Szankowski (1978). "Subspaces without the approximation property". *Israel J. Math.* 30.1-2, pp. 123–129. MR: 508257.

– (2009). "Three-space problems for the approximation property". *J. Eur. Math. Soc. (JEMS)* 11.2, pp. 273–282. MR: 2486934 (cit. on p. 1704).

Andrzej Szankowski (1981). "$B(\mathcal{H})$ does not have the approximation property". *Acta Math.* 147.1-2, pp. 89–108. MR: 631090 (cit. on pp. 1704, 1705).

Matthew Tarbard (2012). "Hereditarily indecomposable, separable $\mathcal{L}_\infty$ Banach spaces with ℓ_1 dual having few but not very few operators". *J. Lond. Math. Soc. (2)* 85.3, pp. 737–764. MR: 2927806 (cit. on p. 1694).

L Weis and M Wolff (1984). "On the essential spectrum of operators on L^1". *Semesterbericht Funk. Anal. Tuebingen*, pp. 103–112 (cit. on p. 1701).

R. J. Whitley (1964). "Strictly singular operators and their conjugates". *Trans. Amer. Math. Soc.* 113, pp. 252–261. MR: 0177302 (cit. on pp. 1693, 1695).

Helmut Wielandt (1949). "Über die Unbeschränktheit der Operatoren der Quantenmechanik". *Math. Ann.* 121, p. 21. MR: 0030701 (cit. on p. 1693).
Aurel Wintner (1947). "The unboundedness of quantum-mechanical matrices". *Physical Rev. (2)* 71, pp. 738–739. MR: 0020724 (cit. on p. 1693).
Bentuo Zheng (2014). "Commutators on $(\Sigma \ell_q)_{\ell_1}$". *J. Math. Anal. Appl.* 413.1, pp. 284–290. MR: 3153585 (cit. on p. 1694).

Received 2017-12-01.

William B. Johnson
Department of Mathematics
Texas A&M University
College Station, TX 77843
USA
johnson@math.tamu.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1709–1730)

THE EFFECT OF DISORDER AND IRREGULARITIES ON SOLUTIONS TO BOUNDARY VALUE PROBLEMS AND SPECTRA OF DIFFERENTIAL OPERATORS

Svitlana Mayboroda (Світлана Майборода)

Abstract

This note describes the impact of disorder or irregularities in the ambient medium on the behavior of stationary solutions to elliptic partial differential equations and on spatial distribution of eigenfunctions, as well as the profound and somewhat surprising connections between these two topics which have been revealed in the past few years.

1 Introduction

Irregularities in boundary value problems appear in different forms: for example, in the non-smooth boundary of the ambient domain, in a lack of sufficient geometric structure or connectedness, and in disordered potentials and other features of the rough media. It is well known that such irregularities may have a drastic impact on the solutions of the equations or eigenfunctions of the underlying operators and completely change their behavior from what is known to occur in a smooth uniform setting.

A survey of the entirety of the circle of such phenomena would be impossible. We concentrate here on two recent research directions in this area and the somewhat unexpected connections between them that have become evident in the past few years. We shall demonstrate that in many situations, different types of irregularities may be treated using the same mechanisms.

The first direction pertains to fundamental questions about regularity of solutions to elliptic equations in arbitrary domains. In the context of the Laplacian, these questions were settled using the maximum principle, which is a simple consequence of the mean value property, and with the Wiener criterion proved in 1924. However, even the simplest higher order generalization of the Laplacian, the polyharmonic operator $(-\Delta)^m$, $m \in \mathbb{N}$,

Mayboroda is supported in part by the Simons Fellowship, the NSF INSPIRE Award DMS 1344235, and the NSF CAREER Award DMS 1220089.

MSC2010: primary 35J40; secondary 35J10, 35P15, 35P99.

has remained resistant to known techniques. The first part of this note is based on a series of papers from 2009–2017 which pioneered a new method of weighted integral inequalities and achieved a complete description of the boundary regularity of polyharmonic functions in arbitrary domains, including sharp dimensional restrictions on boundedness of the derivatives of the solution, and also geometric conditions on the domain necessary and sufficient for continuity of the derivatives, an analogue of the Wiener test.

The second part of this paper focuses on the global spatial distribution of waves, or eigenfunctions of elliptic operators, and on the perplexing effect of wave localization – a confinement of wave propagation to a small portion of the original domain, triggered by irregularities of the boundary or of the coefficients of the underlying operator. Despite the profoundly interferential nature of this phenomenon, there turns out to be a magical stationary solution to an elliptic boundary value problem which governs both spatial distribution of eigenfunctions and the location of eigenvalues near the bottom of the spectrum, and connects certain wave propagation mechanisms to the properties of individual solutions of elliptic equations. Curiously, one of the main components of this new approach involves once again some intricate weighted integral inequalities, albeit different from those used for the higher order Wiener test.

2 Regularity of solutions to higher order equations

Most of the material in this Section is based on the work of Mayboroda and V. Maz'ya [2009, 2014, 2018]. However, there is an extended historical discussion, and, while the restrictions of the format of this survey make putting local references rather difficult, we will list the links to the corresponding literature in the end of the Section. The reader can find a complete version of this paper on ArXiv.

2.1 Regularity of polyharmonic functions in general domains. Elliptic operators of order greater than two arise in many areas of mathematics, including conformal geometry (the Paneitz operator, Q-curvature), free boundary problems, and non-linear elasticity. They have fundamental applications in physics and engineering, ranging from standard models of elasticity to cutting-edge research on Bose-Einstein condensation in graphene, and have enjoyed persistent attention from physicists and mathematicians alike. However, the properties of higher order PDEs on general domains remained largely beyond reach.

The prototypical example of a higher-order elliptic operator, well known from the theory of elasticity, is the bilaplacian $\Delta^2 = \Delta(\Delta)$; a more general example is the polyharmonic operator $(-\Delta)^m$, $m \geq 2$. Already for these model operators, known results defied intuition and classical mathematical approaches, as the XXth century has slowly revealed a sequence of fascinating counterexamples. For instance, Hadamard's 1908 conjecture regarding positivity of the biharmonic Green's function was refuted in 1949. The bilaplacian, which is commonly used to model the deflection of a clamped plate, has a Green's function which can change sign in an elongated smooth convex domain, and actually oscillates near a corner of a rectangle. Szegö's 1950 conjecture about the positivity of the fundamental eigenfunction met a similar counterexample in 1982. Later on, it was shown that the weak maximum principle may fail as well, at least in high dimensions. One could also mention the Babuška paradox: the solution with "hard" supported plate conditions on a smooth domain cannot be approximated by solutions in polyhedra. This makes straightforward finite element methods intrinsically inapplicable.

The positive results of this nature for higher order PDEs have been essentially restricted to three main settings: smooth domains, domains with isolated singularities, and domains with a Lipschitz boundary. Unfortunately, none of the emerging methods and techniques could be extended to domains with more complicated geometry.

To see the extent of this shortfall, let us recall the second order case. A fundamental result of elliptic theory is the maximum principle for harmonic functions. It holds in arbitrary domains and guarantees that every solution to the Dirichlet problem for Laplace's equation achieves its maximum on the boundary and, in particular, that a solution with bounded data is bounded. A similar statement extends to all divergence form second order elliptic operators.

The situation for higher order PDEs is different. On smooth domains the study of higher order differential equations went hand-in-hand with the second order theory and a weak version of the maximum principle was established in 1960. Roughly speaking, if u is a solution to a uniformly elliptic differential equation of order $2m$ with smooth coefficients on a smooth domain Ω, the weak maximum principle gives the estimate

$$\max_{0\leq k\leq m-1}\left\|\nabla^k u\right\|_{L^\infty(\overline{\Omega})} \leq C \max_{0\leq k\leq m-1}\left\|\nabla^k u\right\|_{L^\infty(\partial\Omega)}, \tag{2-1}$$

where $\nabla^k u = \{\partial^\alpha u\}_{|\alpha|=k}$ is a vector of all partial derivatives of u of order k and we adopt the usual convention that the derivative of order zero is u itself. In the early 1990s, (2-1) was extended to three-dimensional domains diffeomorphic to polyhedra or with Lipschitz boundary. The formulation in non-smooth domains could be slightly different from above. However, no matter the setting, the weak maximum principle always guarantees that a solution with "nice" data has bounded derivatives of order $m-1$.

In striking contrast with the case of harmonic functions, for every elliptic operator of order greater than two there exists a cone in $\mathbb{R}^n$, $n\geq 4$, in which the maximum principle is violated. In particular, there are variational solutions to the polyharmonic equation in certain domains Ω in dimensions $n\geq 4$, with bounded Dirichlet data, such that $\nabla^{m-1}u$ is unbounded in Ω.

These positive and negative results raise a number of fundamental questions. Can (2-1) be extended to arbitrary domains in dimension 3, on par with the maximum principle for the Laplacian? Can one establish results of similar generality and strength in higher dimensions, for lower order derivatives? And finally, if any of these answers is positive, might it be possible to find the capacitory conditions which govern continuity of the appropriate derivatives?

The main results of Mayboroda and V. Maz'ya [2009, 2014] establish sharp pointwise estimates on variational solutions to the polyharmonic equation and their derivatives in arbitrary bounded open sets, without any restrictions on the geometry of the underlying domain:

Theorem 2-2. *Let Ω be a bounded domain in $\mathbb{R}^n$, $2 \le n \le 2m+1$, and*

$$(-\Delta)^m u = f \text{ in } \Omega, \quad f \in C_0^\infty(\Omega), \quad u \in \mathring{W}^{m,2}(\Omega). \tag{2-3}$$

Then

$$\nabla^{m-n/2+1/2} u \in L^\infty(\Omega) \text{ when } n \text{ is odd} \quad \text{and} \quad \nabla^{m-n/2} u \in L^\infty(\Omega) \text{ when } n \text{ is even}. \tag{2-4}$$

In particular,

$$\nabla^{m-1} u \in L^\infty(\Omega) \text{ when } n = 2, 3. \tag{2-5}$$

Here the space $\mathring{W}^{m,2}(\Omega)$ is the completion of $C_0^\infty(\Omega)$ with respect to the norm $\|u\|_{\mathring{W}^{m,2}(\Omega)} = \|\nabla^m u\|_{L^2(\Omega)}$, so (2-3) is a weak form of the Dirichlet boundary condition which in smooth domains corresponds to $\nabla^k u = 0$ on $\partial\Omega$, $k = 0, .., m-1$, and $(-\Delta)^m$ is interpreted in the usual weak sense. We note that $\mathring{W}^{m,2}(\Omega)$ embeds into $C^k(\Omega)$ for $k < m - \frac{n}{2}$, $n < 2m$. Thus Theorem 2-2 gains one classical derivative over the outcome of Sobolev embedding.

The results of Theorem 2-2 cannot be improved, for in general domains solutions need not exhibit smoothness higher than that given in (2-4)–(2-5). This fact is a straightforward consequence of the Wiener-type test that we will discuss in Section 2.3, but let us nonetheless show some simpler counterexamples.

When $n \in [3, 2m-1]$ is odd, the main example reduces simply to a punctured ball in $\mathbb{R}^n$. Indeed, let $\Omega = B_1 \setminus \{O\}$, where $B_r = \{x \in \mathbb{R}^n : |x| < r\}$ and O stands for the origin. Suppose $\eta \in C_0^\infty(B_{1/2})$ is such that $\eta = 1$ on $B_{1/4}$ and let

$$u(x) := \eta(x)\, \partial_i^{m-\frac{n}{2}-\frac{1}{2}}(|x|^{2m-n}), \qquad x \in B_1 \setminus \{O\}, \tag{2-6}$$

where ∂_i stands for the derivative in the direction of x_i for some $i = 1, ..., n$. Then $u \in \mathring{W}^{m,2}(\Omega)$ and $(-\Delta)^m u \in C_0^\infty(\Omega)$. Furthermore, while $\nabla^{m-\frac{n}{2}+\frac{1}{2}} u$ is bounded, the vector of higher order classical derivatives, $\nabla^{m-\frac{n}{2}+\frac{3}{2}}u$, is not, and in fact $\nabla^{m-\frac{n}{2}+\frac{1}{2}}u$ is not continuous at the origin. Therefore, estimates (2-4)–(2-5) are optimal.

When $n = 2m+1$, the sharpness of (2-4)–(2-5) follows from a possible lack of continuity of a solution that will be proved with the capacitary estimates in Section 2.3.

When $n \in [2, 2m]$ is even, one can use previously available results on the asymptotic character of a solution in the complement of a ray, and show that there is an m-harmonic function behaving like $|x|^{m-\frac{n}{2}+\frac{1}{2}}$. Thus, truncating by the same cut-off η as above, one obtains a solution u to (2-3) in $B_1 \setminus \{x_1 = 0, \ldots, x_{n-1} = 0, 0 \le x_n < 1\}$ with the property that $\nabla^{m-\frac{n}{2}+1}u$ is unbounded. Therefore, (2-4) need not hold for higher order classical derivatives. In Section 2.3 we exhibit refined counterexamples confirming sharpness of our results at the precise level of fractional derivatives and even confirming that $\nabla^{m-\frac{n}{2}}u$ need not be continuous when n is even.

2.2 Estimates for the Green's function. Theorem 2-2 has a quantitative aspect as well. It has several possible manifestations, including monotonicity-type formulas that will be discussed in Section 2.4. Here we concentrate on the kernel of the representation formulas, the Green's function of the polyharmonic operator. Much like the maximum principle, pointwise estimates on the Green's function lie at the foundations of elliptic theory. In the second order case the basic bound $G(x,y) \le C\,|x-y|^{2-n}$, $x \ne y \in \Omega$, is indeed a straightforward consequence of the maximum principle. As for the higher order PDEs, already the counterexamples to Hadamard's conjecture that we have discussed above show that the biharmonic Green's function is an intricate object, with a very peculiar behavior at singular points of the boundary. Nonetheless, the estimates behind the main results in Section 2.1 yield roughly the same order-of-magnitude bounds on the derivatives of a polyharmonic Green's function as one would expect in a regular setting, of course with the same restrictions on the maximal order of differentiability as in Theorem 2-2. The list of all such estimates is extensive, so we highlight the highest order case.

To this end, let $\Omega \subset \mathbb{R}^n$ be an arbitrary bounded domain with $n \in [3, 2m+1]$ odd. We denote by G the Green's function for the polyharmonic equation, i.e., a solution to $(-\Delta_x)^m G(x,y) = \delta(x-y)$, $x \in \Omega$, in $\mathring{W}^{m,2}(\Omega)$, and by Γ the fundamental solution in $\mathbb{R}^n$. The difference $G(x,y) - \Gamma(x-y)$ is the regular part of the Green's function. Here Δ_x denotes the Laplacian in x, and similarly, ∇_x, ∇_y denote the gradient in x, and in y, respectively. The function $d(x)$ is the distance from $x \in \Omega$ to $\partial\Omega$.

The theorems in the work of Mayboroda and V. Maz'ya [2014] include the following estimates. If $n \in [3, 2m+1]$ is odd, then

$$\left|\nabla_x^{m-\frac{n}{2}+\frac{1}{2}} \nabla_y^{m-\frac{n}{2}+\frac{1}{2}} \big(G(x,y) - \Gamma(x-y)\big)\right| \le \frac{C}{\max\{d(x), d(y), |x-y|\}}, \tag{2-7}$$

and

$$\left|\nabla_x^{m-\frac{n}{2}+\frac{1}{2}} \nabla_y^{m-\frac{n}{2}+\frac{1}{2}} G(x,y)\right| \leq \frac{C}{|x-y|}, \tag{2-8}$$

for every $x, y \in \Omega$. The constant C here depends on the dimension and the order of the operator, but not on the size or geometry of Ω. Similar estimates have been obtained in even dimensions, with mixed derivatives of the Green's function of order $2m - n$ on the left-hand side and appropriate logarithmic terms on the right-hand side.

Much as for general solutions, previous results have been available in smooth domains, in conical domains, and in polyhedra. In addition, for arbitrary domains, pointwise bounds on the Green's function (rather than its derivatives) in dimensions $2m+1$ and $2m+2$ for $m > 2$ and dimensions $5, 6, 7$ for $m = 2$ can be found in existing literature (see Section 2.5).

Estimates (2-7) and (2-8) for the derivatives of the Green's function have been treated by Mayboroda and V. Maz'ya [2014] for the first time and are essentially the best possible. The higher derivatives may be unbounded, due to the counterexamples in Section 2.1.

2.3 The Wiener test. As we discussed above, the maximum principle for harmonic functions, guaranteeing that solutions with bounded data are always bounded, is a beautifully simple result that requires no assumptions on the boundary of the domain and enjoys a fairly straightforward proof. By contrast, even for the Laplacian, the question of continuity of solutions at the boundary stood out as an important, highly non-trivial open problem in the beginning of the XXth century. The results of Poincaré, Zaremba and Lebesgue showed that harmonic functions are always continuous at the vertex of a cone, while in the complement of a sufficiently thin cusp this property may fail. In 1924 Wiener introduced the harmonic capacity, which now bears his name, and established his famous criterion for regularity of a boundary point. Wiener's interest in the problem was primarily guided by a physical example, the question of breakdown of an electrostatic field at the tip of a lightning rod, and the now famous Wiener capacity was an extension of the physical notion of a capacitor in electrostatics. However, over the century the Wiener criterion has made a profound impact in many branches of mathematics. Most notably, the notion of capacity gave a non-linear analogue to the Lebesgue measure, suitable for measuring exceptional sets in Sobolev spaces much as the Lebesgue measure does in L^p, and seamlessly entered probability, potential theory, and function space theory.

To formulate the Wiener criterion, we recall that a point $O \in \partial\Omega$ is referred to as regular if every solution to the Dirichlet problem for the Laplacian, with continuous data, is continuous at O. The Wiener criterion states that $O \in \partial\Omega$ is a regular point if and only if

$$\sum_{j=0}^{\infty} 2^{j(n-2)} \operatorname{cap}\left(\overline{B_{2^{-j}}} \setminus \Omega\right) = +\infty, \tag{2-9}$$

where $B_{2^{-j}}$ is the ball of radius 2^{-j} centered at O and the harmonic capacity of a compactum $K \subset \mathbb{R}^n$, $n \geq 3$, is defined as

$$\operatorname{cap}(K) := \inf\Big\{ \|\nabla u\|^2_{L^2(\mathbb{R}^n)} : \ u \in C_0^\infty(\mathbb{R}^n),\ u = 1 \text{ in a neighborhood of } K \Big\}. \tag{2-10}$$

In other words, harmonic functions are always continuous at the boundary point O if and only if the complement of the domain near the point O, measured in terms of the Wiener capacity, is sufficiently massive. An appropriately modified version of this condition is also available in dimension $n = 2$.

We remark that we have not specified which concept of solution to the Dirichlet problem is used in (2-9)–(2-10). Below we will continue working with an equivalent reformulation in terms of (2-3), although it was not Wiener's definition, but it is instructive to point out that a part of the appeal and the original goal of the Wiener criterion was to separate the issues of definition and boundary regularity of solutions and to give a comprehensive description of situations when a solution is "classical", that is, continuously achieves its boundary data.

The Wiener test has been extended to a large variety of second order differential equations, including general divergence form elliptic equations, degenerate and parabolic PDEs, the Schrödinger operator, and also celebrated notoriously difficult quasilinear and nonlinear problems. Nonetheless, none of the previous methods could handle the Wiener criterion or Wiener capacity in a higher order context, and moreover, as we discussed above, even modest analogues of the maximum principle remained an open problem impeding any further progress. Indeed, to address a Wiener-type criterion for continuity of higher order derivatives, one must establish their boundedness first. Theorem 2-2 establishes the exact order of smoothness for polyharmonic functions on domains with no geometrical restrictions and sets the stage for a discussion of the Wiener test for continuity of the corresponding derivatives of the solution. The first obstacle in carrying this out is the lack of a definition of an appropriate capacity since (2-10) is tailored to second order equations and, moreover, the "boundary data" $u = 1$ on the compactum K in (2-10) is geared to understanding the continuity of u rather than its derivatives at the boundary. The following solution has been presented in the work of Mayboroda and V. Maz'ya [2018].

Let $n \in [2, 2m+1]$ and denote by Z the following set of indices:

$$Z = \{0, 1, \ldots, m - n/2 + 1/2\}, \text{ if } n \text{ is odd}, \tag{2-11}$$

$$Z = \{m - n/2,\ m - n/2 - 2,\ m - n/2 - 4, \dots\} \cap (\mathbb{N} \cup \{0\}),\ \text{if } n \text{ is even.} \tag{2-12}$$

Let Π be the space of linear combinations of spherical harmonics

$$P(x) = \sum_{p \in Z} \sum_{l} b_{pl} Y_l^p(x/|x|), \qquad b_{pl} \in \mathbb{R}, \quad x \in \mathbb{R}^n \setminus \{O\}, \tag{2-13}$$

where l is an integer parameter running from 1 to $N(p,n)$, the number of linearly independent homogeneous harmonic polynomials of degree p in n variables. The space Π is equipped with the norm $\|P\|_\Pi$ given by the ℓ^2 norm of the coefficients $\{b_{pl}\}_{p,l}$; we also set $\Pi_1 := \{P \in \Pi : \|P\|_\Pi = 1\}$. Given $P \in \Pi_1$, an open set D in $\mathbb{R}^n$ such that $O \in \mathbb{R}^n \setminus D$, and a compactum K in D, we define

$$\operatorname{Cap}_P(K, D) := \inf\left\{ \int_D |\nabla^m u(x)|^2\, dx :\ u \in \mathring{W}^{m,2}(D),\ u = P \text{ in a neigh. of } K \right\}, \tag{2-14}$$

and

$$\operatorname{Cap}(K, D) := \inf_{P \in \Pi_1} \operatorname{Cap}_P(K, D). \tag{2-15}$$

Since we will be primarily working in dyadic annuli $C_{2^{-j},2^{-j+2}}$, $j \in \mathbb{N}$, where $C_{s,as} := \{x \in \mathbb{R}^n : s < |x| < as\}$, $s, a > 0$, it is convenient to abbreviate the notation and drop reference to the "ambient" set, thus writing

$$\operatorname{Cap}_P(\overline{C_{2^{-j},2^{-j+2}}} \setminus \Omega) := \operatorname{Cap}_P(\overline{C_{2^{-j},2^{-j+2}}} \setminus \Omega, C_{2^{-j-2},2^{-j+4}}), \quad j \in \mathbb{N}, \tag{2-16}$$

and similarly for Cap.

Given a domain $\Omega \subset \mathbb{R}^n$, $n \geq 2$, the point $Q \in \partial\Omega$ is called *k-regular* with respect to Ω and the operator $(-\Delta)^m$, $m \in \mathbb{N}$, if any weak solution to the Dirichlet boundary problem for the m-Laplacian (2-3) satisfies the condition

$$\nabla^k u(x) \to 0 \text{ as } x \to Q,\ x \in \Omega, \tag{2-17}$$

that is, all partial derivatives of u of order k are continuous at Q. Notice that, in the spirit of our discussion for the Laplacian, the case $k = m-1$ addresses the "classical" solutions, that is, the case when a weak solution achieves its Dirichlet boundary data continuously.

Theorem 2-18 (Mayboroda and V. Maz'ya [2009, 2018])**.** *Let Ω be an arbitrary open set in $\mathbb{R}^n$, $m \in \mathbb{N}$, $2 \leq n \leq 2m+1$. Assume first that n is odd. If*

$$\sum_{j=0}^{\infty} 2^{-j(2m-n)} \inf_{P \in \Pi_1} \operatorname{Cap}_P(\overline{C_{2^{-j},2^{-j+2}}} \setminus \Omega) = +\infty, \tag{2-19}$$

then the point O is $(m-n/2+1/2)$-regular with respect to the domain Ω and the operator $(-\Delta)^m$. Conversely, if the point $O \in \partial\Omega$ is $(m-n/2+1/2)$-regular then

$$\inf_{P\in\Pi_1} \sum_{j=0}^{\infty} 2^{-j(2m-n)} \operatorname{Cap}_P\left(\overline{C_{2^{-j},2^{-j+2}}} \setminus \Omega\right) = +\infty. \tag{2-20}$$

Assume next that n is even. If the condition (2-19) holds with $j\,2^{-j(2m-n)}$ in place of $2^{-j(2m-n)}$, then the point O is $(m-n/2)$-regular with respect to the domain Ω and the operator $(-\Delta)^m$. Conversely, if the point $O \in \partial\Omega$ is $(m-n/2)$-regular then the condition (2-20) is satisfied with $j\,2^{-j(2m-n)}$ in place of $2^{-j(2m-n)}$.

A byproduct of Theorem 2-18 is a strengthened version of the Green's function bounds in Section 2.2. For instance, for $m=2$ and $n=3$, inequality (2-8) can be strengthened to

$$|\nabla_x\nabla_y G(x,y)| \le \frac{C}{|x-y|} \times \exp\Big(-c\sum_{j=2}^{l_{yx}} 2^{4j}|y| \operatorname{Cap}\left(\overline{C_{2^{4j+1}|y|,2^{4j+5}|y|}} \setminus \Omega\right)\Big),$$

where $|y| \le c|x|$ and $l_{yx} \ge 2$, $l_{yx} \in \mathbb{N}$, is such that $|x| \ge 2^{4l_{yx}+5}|y|$. In particular, this yields Hölder continuity of the Green's function in a class of domains satisfying uniform capacity conditions.

Turning back to Theorem 2-18, we note that this result was the first treatment of necessary and sufficient conditions for boundary continuity of derivatives of an elliptic equation of order $2m>2$ and the first time the capacity (2-14) for $m>2$ appeared in literature. Continuity of the solution itself (but not of its derivatives) was previously treated for the polyharmonic operator, and the resulting criterion follows from Theorem 2-18. In particular, when $n=2m$, the new notion of capacity (2-15) coincides with the potential-theoretical Bessel capacity. When applied to the case $m=1$, $n=3$, Theorem 2-18 yields the classical Wiener criterion (2-9)–(2-10). In fact, for $n=2m$ or $n=2m+1$ the necessary and sufficient conditions in Theorem 2-18 are trivially the same, as $P\equiv 1$. In lower dimensions, the discrepancy is not artificial: for example, (2-19) may fail to be necessary. One could furthermore treat the aforementioned results in Lipschitz and smooth domains as sufficient conditions for the continuity of the corresponding derivatives of polyharmonic functions. It is not difficult to verify that they fall into the scope of Theorem 2-18 as well since the capacity of a cone, and hence the capacity of an intersection with a complement of a Lipschitz domains, is large enough to assure divergence of the series in (2-19).

Finally, let us come back to the question of sharpness of Theorem 2-2. Resting on Theorem 2-18 and choosing sufficiently small balls in the consecutive annuli as a complement to the domain, we can build a set with a convergent capacitory integral and, respectively, an irregular solution with discontinuous derivatives of order $m - n/2$ or $m - n/2 + 1/2$ at the point O in even and odd dimensions, respectively. This provides additional, albeit somewhat more technical, counterexamples, in concert with the discussion at the end of Section 2.1.

At the heart of Theorem 2-18, much as at the heart of the original Wiener criterion, lies the challenge of finding a correct notion of polyharmonic capacity and of understanding its key properties. While seemingly exotic, a careful choice of linear combinations of spherical harmonics (see (2-11)–(2-12) and (2-13)), dependent on the parity of m and n, is crucial at several stages of the proof. Needless to say, it necessitates quite different arguments than the original Wiener criterion, and moreover, this new capacity and the notion of higher-order regularity sometimes exhibits unexpected properties. To mention only a few examples, we recall that contrary to the classical theory 1-regularity of a boundary point for the bilaplacian may be unstable under the affine changes of coordinates. Perhaps even more surprising is the fact that in sharp contrast to the second order case, the same geometric conditions are not responsible for regularity of solutions to all higher order equations.

The latter point is important and we elaborate further. The Wiener criterion for second order partial differential equations is exceptionally versatile. A boundary point is 0-regular for the Laplacian (any weak solution is continuous) if and only if it is 0-regular for any other elliptic operator $-\operatorname{div} A\nabla$ with real bounded measurable coefficients. One might thus conjecture that Theorems 2-2 and 2-18 hold for general higher order elliptic operators. However, this is false, even for constant coefficient operators. For instance, a solution to the biharmonic equation $(-\Delta)^2 u = 0$ is continuous at the vertex of a cone in any dimension, while for $[(-\Delta^2) + a(\partial/\partial x_n)^4]u = 0$ in dimensions $n \geq 8$, the solution need not bounded, much less continuous, for any $a > 0$ such that $(n-3)\arctan\sqrt{a} \in (2\pi, 4\pi)$, e.g., for $a > 5+2\sqrt{5}$ when $n = 8$. On the other hand, for $n \geq 2m$, the higher order Wiener criterion for 0-regularity of a boundary point has been extended to constant coefficient operators satisfying a certain weighted integral positivity property. Such properties are explained in the next section, but we remark now that when $n = 2m$, this weighted integral positivity property (and hence, the Wiener criterion) holds for every higher order elliptic equation with real constant coefficients. This leads to a number of interesting questions.

Another important direction of future research is to understand the new capacity for further common examples. For instance, suppose $m = 2$, $n = 3$, and the domain Ω has an inner cusp, so that in a neighborhood of the origin, $\Omega = \{(r,\theta,\phi) : 0 < r < c,\ h(r) < \theta \leq \pi,\ 0 \leq \phi < 2\pi\}$, where h is a non-decreasing function such that $h(br) \leq h(r)$ for some $b > 1$. For such a domain Theorem 2-18 yields the following criterion:

$$\text{(2-21)} \qquad \text{the point } O \text{ is 1-regular} \quad \text{if and only if} \quad \int_0^1 s^{-1} h(s)^2\, ds = +\infty.$$

Overall, our understanding of higher order capacity is far from complete.

2.4 Weighted integral inequalities. A technical core of the proof of Theorem 2-2, and an important part of the work of Mayboroda and V. Maz'ya [2009, 2014, 2018], is a new method which builds on some intricate weighted integral identities. It does not require any a priori information on the geometry of the domain and allows one to estimate the growth of a solution in spherical sections centered at a given boundary point, much as monotonicity formulas do for nice second order operators. The main difficulty is the proper choice of the weight function w, which is finely tuned to the underlying elliptic operator in such a way that suitable lower bounds can be derived from $\int_\Omega (-\Delta)^m u\, u\, w dx$, which we write as $\langle Lu, u\, w\rangle$, $L = (-\Delta)^m$.

To give a sense of the key estimates, let us concentrate on the technically simpler case when the dimension is odd. We assume, as before, that $m \in \mathbb{N}$, $n \in [3, 2m+1]$ (odd for the moment), and Ω is a bounded domain in $\mathbb{R}^n$, $O \in \mathbb{R}^n \setminus \Omega$, and $u \in C_0^\infty(\Omega)$. Our key tool is the weighted integral inequality

$$C\,\frac{1}{|\xi|^{n-1}} \int_{x\in\Omega:\,|x|=|\xi|} \Big(\frac{u(x)}{|x|^{m-\frac{n}{2}+\frac{1}{2}}}\Big)^2 d\sigma_x \le \langle Lu, u\, w_\xi\rangle, \tag{2-22}$$

for every $u \in C_0^\infty(\Omega)$, $\xi \in \Omega$, and

$$w_\xi(x) = |x|^{-1}\Big(C_1 h\Big(\log\frac{|\xi|}{|x|}\Big) + C_2\Big), \qquad x, \xi \in \Omega. \tag{2-23}$$

Here C, C_1, C_2 are constants depending on m and n only, and

$$h(t) = \begin{cases} \sum_{j=1}^m \nu_j\, e^{-\alpha_j t}, & t > 0, \\ \sum_{j=1}^m \mu_j\, e^{\beta_j t}, & t < 0. \end{cases} \tag{2-24}$$

The constants $\alpha_j > 0$, $j = 1, 2, \dots, m$, $\beta_1 = 0$, and $\beta_j > 0$ for $j = 2, \dots, m$ are given by

$$\{-\alpha_j\}_{j=1}^m \bigcup \{\beta_j\}_{j=1}^m = \Big\{-m + \frac{n}{2} - \frac{1}{2} + 2j\Big\}_{j=0}^{m-1} \cup \Big\{-\frac{n}{2} - \frac{1}{2} + m - 2j\Big\}_{j=0}^{m-1}, \tag{2-25}$$

and writing $\vec{\gamma} = (-\alpha_1, \dots, -\alpha_m, \beta_1, \dots, \beta_m)$, $\vec{\kappa} = (\nu_1, \dots, \nu_m, -\mu_1, \dots, -\mu_m)$, the coefficients $\nu_j, \mu_j \in \mathbb{R}$ satisfy $\kappa_i = (-1)^{m+1} \prod_{j\neq i} (\gamma_j - \gamma_i)^{-1}$.

The expression on the left in (2-22) can be written more concisely, but the form here emphasizes its key features. This inequality is a statement about uniform control of spherical averages of $\frac{u(x)}{|x|^{m-\frac{n}{2}+\frac{1}{2}}}$ for solutions in a neighborhood of a boundary point $O \in \partial\Omega$.

By interior regularity estimates, it also yields uniform control on $\nabla^{m-\frac{n}{2}+\frac{1}{2}}u$, which is our goal. Indeed, using (2-22) we prove that for any $Q \in \mathbb{R}^n \setminus \Omega$ a solution to

$$(2\text{-}26) \qquad (-\Delta)^m u = f \text{ in } \Omega, \quad f \in C_0^\infty(\Omega \setminus B_{4R}(Q)), \quad u \in \mathring{W}^{m,2}(\Omega), \quad R > 0,$$

satisfies the monotonicity-type formula

(2-27)
$$\frac{1}{\rho^{2m}} \int_{S_\rho(Q)\cap\Omega} |u(x)|^2 \, d\sigma_x \le \frac{C}{R^{2m+1}} \int_{C_{R,4R}(Q)\cap\Omega} |u(x)|^2 \, dx \quad \text{for every} \quad \rho < R,$$

where C is a constant depending only on m and n, and $S_\rho(Q) = \{x \in \mathbb{R}^n : |x-Q| = \rho\}$, $B_\rho(Q) = \{x \in \mathbb{R}^n : |x-Q| < \rho\}$, and $C_{R,4R}(Q) = \{x \in \mathbb{R}^n : R < |x-Q| < 4R\}$. Thus solutions do not grow faster than $|x-Q|^{m-\frac{n}{2}+\frac{1}{2}}$ near $Q \in \partial\Omega$, and

(2-28)
$$|\nabla^i u(x)|^2 \le C \, \frac{|x-Q|^{2m-n+1-2i}}{R^{2m+1}} \int_{C_{R/4,4R}(Q)\cap\Omega} |u(y)|^2 \, dy, \quad 0 \le i \le m - \frac{n}{2} + \frac{1}{2},$$

for every $x \in B_{R/4}(Q) \cap \Omega$. In particular, for every bounded domain $\Omega \subset \mathbb{R}^n$, the solution to the boundary value problem (2-26) satisfies $|\nabla^{m-n/2+1/2}u| \in L^\infty(\Omega)$.

The choice of the weight function w_ξ in (2-22) is dictated by several considerations. It must be a fundamental solution of a suitable ordinary differential equation so that the integration by parts leading from the right side of (2-22) to a weighted quadratic form produces a delta function. This restricts the corresponding portion of the integral to the sphere $|x| = |\xi|$, which yields the term on the left in (2-22). In addition, w_ξ must also be chosen so that all the terms of the resulting quadratic form which do not contribute to the aforementioned integral on the sphere give a positive quantity. For the bilaplacian on $\mathbb{R}^3$the computation is explicit, but appears somewhat magical, and the difficulties of handling general m and n seemed at first insurmountable. Our most recent work introduces a novel systematic construction of the weight leading to (2-22), which uses, in particular, an induction in the eigenvalues of the Laplace-Beltrami operator on the sphere, preservation of some positivity properties under a change of underlying higher order operator, and the exploitation of delicate features of $(-\Delta)^m$ which depend on the parity of $m, n, m-n/2$. While the role of parity of m and $m-n/2$ is somewhat peculiar, one can imagine a difference between the analysis in odd and even dimensions in parallel with the Laplacian, for which the fundamental solution in $\mathbb{R}^2$ is logarithmic, while in $\mathbb{R}^3$ it is a power function.

2.5 Historical references. The counterexamples to positivity and local regularity of solutions to higher order equations mentioned in Subsection 2.1 are due to Duffin [1949], Coffman [1982], Babuška [1963], V. G. Maz'ya and Rossmann [1992], Pipher and Verchota [1995]. The results on smooth domains discussed in connection with the maximum principle can be found in the work of Agmon [1960], while for domains with isolated singularities we refer to Kozlov, V. G. Maz'ya, and Rossmann [2001], V. G. Maz'ya and Rossmann [1991], and for domains with a Lipschitz boundary we refer to Pipher and Verchota [1995].

Estimates for the Green function of the general second order elliptic operators, as well as the corresponding Wiener criterion can be found in the work of Littman, Stampacchia, and Weinberger [1963]. The original Wiener criterion for the Laplacian is due to Wiener [1924]. Furthermore, pointwise estimates for solutions to the higher order equations (rather than derivatives of the solutions) and the corresponding Wiener criterion have been obtained in the work of V. G. Maz'ya [1999, 2002] and V. Maz'ya [1997]. In these papers one can also find some of the aforementioned examples.

3 Wave localization

This section is devoted to the work of author and her collaborators on localization of eigenfunctions. The flagship references are Filoche and Mayboroda [2009, 2012], and Arnold, David, Jerison, Mayboroda, and Filoche [2018b, 2016], and some others will be mentioned throughout the text. Much as before, the historical references will placed in the end of the Section. The complete version of this paper can be found on ArXiv.

3.1 Dirichlet problem and the birth of the landscape function. The regularity results for solutions of the Dirichlet problem described in Section 2 have yielded far-reaching implications well beyond their original scope. In particular, some aspects of the emerging intuition are at the heart of the first ideas in a new approach to wave localization.

To illustrate this at a rudimentary level, consider the eigenfunctions φ of the bilaplacian. These satisfy $\Delta^2\varphi = \lambda\varphi$ in a bounded domain Ω with the Dirichlet boundary condition, $\varphi \in \mathring{W}^{2,2}(\Omega)$ (cf. (2-3)). Unlike the case of the Laplacian, explicit solutions of this eigenvalue problem, when Ω is a rectangle $\mathcal{R} = (a,b) \times (c,d) \subset \mathbb{R}^2$, are not available (recall Szegö's conjecture discussed in Section 2.1). However, known formulas and numerical experiments confirm that eigenfunctions in $\mathcal{R}$ still behave at large like linear combinations of products of sines and cosines and therefore are uniformly distributed throughout $\mathcal{R}$.

It turned out that in a *punctured* rectangle, $\mathcal{R} \setminus P$, where P is a point in $\mathcal{R}$, all eigenfunctions at the bottom of the spectrum may be localized. For instance, if $\mathcal{R} = (0, \sqrt{20}) \times (0, 1/\sqrt{20})$ and the point P has coordinates $(1/5, 1/2)$, then essentially the first 200 eigenfunctions of $\mathcal{R} \setminus P$ stay strictly on the right or on the left of the clamped point (see Figure 1). In other words, one nail strategically placed in a rectangular bench could virtually bring to a halt the transfer of energy and induce severe wave localization.

Inspired both by the positive and negative results described in Section 2, Filoche and Mayboroda [2009] discovered a completely different behavior in domains infinitesimally close to $\mathfrak{R}$.

But what is wave localization? It is an astonishing ability of physical systems to maintain vibrations in small portions of their original domains of activity, preventing extended propagation. In this context, one should not think solely in terms of mechanical vibrations. Light is a particular example of an electromagnetic wave, wifi is delivered by waves, sound is a pressure wave, and, from the vantage point of quantum physics, even matter can be perceived as a type of wave. Localization of waves plays a paramount role in each of its physical manifestations. However, many aspects of this phenomenon remain a mystery.

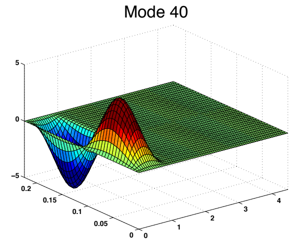


Figure 1: An eigenfunction of the bilaplacian in a rectangular plate with one interior clamped point.

There are some simple and well-understood examples. For instance, in a dumbbell domain given by two non-identical balls connected by a long thin passage, there are Dirichlet eigenfunctions of the Laplacian which are essentially confined to one of the balls and for all practical purposes vanish in the other, visually dying off along the thin passage.

Another better quantified example is the harmonic oscillator $-\Delta + V$ on $\mathbb{R}^n$, with $V(x) = |x|^2$. In this case, the eigenfunctions are known explicitly:

$$u_\alpha(x) = \prod_{j=1}^{n} H_{\alpha_j}(x_j)\, e^{-\frac{|x|^2}{2}}, \quad x \in \mathbb{R}^n, \quad \alpha \in \mathbb{N}_0^n, \tag{3-1}$$

where H_{α_j} are the Hermite polynomials of degree α_j. The exponential decay is a rigorous manifestation of localization. This pivotal example has been generalized significantly in the work of Agmon [1982], who proved the exponential decay of eigenfunctions with eigenvalues $\lambda < M$ for any potential V such that $V \geq M$ outside of some compact set. This provided the foundation for many important developments in semiclassical analysis. While there are virtually no other restrictions, Agmon's potentials must be strongly confining, i.e., they must stay above some $M > 0$ in a neighborhood of infinity.

However, localization does not have to be triggered by confinement; the reader can witness this in the example in Figure 1. From now on we concentrate on more involved cases where no bottleneck and no visible confining potential is present. In this context localization has long been attributed to purely interferential mechanisms, and as such it has seemed barely related to the local properties of stationary solutions discussed in Section 2. In this exposition we present a new approach to wave localization which shows that significant information about the eigenvalues and eigenfunctions near the bottom of the spectrum is encoded in a solution to one special stationary boundary value problem akin to (2-3). In particular, localization can be rather accurately predicted from the geometric properties of this solution.

The work of Filoche and Mayboroda [2012] introduced the *landscape function*. At first, the key to their approach was the following inequality. Assume that L is a positivity-preserving linear operator. Then all eigenfunctions, i.e., solutions to $L\varphi = \lambda\varphi$ on a bounded domain Ω, satisfy the bound

$$\frac{|\varphi(x)|}{\sup_\Omega |\varphi|} \leq \lambda\, u(x), \quad x \in \Omega, \tag{3-2}$$

where u is the solution to $Lu = 1$ in Ω. We intentionally leave undefined a notion of a positivity-preserving operator as the proof requires only positivity of the solution to $Lu = f$ for $f > 0$. In particular, the Laplacian $-\Delta$ or more generally the Schrödinger operator $-\Delta + V$, $V \geq 0$, as well as the divergence form operator $-\operatorname{div} A\nabla$ for an elliptic matrix A with bounded measurable coefficients, with Dirichlet or Neumann boundary data, fit the profile. As pointed out in Section 2.1 the bilaplacian is not positivity-preserving due to the failure of the Hadamard's conjecture, but for many examples this distinction yields only marginal errors.

Inequality (3-2) postulates that all *eigenfunctions* are controlled by a solution to one *stationary* problem,

$$Lu = 1 \quad \text{in} \quad \Omega.$$

In the context of a membrane vibration, u can be envisioned as a deflection of the membrane under a uniform load. Thus, quite naturally, in this and many other applications u exhibits a structure with clearly defined "high mountains" and "low valleys"; this will be referred to as a landscape. Due to (3-2), along the valleys of u the amplitude of any φ is small, as long as $\lambda\, u < 1$. As a result, the valleys indicate separation into localization regions. For instance, in Figure 1 a valley would indeed be a vertical line passing through P.

The proof of (3-2) is so simple that it leaves one wondering why it was not discovered before: denoting by G the Green's function of L, i.e., a solution to $L_x G(x, y) = \delta_y(x)$, we have $\varphi(x) = \int G(x, y)\lambda\varphi(y)\, dy$, $u(x) = \int G(x, y)\, dy$, and hence, by positivity of the Green's function,

$$|\varphi(x)| = \left| \int G(x, y)\lambda\varphi(y)\, dy \right| \leq \lambda \|\varphi\|_{L^\infty(\Omega)} \int |G(x, y)|\, dy = \lambda \|\varphi\|_{L^\infty(\Omega)} u(x),$$

for every $x \in \Omega$. However, this landscape function turns out to be strikingly useful. In the paper of Filoche and Mayboroda [ibid.] the reader can find a numerical study of the accuracy of the method which uses valleys of the landscape function to predict localization regions for eigenfunctions of the Dirichlet problem for a variety of elliptic operators. This work numerically demonstrates a much better precision than what inequality (3-2) would warrant. Part of this phenomenon will be explained in the next section in the context of the Anderson model. For now we mention simply that results for the bilaplacian have been confirmed experimentally for thin duraluminium plates and have provided a foundation for systematic use of the landscape function in mechanical engineering (Lefebvre et al. [2016]).

3.2 The Schrödinger equation and the effective potential of localization. In 1958 Anderson introduced the concept of suppression of electron transport due to disorder. This idea has shaped the development of mathematical and condensed matter physics over the past 50 years and continues to be a topic of active research. In Anderson's interpretation, a disorder in the potential V induces exponential decay of the eigenfunctions of the Schrödinger operator $-\Delta + V$, due to destructive interferences between waves traveling from an initial source along different propagation pathways. Over the years, Anderson localization has been confirmed experimentally and celebrated results of mathematical physics have provided rigorous proofs for many types of random potentials.

Of course the phenomenon of Anderson localization is completely different from any confinement considerations as well as from the decay of Dirichlet eigenfunctions in a bounded domain; indeed, it pertains to dense pure point spectrum rather than discrete spectrum. We will show, however, that some signatures of exponential decay in the Anderson model can also be predicted by means of the landscape function. More importantly, the landscape magically reveals a hidden "ordered" structure perceived by the eigenfunctions in the presence of a disordered potential V.

We start with a few definitions. Consider an elliptic operator

$$L = -\frac{1}{m}\,\mathrm{div}\,(m\,A\nabla) + V, \tag{3-3}$$

with Neumann, Dirichlet, or periodic boundary conditions on an open, connected, bounded, Lipschitz domain $\Omega \subset \mathbb{R}^n$. (One could equally well replace $\mathbb{R}^n$ with a compact connected C^1 manifold.) Assume that the potential $V \in L^\infty(\Omega)$ is a non-negative, bounded, real-valued function, which is non-degenerate in the sense that it is strictly positive on a subset of positive measure of Ω. We denote by $V_{\max}$ the maximum of V on Ω. As usual, $A = (a_{ij}(x))_{i,j=1}^n$ is an elliptic symmetric matrix of real bounded measurable coefficients satisfying

$$C^{-1}|\xi|^2 \leq A(x)\,\xi\cdot\xi \;\leq C\,|\xi|^2, \quad x \in \Omega, \quad \xi \in \mathbb{R}^n; \tag{3-4}$$

we also fix $m \in L^\infty(\Omega)$, a real-valued density satisfying uniform upper and lower bounds $C^{-1} \le m(x) \le C$, for some positive constant C, and set $M = \bar{\Omega}$.

A disordered structure of the potential V and/or of the matrix of coefficients A can produce localization of eigenfunctions of the operator L. Consider, for instance, $L = -\Delta + V$ on a square domain with periodic boundary conditions, where V is obtained by dividing the square into 80×80 unit subsquares and assigning to each randomly a value 0 with probability 0.6 and 4 with probability 0.4. An example of such a potential and its fundamental eigenfunction are depicted in Figure 2, (a) and (b). Led by (3-2) we have computed the landscape for this equation, determined the valleys, and superimposed the valley network (thin white lines) on the map of the eigenfunction in Figure 2, (b). Beautifully, the valleys convincingly outline the actual localization region.

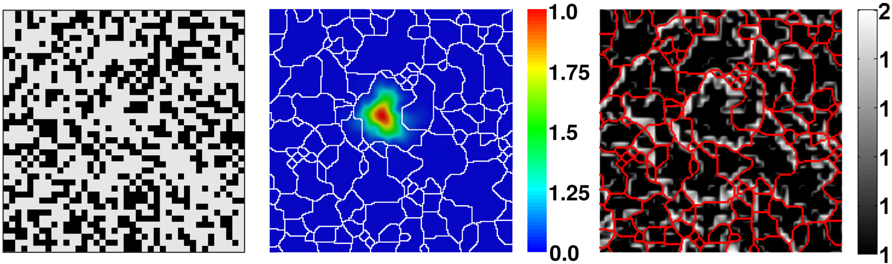


Figure 2: (a) Anderson-Bernoulli potential; (b) an eigenfunction φ together with the superimposed network of valleys; (c) $\left|\nabla \log |\varphi|\right|$ together with the superimposed network of valleys.

There are much deeper mechanisms at play than (3-2), and these are of a nonlinear nature. The work of Arnold, David, Jerison, Mayboroda, and Filoche [2018b] shows that the reciprocal of the landscape function, $1/u$, acts as an *effective potential*. The eigenfunctions are constrained to the wells of $1/u$ and undergo quantum tunneling, i.e., *exponential* decay, across the barriers of $1/u$ (in the terminology of Section 3.1, across the valleys of u). A strong manifestation of this appears in Figure 2, (c); this demonstrates that the exponential decay of the eigenfunction φ, interpreted through large values of $\left|\nabla \log |\varphi|\right|$ (white regions), holds precisely in the neighborhood of valleys of $1/u$ (thin red lines). Furthermore, the landscape function splits the domain into independently vibrating regions, and the spectrum of the original domain maps bijectively to the combined spectra of the subregions delimited by the barriers.

To state these results rigorously, denote by $\mathring{W}^{1,2}(\Omega)$, $\Omega \subset M$, the closure of the space of smooth functions compactly supported in Ω with respect to the norm $\left(\int |\nabla\varphi|^2 +\right.$

$\varphi^2\,dx\big)^{\frac{1}{2}}$. We say that $v \in \mathring{W}^{1,2}(\Omega)$ is a weak solution to $Lv = f$, $f \in L^2(\Omega)$, if

$$\int_\Omega [(A\nabla v)\cdot\nabla\eta + Vv\eta]\,m\,dx = \int_\Omega f\eta m\,dx \tag{3-5}$$

for every $\eta \in \mathring{W}^{1,2}(\Omega)$. Note that we do not require Ω to be necessarily an open or a closed set, and thus the definition above prescribes Neumann boundary conditions on the parts of $\partial\Omega$ that belong to Ω and Dirichlet boundary conditions on the parts of $\partial\Omega$ that belong to its complement.

We define the landscape function as the unique solution to $Lu = 1$ on M. By the maximum principle u is strictly positive. Now fix some small $\delta > 0$, and let

$$E(\lambda+\delta) := \{x \in M : 1/u(x) \le \lambda+\delta\}, \quad \lambda > 0.$$

The (typically disconnected) set $E(\lambda+\delta)$ can be envisioned as a collection of subsets of several regions bounded by the valleys of $1/u$ and referred to as *wells* below. The work of the author and her collaborators shows that an eigenfunction with eigenvalue λ is necessarily localized inside those wells and decays exponentially in the complement of $E(\lambda+\delta)$. To quantify this statement, let $\big(\frac{1}{u}-\lambda\big)_+ := \max\big(\frac{1}{u}-\lambda, 0\big)$, $B = A^{-1} = (b_{ij})_{i,j=1}^n$, and ρ_λ be the distance in the Riemannian metric $ds^2 = \big(\frac{1}{u}-\lambda\big)_+ B\,dx\cdot dx$, that is,

$$\rho_\lambda(x,y) = \inf_\gamma \int_0^1 \Big(\Big(\frac{1}{u}-\lambda\Big)_+(\gamma(t)) \sum_{i,j=1}^n b_{ij}(\gamma(t))\dot\gamma_i(t)\dot\gamma_j(t)\Big)^{1/2} dt, \tag{3-6}$$

where the infimum is taken over all absolutely continuous paths $\gamma : [0,1] \to \Omega$ such that $\gamma(0) = y$ and $\gamma(1) = x$. Finally, let

$$\rho(x, E(\lambda+\delta)) = \inf\{\rho_\lambda(x,y) : y \in E(\lambda+\delta)\}$$

be the Agmon distance from x to $E(\lambda+\delta)$. With this at hand, we state (a simplified version of) one of the main results in Arnold, David, Jerison, Mayboroda, and Filoche [2018b].

Theorem 3-7. *Let $\lambda > 0$ and $0 < \delta \le V_{\max}/10$ be such that $\lambda+\delta \le V_{\max}$. Let ψ be an eigenfunction of L satisfying $L\psi = \mu\psi$ on M, $\mu \le \lambda$. Then*

$$\int_{\{\rho(x,E(\lambda+\delta))\ge 1\}} e^{\rho(x,E(\lambda+\delta))}(|\nabla\psi|^2 + V_{\max}\psi^2)\,dx \le \tfrac{50\,V_{\max}}{\delta} \int_M V_{\max}\psi^2\,dx. \tag{3-8}$$

Thus, roughly speaking, an eigenfunction with the eigenvalue λ decays as $e^{-\rho(x,E(\lambda+\delta))/2}$ away from $E(\lambda+\delta)$, which, in particular, explains the exponential decay across the "valleys" in Figure 2, (c). Note that the constant in (3-8) does not depend on the ellipticity constants of A, m, or on the oscillations of V.

Underpinning these results is the identity

$$\int_M |\nabla_A f|^2 + V f^2\,dx = \int_M u^2 \Big|\nabla_A\Big(\frac{f}{u}\Big)\Big|^2 + \frac{1}{u} f^2\,dx, \quad f \in \mathring{W}^{1,2}(M), \tag{3-9}$$

where $|\nabla_A f|^2 = A\nabla f\,\nabla f$, which implies that

$$\langle Lf, f\rangle \geq \langle (1/u) f, f\rangle. \tag{3-10}$$

Here $\langle\cdot,\cdot\rangle$ denotes the inner product on $L^2(M)$.

Although very different in nature from (2-22), the weighted inequality (3-10) provides very accurate information about the solutions (this time, eigenfunctions). The crux of the matter, once again, is to choose the weight function very carefully.

The bound (3-10) is a form of the uncertainty principle. Indeed, $\langle (1/u) f, f\rangle$ combines the impact of the kinetic energy coming from $|\nabla_A f|^2$ and the potential energy given by the integral of Vf^2. Agmon already realized that a positive lower bound on the form $\langle Lf, f\rangle$ yields exponential decay of eigenfunctions, but he used the trivial inequality $\langle Lf, f\rangle \geq \langle V_+ f, f\rangle$. Later, these rudimentary ideas were advanced considerably in the framework of semiclassical analysis. In that formalism one approximates a smooth potential V near its minimum by the potential of a harmonic oscillator $C|x|^2$. In our setting, however, the potentials are not smooth, and even more importantly, the estimate for the exponential decay in terms of the Agmon distance using $(V-\lambda)_+$ is virtually useless; for example, $(V-\lambda)_+ = 0$, and hence the corresponding Agmon distance is zero, in the entire grey area in Figure 2, (a), so this estimate would give no decay whatsoever. By contrast, $\big(\frac{1}{u}-\lambda\big)_+$ is extremely efficient, as in the disordered regimes (akin to the Anderson model) a plot of $\big(\frac{1}{u}-\lambda\big)_+$ typically demonstrates a clear separation into disjoint regions, at least for smaller values of λ.

We now prove a diagonalization of the operator L modulo an exponentially small error. To this end, retaining the notation of Theorem 3-7, consider a finite decomposition of $E(\lambda+\delta) = \bigcup_\ell E_\ell$ into disjoint sets. Denote by S the Agmon distance between these sets,

$$S = \inf\{\rho_\lambda(x,y) : x \in E_\ell,\ y \in E_{\ell'},\ \ell \neq \ell'\}.$$

We regard the sets E_ℓ as individual wells of $E(\lambda+\delta)$, but note that that it might be advantageous to combine several nearby connected components into one E_ℓ in the interest of maximizing the pairwise distance S. Going further, we denote by Ω_ℓ a collection of reasonably regular sets, each containing an $(S'/2)$-neighborhood of E_ℓ, $S' < S$, and so that Ω_ℓ is disjoint from an $(S/2)$-neighborhood of $E_{\ell'}$ for any $\ell' \neq \ell$. (These are chosen "in the spirit of" Voronoi sets.) We also ensure that $\partial\Omega_\ell \cap \Omega_\ell = \partial\Omega_\ell \cap \partial\Omega$. Denote by $\varphi_{\ell,j}$

the eigenfunctions of L in $\mathring{W}^{1,2}(\Omega_\ell)$ with eigenvalues $\mu_{\ell,j}$. By assumption, they satisfy Neumann boundary conditions at $\partial\Omega_\ell \cap \partial\Omega$ and Dirichlet boundary conditions on the remaining parts of the boundary. The second main theorem in the work of Arnold, David, Jerison, Mayboroda, and Filoche [2018b] states that the eigenfunctions of the operator L on the original domain Ω are exponentially close to the functions $\varphi_{\ell,j}$, and vice versa. More precisely, for every eigenfunction ψ_j with an eigenvalue $\mu = \mu_j \le \lambda - \delta$,

$$\big\|\psi - \Phi_{(\mu-\delta,\,\mu+\delta)}\psi\big\|^2_{L^2(M,m\,dx)\to L^2(M,m\,dx)} \le 300\,(V_{\max}/\delta)^3\,e^{-S/2}, \tag{3-11}$$

where $\Phi_{(\mu-\delta,\,\mu+\delta)}$ denotes the orthogonal projection in $L^2(M, m\,dx)$ onto the span of eigenvectors $\varphi_{\ell,j}$ with eigenvalues $\mu_{\ell,j} \in (\mu-\delta, \mu+\delta)$ (extended by 0 to all of M). Conversely, the same upper bound holds for $\varphi - \Psi_{(\mu-\delta,\,\mu+\delta)}\varphi$ for any $\varphi = \varphi_{\ell,j}$ with $\mu = \mu_{\ell,j} \le \lambda - \delta$, where Ψ denotes the orthogonal projection in $L^2(M, m\,dx)$ onto the span of the eigenfunctions ψ_j with eigenvalues within $(\mu-\delta,\ \mu+\delta)$.

Moreover, if $N_0(\lambda) = \#\{\mu_{\ell,j} \le \lambda\}$ is the counting function for the combined spectrum of the domains Ω_ℓ, and $N(\lambda)$ is the spectral counting function for the original operator L on M, then for $\mu \le \lambda$ we have

$$\min(\bar{N}, N_0(\mu-\delta)) \le N(\mu) \quad \text{and} \quad \min(\bar{N}, N(\mu-\delta)) \le N_0(\mu). \tag{3-12}$$

with the threshold $\bar{N}$ such that $300\,\bar{N}\ (V_{\max}/\delta)^3\,e^{-S/2} \le 1$. It follows that, e.g., the fundamental eigenvalue can be identified with a precision $\delta \sim e^{-S/6}$ and the fundamental eigenfunction is a linear combination of eigenfunctions of individual subregions with eigenvalues roughly within a band $\big(\mu - e^{-S/6},\ \mu + e^{-S/6}\big)$.

Taking a limit as the domain becomes infinite, a potential of the type described above induces Anderson localization, that is, the system almost surely exhibits dense pure point spectrum and eigenfunctions are exponentially localized. The limiting case of our theorems depends on the dynamics of the separation between the wells and on the probability of resonances. We rigorously demonstrate the pattern of exponential decay between wells, but do not yet discuss the probability that the eigenvalues of nearby wells are exponentially close. A reasonable probabilistic conjecture in the spirit of Anderson localization would say that almost surely for some family of disordered potentials V, the effective potential $1/u$ has a well-defined structure of walls and wells which are nearly independent.

Meanwhile, the idea of diagonalization of L based on the uncertainty principle (3-10) has proved to be remarkably powerful. We have already compared our theorems with various results from the 1980s designed to treat smooth potentials. A related step in this direction, albeit one with different goals, was taken by Fefferman and Phong, who introduced an improved uncertainty principle

$$\int_\Omega |\nabla f|^2 + Vf^2\,dx \ge C_0 \int_\Omega m(x,V)\,f^2\,dx \quad \text{for all} \quad f \in C_0^\infty(\Omega), \tag{3-13}$$

where the maximal function is $m(x,V)=\inf_{r>0}\{\frac{1}{r}:\frac{1}{r^{n-2}}\int_{B(x,r)}V(y)\,dy\leq 1\}$. That result is valid when V lies in the reverse Hölder class $B^{\frac{n}{2}}(\mathbb{R}^n)$, so $\left(\int_B V^{\frac{n}{2}}(x)\,dx\right)^{\frac{2}{n}}\leq C_V\int_B V(x)\,dx$ on all balls $B\subset\mathbb{R}^n$, and the constant C_0 in (3-13) depends on the constant C_V in the $B^{\frac{n}{2}}$ condition, that is, on the oscillations of V on M. Thus, if one tries to diagonalize the operator L on M starting from (3-13), the resulting correspondence would invoke a large band of frequencies, and would produce considerably weaker bounds on eigenfunctions. In some sense, the estimate with $m(\cdot,V)$ treats all eigenfunctions as collections of bumps, and this typically becomes reasonably accurate only for large eigenvalues, whereas the effective potential function $1/u$ is sensitive to the precise shape of low-energy eigenfunctions.

An important correlate of the uncertainty principle is the Weyl law, which asserts that as $\lambda\to\infty$, the counting function for the eigenvalues of $-\Delta+V$, $N(\lambda)$, behaves like the volume $N_V(\lambda):=\left|\{(\xi,x):|\xi|^2+V(x)\leq\lambda\}\right|$. Roughly speaking, this holds because by the uncertainty principle every eigenvalue contributes a box of size at least 1 in (ξ,x) phase space. The Fefferman-Phong uncertainty principle improved this asymptotic result by developing a method to estimate the number of distorted boxes of size 1 in $\{|\xi|^2+V(x)\leq\lambda\}$ rather than simply using the volume of this set. In many examples our uncertainty principle yields a significantly better estimate for small eigenvalues. For instance, Figure 3 displays the data for a 1D Schrödinger operator on an interval $[0,256]$ obtained by randomly assigning values between 0 and 1 on each interval of size 1. The graph displays the actual counting function N (in black), the classical Weyl law estimate N_V (in blue), and the estimate using the volume $N_{1/u}(\lambda):=\left|\{(\xi,x):|\xi|^2+1/u(x)\leq\lambda\}\right|$ (in red) which treats $1/u$ rather than V as the relevant potential. Asymptotically, when λ is large, these all give approximately the same result, but near the bottom of the spectrum the estimate with $1/u$ is far more accurate. In the work of Arnold, David, Jerison, Mayboroda, and Filoche [2018a], we numerically study this phenomenon and show that in fact the local minima of the effective potential, properly normalized, already provide a good and computationally efficient approximation to the eigenvalue distribution. A mathematical treatment of these results remains open. However, these considerations, along with rigorous estimates of the exponential decay outlined above, have already initiated a transformative change in the treatment of quantum

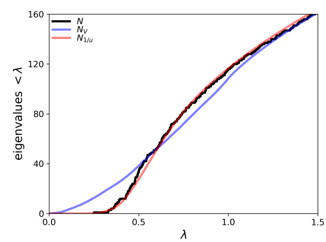


Figure 3: Counting function for the eigenvalues of a 1D Schrödinger operator with the old and new Weyl-type estimates.

localization effects in the physics of semiconductors and in the LED engineering (Filoche, Piccardo, Wu, Li, Weisbuch, and Mayboroda [2017]).

3.3 Historical references. Anderson localization is a vast and celebrated subject and we do not aim to provide a reasonable historical account. Among the major highlights in concert with this paper one could mention the pivotal work of Anderson Anderson [1958], the mathematical proofs for several types of random potentials, most notably, Aizenman and Molchanov [1993], Fröhlich and Spencer [1983], and Bourgain and Kenig [2005], and a general review of related progress in theoretical and experimental physics Abrahams [2010].

Similarly, it would be impossible to survey related achievements in semiclassical analysis. Closest to our investigation are perhaps the works of Helffer and Sjöstrand [1984] and Simon [1984].

Fefferman-Phong maximal principle discussed in Section 3.2 can be found in Fefferman [1983] and Shen [1994].

Picture credits. Figure 1 first appeared in Filoche and Mayboroda [2009] and Figure 2, (a), (c), in Arnold, David, Jerison, Mayboroda, and Filoche [2016].

References

Elihu Abrahams, ed. (2010). *50 years of Anderson localization*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, pp. xii+597. MR: 2848302 (cit. on p. 1730).

David R. Adams (1997). "Potential and capacity before and after Wiener". In: *Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994)*. Vol. 52. Proc. Sympos. Appl. Math. Amer. Math. Soc., Providence, RI, pp. 63–83. MR: 1440907.

Shmuel Agmon (1960). "Maximum theorems for solutions of higher order elliptic equations". *Bull. Amer. Math. Soc.* 66, pp. 77–80. MR: 0124618 (cit. on p. 1721).

– (1982). *Lectures on exponential decay of solutions of second-order elliptic equations: bounds on eigenfunctions of N-body Schrödinger operators*. Vol. 29. Mathematical Notes. Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, p. 118. MR: 745286 (cit. on p. 1722).

Michael Aizenman and Stanislav Molchanov (1993). "Localization at large disorder and at extreme energies: an elementary derivation". *Comm. Math. Phys.* 157.2, pp. 245–278. MR: 1244867 (cit. on p. 1730).

Philip W Anderson (1958). "Absence of diffusion in certain random lattices". *Physical review* 109.5, p. 1492 (cit. on p. 1730).

Douglas Arnold, Guy David, David Jerison, Svitlana Mayboroda, and Marcel Filoche (2016). "Effective confining potential of quantum states in disordered media". *Physical review letters* 116.5, p. 056602 (cit. on pp. 1721, 1730).

– (2018a). "Computing spectra without solving eigenvalue problems". Submitted (cit. on p. 1729).

– (2018b). "Localization of eigenfunctions via an effective potential". Submitted (cit. on pp. 1721, 1725, 1726, 1728).

Ivo Babuška (1963). "The theory of small changesin the domain of existence in the theory of partial differential equations and its applications". *Differential equations and their applications*, pp. 13–26 (cit. on p. 1720).

Jean Bourgain and Carlos E. Kenig (2005). "On localization in the continuous Anderson-Bernoulli model in higher dimension". *Inventiones Math.* 161.2, pp. 389–426. MR: 2180453 (cit. on p. 1730).

Charles V. Coffman (1982). "On the structure of solutions $\Delta^2 u = \lambda u$ which satisfy the clamped plate conditions on a right angle". *SIAM J. Math. Anal.* 13.5, pp. 746–757. MR: 668318 (cit. on p. 1720).

R. J. Duffin (1949). "On a question of Hadamard concerning super-biharmonic functions". *J. Math. Physics* 27, pp. 253–258. MR: 0029021 (cit. on p. 1720).

Charles L. Fefferman (1983). "The uncertainty principle". *Bull. Amer. Math. Soc. (N.S.)* 9.2, pp. 129–206. MR: 707957 (cit. on p. 1730).

Marcel Filoche and Svitlana Mayboroda (2009). "Strong localization induced by one clamped point in thin plate vibrations". *Physical review letters* 103.25, p. 254301 (cit. on pp. 1721, 1722, 1730).

– (2012). "Universal mechanism for Anderson and weak localization". *Proc. Natl. Acad. Sci. USA* 109.37, pp. 14761–14766. MR: 2990982 (cit. on pp. 1721, 1723).

Marcel Filoche, Marco Piccardo, Yuh-Renn Wu, Chi-Kang Li, Claude Weisbuch, and Svitlana Mayboroda (2017). "Localization landscape theory of disorder in semiconductors. I. Theory and modeling". *Physical Review B* 95.14, pp. 144–204 (cit. on p. 1730).

Jürg Fröhlich and Thomas Spencer (1983). "Absence of diffusion in the Anderson tight binding model for large disorder or low energy". *Comm. Math. Phys.* 88.2, pp. 151–184. MR: 696803 (cit. on p. 1730).

B. Helffer and J. Sjöstrand (1984). "Multiple wells in the semiclassical limit. I". *Comm. Partial Differential Equations* 9.4, pp. 337–408. MR: 740094 (cit. on p. 1730).

V. A. Kozlov, V. G. Maz'ya, and J. Rossmann (2001). *Spectral problems associated with corner singularities of solutions to elliptic equations*. Vol. 85. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, pp. x+436. MR: 1788991 (cit. on p. 1721).

Gautier Lefebvre et al. (2016). "One single static measurement predicts wave localization in complex structures". *Physical review letters* 117.7, p. 074301 (cit. on p. 1724).

W. Littman, G. Stampacchia, and H. F. Weinberger (1963). "Regular points for elliptic equations with discontinuous coefficients". *Ann. Scuola Norm. Sup. Pisa (3)* 17, pp. 43–77. MR: 0161019 (cit. on p. 1721).

Svitlana Mayboroda and Vladimir Maz'ya (2009). "Boundedness of the gradient of a solution and Wiener test of order one for the biharmonic equation". *Invent. Math.* 175.2, pp. 287–334. MR: 2470109 (cit. on pp. 1710, 1712, 1716, 1719).

– (2014). "Regularity of solutions to the polyharmonic equation in general domains". *Invent. Math.* 196.1, pp. 1–68. MR: 3179572 (cit. on pp. 1710, 1712–1714, 1719).

– (2018). "Polyharmonic capacity and Wiener test of higher order". *Invent. Math.* 211.2, pp. 779–853. MR: 3748317 (cit. on pp. 1710, 1715, 1716, 1719).

Vladimir Maz'ya (1997). "Unsolved problems connected with the Wiener criterion". In: *The Legacy of Norbert Wiener: A Centennial Symposium (Cambridge, MA, 1994)*. Vol. 60. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 199–208. MR: 1460283 (cit. on p. 1721).

Vladimir G. Maz'ya (1999). "On Wiener's type regularity of a boundary point for higher order elliptic equations". *Nonlinear analysis, function spaces and applications* 6, pp. 119–155 (cit. on p. 1721).

– (2002). "The Wiener test for higher order elliptic equations". *Duke Math. J.* 115.3, pp. 479–512. MR: 1940410 (cit. on p. 1721).

Vladimir G. Maz'ya and Jürgen Rossmann (1991). "On the Agmon-Miranda maximum principle for solutions of elliptic equations in polyhedral and polygonal domains". *Ann. Global Anal. Geom.* 9.3, pp. 253–303. MR: 1143406 (cit. on p. 1721).

– (1992). "On the Agmon-Miranda maximum principle for solutions of strongly elliptic equations in domains of $\mathbf{R}^n$ with conical points". *Ann. Global Anal. Geom.* 10.2, pp. 125–150. MR: 1175915 (cit. on p. 1720).

J. Pipher and G. C. Verchota (1995). "Maximum principles for the polyharmonic equation on Lipschitz domains". *Potential Anal.* 4.6, pp. 615–636. MR: 1361380 (cit. on pp. 1720, 1721).

Zhong Wei Shen (1994). "On the Neumann problem for Schrödinger operators in Lipschitz domains". *Indiana Univ. Math. J.* 43.1, pp. 143–176. MR: 1275456 (cit. on p. 1730).

Barry Simon (1984). "Semiclassical analysis of low lying eigenvalues. II. Tunneling". *Ann. of Math. (2)* 120.1, pp. 89–118. MR: 750717 (cit. on p. 1730).

N. Wiener (1924). "The Dirichlet problem". *J. Math. Phys.* 3, pp. 127–146 (cit. on p. 1721).

Received 2017-11-30.

SVITLANA MAYBORODA (СВІТЛАНА МАЙБОРОДА)
svitlana@math.umn.edu

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1731–1750)

MEASURABLE EQUIDECOMPOSITIONS

ANDRÁS MÁTHÉ

Abstract

The famous Banach–Tarski paradox and Hilbert's third problem are part of story of paradoxical equidecompositions and invariant finitely additive measures. We review some of the classical results in this area including Laczkovich's solution to Tarski's circle-squaring problem: the disc of unit area can be cut into finitely many pieces that can be rearranged by translations to form the unit square.

We also discuss the recent developments that in certain cases the pieces can be chosen to be Lebesgue measurable or Borel: namely, a measurable Banach–Tarski 'paradox' and the existence of measurable/Borel circle-squaring.

1 Paradoxical equidecompositions and invariant measures

In the plane, any polygon can be transformed into any other polygon of the same area by cutting it into polygonal pieces and recomposing these after applying translations and rotations (isometries of the plane). This is the Bolyai–Gerwien–Wallace theorem from the nineteenth century. The analogous problem about polyhedra in $\mathbb{R}^3$ became known as Hilbert's third problem in 1900. This was solved by Dehn by inventing an algebraic invariant that shows that the unit cube cannot be cut into finitely many polyhedral pieces that, after applying isometries, reassemble into the regular tetrahedron of unit volume.

It makes sense to study analogous questions where we may cut geometric objects into arbitrary sets, not just polygons or polyhedra. For sets $A, B \in \mathbb{R}^n$ let $A \cong B$ denote that they are congruent; that is, there is a distance-preserving bijection from A to B, or equivalently, there is an Euclidean motion (isometry) moving A to B.

Definition 1.1. We say that sets $A, B \in \mathbb{R}^n$ are *equidecomposable* if there are finite partitions $A = \cup_{i=1}^k A_i$ and $B = \cup_{i=1}^k B_i$ such that $A_i \cong B_i$ for every $i = 1, \ldots, k$.

The most famous result about equidecompositions is the following.

Partially supported by the Hungarian National Research, Development and Innovation Office – NKFIH.
MSC2010: primary 28A75; secondary 05C70, 43A07, 51M25, 52B45.

Theorem 1.2 (Banach–Tarski paradox Banach and Tarski [1924]). *If $A, B \subset \mathbb{R}^n$, $n \geq 3$, are bounded sets with non-empty interior, then A and B are equidecomposable.*

Their result is based on earlier work of Hausdorff that essentially says that the unit sphere in $\mathbb{R}^3$ is equidecomposable to the disjoint union of two unit spheres, modulo countable sets.

Theorem 1.3 (Hausdorff paradox, 1914). *Let $\mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$. Then there are partitions*

$$\mathbb{S}^2 = A_1 \cup A_2 \cup C_1 \quad \text{and} \quad \mathbb{S}^2 = A_3 \cup A_4 \cup A_5 \cup C_2$$

where the sets A_i are congruent to each other and C_1, C_2 are countable.

Hausdorff's proof is based on his discovery that $SO(3)$, the group of rotations of $\mathbb{S}^2$, contains a free subgroup of rank 2.

The interest in equidecompositions originates from questions about the existence of certain invariant measures. In 1905 Vitali proved that there are no non-trivial isometry-invariant σ-additive measures defined on all subsets of $\mathbb{R}$ (that assign measure 1 to the unit interval). Hausdorff raised the question whether at least *finitely additive* isometry-invariant measures defined on all subsets of $\mathbb{R}^n$ exist. An immediate corollary of his Theorem 1.3 is that the analogous question for $\mathbb{S}^2$ has a negative answer. His question, for $\mathbb{R}$ and $\mathbb{R}^2$, was solved by Banach in 1923.

Theorem 1.4 (Existence of Banach measures in $\mathbb{R}$ and $\mathbb{R}^2$ Banach [1923].). *In $\mathbb{R}$ and $\mathbb{R}^2$ the Lebesgue measure can be extended to all subsets as an isometry-invariant finitely additive measure.*

An immediate corollary of this theorem is that if two Lebesgue measurable sets $A, B \subset \mathbb{R}$ or $\mathbb{R}^2$ are equidecomposable, then they must have equal Lebesgue measure. Similarly, the Banach–Tarski paradox immediately implies that there are no non-trivial isometry-invariant finitely additive measures defined on all subsets of $\mathbb{R}^n$ for $n \geq 3$ (that assign positive and finite measure to the unit cube).

So clearly, the existence of paradoxical equidecompositions imply the non-existence of invariant measures. It is a fundamental theorem of Tarski that the other direction holds as well in a very general setting.

Definition 1.5. Assume that a group G acts on a set X. We say that $A, B \subset X$ are G-equidecomposable if there are finite partitions $A = \cup_{i=1}^k A_i$ and $B = \cup_{i=1}^k B_i$ such that, for each i, $B_i = g_i(A_i)$ for some $g_i \in G$.

Theorem 1.6 (Tarski 1929). *Assume that a group G acts on a set X and $A \subset X$. Then there is a G-invariant finitely additive measure μ defined on all subsets of X satisfying $\mu(A) = 1$ if and only if A cannot be written as the union of disjoint sets A', A'' such that A is G-equidecomposable to both A' and A''.*

1.1 Remark on amenability.

The Banach–Tarski paradox holds in $\mathbb{R}^n$ for $n \geq 3$, while Banach measures exist in $\mathbb{R}$ and $\mathbb{R}^2$. This contrast was better understood after von Neumann (1929) studied the behaviour of the isometry groups of these spaces. A group G is called *amenable* if there is a finitely additive probability measure μ defined on all subsets of G that is left invariant under the action of G on itself:

$$\mu(\gamma A) = \mu(A) \quad \text{for every } A \subset G,\ \gamma \in G.$$

He proved that the isometry group of $\mathbb{R}$ and $\mathbb{R}^2$ are amenable (in fact, solvable, and as he proved, every solvable group is amenable), whereas the isometry group of $\mathbb{R}^n$ ($n \geq 3$) and $SO(3)$ are not amenable.

When G is amenable, the existence of G-invariant measures carries to spaces X on which G acts.

Theorem 1.7 (Mycielski [1979]). *Assume that an amenable group G is acting on a set X, and let μ be a G-invariant finitely additive measure defined on a G-invariant algebra $\mathfrak{A}$ of subsets of X. Then μ can be extended to be a G-invariant finitely additive measure defined on all subsets of X.*

1.2 Equidecompositions using sets of the Baire property.

Recall that a set $A \subset \mathbb{R}^n$ is meager (or of the first Baire category) if it is a union of countably many nowhere dense sets. A set is said to have the Baire property if it is the symmetric difference of an open set and a meager set. (All Borel sets have the Baire property.) A set A is called Jordan measurable if it is bounded and the boundary ∂A has Lebesgue measure zero. Its Jordan measure is the same as its Lebesgue measure.

Marczewski proved an analogue of Banach's Theorem 1.4.

Theorem 1.8 (Existence of Marczewski measures in $\mathbb{R}$ and $\mathbb{R}^2$). *In $\mathbb{R}$ and $\mathbb{R}^2$ there is an isometry-invariant finitely additive measure μ defined on all subsets such that μ extends the Jordan measure and vanishes on meager sets.*

Marczewski posed the question, in 1930, whether the same holds in $\mathbb{R}^n$, for $n \geq 3$. This was unsolved until 1992 when Dougherty and Foreman proved that these measures cannot exist in higher dimensions. In fact, they proved the striking result that the Banach–Tarski paradox works with pieces that are Baire measurable.

Theorem 1.9 (Dougherty and Foreman [1992, 1994]). *Let $n \geq 3$. The unit ball in $\mathbb{R}^n$ can be equidecomposed to the union of two disjoint unit balls by sets that have the Baire property.*

The underlying statement is that a dense open subset of the unit ball is equidecomposable by open pieces to a dense open subset of the union of two disjoint unit balls. Theorem 1.9 is then obtained by "combining" this equidecomposition with the one given by the Banach–Tarski paradox restricted to a suitable meager set.

1.3 The Banach–Ruziewicz problem and the measurable Banach–Tarski paradox.

In this section we consider finitely additive measures that are defined only on the σ-algebra of Lebesgue measurable sets, not for all subsets of $\mathbb{R}^n$.

Question 1.10 (Banach–Ruziewicz problem)**.** *Let $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ be the unit sphere.*

1. *Assume that μ is a rotation invariant finitely additive probability measure defined on Lebesgue measurable sets of $\mathbb{S}^{n-1}$. Does μ necessarily coincide with the normalized Lebesgue measure?*

2. *Assume that μ is an isometry invariant finitely additive measure defined on bounded Lebesgue measurable sets of $\mathbb{R}^n$, assigning measure 1 to the unit cube. Does μ necessarily coincide with the Lebesgue measure?*

For $\mathbb{S}^1$, $\mathbb{R}$ and $\mathbb{R}^2$, the questions have a negative answer. These easily follow from the existence of Marczewski measures μ on $\mathbb{R}$ and $\mathbb{R}^2$ (Theorem 1.8), as μ differs from Lebesgue measure on meager sets of positive Lebesgue measure.

On the other hand, for $n \geq 3$, the answers to both questions are positive. This was independently proved by Margulis [1980] and Sullivan [1981] for $n \geq 5$ in 1980 and then by Drinfel'd [1984] for $n = 2, 3$ in 1984. The proofs rely on Kazhdan's property (T) and the existence of a spectral gap for certain averaging operators, see Section 2 for details.

The Banach–Ruziewicz problem also has a connection to paradoxical equidecompositions. The following theorem can be regarded as the measurable version of the Banach–Tarski paradox. It is easy to see that it implies the positive answers to Question 1.10 for $n \geq 3$.

Theorem 1.11 (Grabowski, Máthé, and Pikhurko [2016])**.** *Let $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ be the unit sphere, and let $n \geq 3$.*

1. *Let $A, B \subset \mathbb{S}^{n-1}$ be measurable sets with non-empty interior and equal Lebesgue measure. Then A is equidecomposable to B using rotations with* measurable pieces.

2. *Let $A, B \subset \mathbb{R}^n$ be bounded measurable sets with non-empty interior and equal Lebesgue measure. Then A is equidecomposable to B with* measurable pieces.

The proof is based on work by Lyons and Nazarov [2011], Elek and Lippner [2010] and the spectral gap results of Margulis, Sullivan and Drinfeld. We review this proof in detail in Section 2.

1.4 Tarski's circle-squaring problem.

Question 1.12 (Tarski's circle-squaring problem, 1925). *Are the disc and a square of the same area equidecomposable?*

Dubins, Hirsch, and Karush [1963] proved in 1963 that circle-squaring is not possible by pieces that are Jordan domains (that is, topological discs), even if their boundaries can be ignored.

In 1985 Gardner proved that circle-squaring is not possible if the pieces are arbitrary but they are moved by isometries that generate a locally discrete group. In fact, he proved the following theorem.

Theorem 1.13 (Gardner [1985]). *Let G be a locally discrete group of isometries of $\mathbb{R}^n$. If a convex polytope and a convex set in $\mathbb{R}^n$ are G-equidecomposable, then they are G-equidecomposable with convex pieces.*

Finally, in 1990, Tarski's circle-squaring problem was solved by Laczkovich in the affirmative.

Theorem 1.14 (Laczkovich [1990]). *The disc is equidecomposable to a square of the same area; in fact, it is enough to use translations only.*

His proof extends to more general sets, assuming a condition on their boundary. Recall that the upper Minkowski (or box) dimension of a (non-empty bounded) set $X \subset \mathbb{R}^n$ is

$$\overline{\dim}_M(X) = \limsup_{\delta \to 0} \frac{\log N_\delta(X)}{-\log \delta}$$

where $N_\delta(X)$ denotes the minimum number of (grid) cubes of side δ that cover the set X.

Theorem 1.15 (Laczkovich [1992a]). *Let A and B be bounded measurable sets in $\mathbb{R}^n$ with equal positive Lebesgue measure such that $\overline{\dim}_M(\partial A) < n$ and $\overline{\dim}_M(\partial B) < n$. Then A and B are equidecomposable using translations; that is, there exist partitions $A = \cup_{i=1}^k A_i$, $B = \cup_{i=1}^k B_i$ and translation vectors $x_i \in \mathbb{R}^n$ such that $B_i = A_i + x_i$.*

The condition on the Minkowski dimension is satisfied, for example, when the sets are bounded convex sets, or if they are Jordan domains of the plane with rectifiable boundaries (of finite length).

The condition on the Minkowski dimension of the boundary is necessary. Laczkovich [1993] proved that there exist intervals converging to zero such that their union A is not equidecomposable to an interval. (The boundary of A is a sequence of points converging to zero — such sequences can have Minkowski dimension one, and it has in this case.)

It has turned out recently that circle-squaring is possible with "nice" pieces as well.

Theorem 1.16 (Grabowski, Máthé, and Pikhurko [2017]). *Circle-squaring is possible using translations with pieces that are Lebesgue measurable and have the Baire property.*

This result was quickly superseded by a stronger result of Marks and Unger.

Theorem 1.17 (A. S. Marks and S. T. Unger [2017]). *Circle-squaring is possible using translations with Borel pieces.*

In fact, both Grabowski, Máthé, and Pikhurko [2017] and A. S. Marks and S. T. Unger [2017] prove that these equidecompositions exist not just for the disc and square, but for any sets A, B satisfying the assumptions of Theorem 1.15. Both papers build on Laczkovich's work.

In the next section we discuss the connection of equidecompositions to perfect matchings in (infinite) bi-partite graphs. To briefly summarise these proofs: Theorem 1.15 is proved by checking that Hall's condition holds in the bi-partite graph implying the existence of a perfect matching; Theorem 1.16 considers (an algorithm involving) augmenting paths to provide a measurable perfect matching; and Theorem 1.17 uses (Borel) flows to find a Borel perfect matching.

We will discuss Laczkovich's solution in Section 3, and the proof of Theorem 1.16 (for measurable pieces) in Section 5. (For the elegant proof Theorem 1.17, see the original paper A. S. Marks and S. T. Unger [ibid.].)

Remark 1.18. Before the results on the measurable and Borel circle-squaring, it was already known that Laczkovich's non-measurable circle-squaring implies circle-squaring by *measurable functions*. This was proved independently by Wehrung and Laczkovich. We review the exact statement and its proof in Section 4.

1.5 Equidecompositions and perfect matchings in bi-partite graphs. When we are looking for equidecompositions, the best way is to first fix the finitely many isometries that we are going to use to move the pieces, and then try to find the pieces. This way the problem of finding equidecompositions reduces to finding perfect matchings in certain bi-partite graphs. Let A, B be arbitrary subsets of a set X, and let H be a finite set of bijections $X \to X$. (For example, $X = \mathbb{R}^n$ and H is a finite set of isometries.) To avoid confusion later, let us assume that A and B are disjoint. Define the bi-partite graph

$$\Gamma_H(A, B) = \{(a, b) \in A \times B \,:\, b = f(a) \text{ for some } f \in H\}.$$

Then the two sets of vertices of the bi-partite graph are A and B, and a vertex $a \in A$ is connected to a vertex $b \in B$ by an edge if one of the bijections $f \in H$ move a to b.

Recall that a set $\mathfrak{M}$ of edges of a bi-partite graph is a *matching* if no vertex is covered by two edges in $\mathfrak{M}$, and it is a *perfect matching* if every vertex is covered exactly once. The following is obvious.

Lemma 1.19.

1. *There is a perfect matching in the graph* $\Gamma_H(A, B)$ *if and only if* A *is equidecomposable to* B *by using bijections of* H.
2. *Let* G *act on* X. *Then* $A, B \subset X$ *are* G*-equidecomposable if there is a finite set* $H \subset G$ *such that* $\Gamma_H(A, B)$ *contains a perfect matching.*

As it is noted by Laczkovich [2002], the connection of equidecomposability and perfect matchings was already known and used by König and Tarski.

Hall's marriage theorem extend to the case of infinite bi-partite graphs if all degrees are finite (Rado [1949]). In particular, we obtain the following.

Lemma 1.20. *The graph* $\Gamma_H(A, B)$ *contains a perfect matching if and only if*

$$|N(U)| \geq |U| \tag{1-1}$$

for every finite set of vertices U*, where* $N(U)$ *denotes the set of neighbours of* U.

Therefore proving that Hall's condition (1-1) holds is enough to prove the existence of an equidecomposition. However, it does not yield measurable or Borel equidecompositions as Rado's result relies on the Axiom of Choice.

Remark 1.21. For the full story of equidecompositions see Wagon's excellent book on the Banach–Tarksi paradox Wagon [1985] and the new edition by Tomkowicz and Wagon [2016]. Another excellent reading is Laczkovich's monograph on paradoxes in measure theory Laczkovich [2002].

2 Measurable Banach–Tarski

In this section we sketch the proof of Theorem 1.11 Grabowski, Máthé, and Pikhurko [2016] following Grabowski, Máthé, and Pikhurko [2014]. We focus on the technically easier case of the sphere $\mathbb{S}^{n-1}$. First we consider measurable equidecompositions modulo nullsets.

Theorem 2.1. *Let* $n \geq 3$ *and let* $A, B \subset \mathbb{S}^{n-1}$ *be measurable sets with non-empty interiors and of the equal Lebesgue measure. Then there are measurable sets* $A' \subset A$ *and* $B' \subset B$ *that are equidecomposable with measurable pieces such that* $A \setminus A'$ *and* $B \setminus B'$ *have measure zero.*

Sketch of proof. Let us assume, to simplify notation, that A and B are disjoint. We may assume that A, B are Borel sets (as we can forget about nullsets). Let λ denote the normalised Lebesgue measure on $\mathbb{S}^{n-1}$. The key ingredient of the proof is the following *spectral gap property*.

Lemma 2.2 (Margulis [1980], Sullivan [1981] for $n \geq 5$, Drinfel'd [1984] for $n \geq 3$). *There exist rotations $\gamma_1, \dots, \gamma_k \in SO(n)$ and $\varepsilon > 0$ such that the averaging operator $T : L^2(\mathbb{S}^{n-1}, \lambda) \to L^2(\mathbb{S}^{n-1}, \lambda)$ defined by*

$$(Tf)(x) = \frac{1}{k}\sum_{i=1}^{k} f(\gamma_i(x)) \qquad (f \in L^2(\mathbb{S}^{n-1}, \lambda),\ x \in \mathbb{S}^{n-1})$$

satisfies $\|Tf\|_2 \leq (1-\varepsilon)\|f\|_2$ for every $f \in L^2(\mathbb{S}^{n-1}, \lambda)$ with $\int f\, d\lambda = 0$.

It is easy to show that this lemma implies the following *expansion property*. For any large constant $C > 1$ there is a finite set S of rotations such that for every measurable set $U \subset \mathbb{S}^{n-1}$ we have

$$\lambda\big(\cup_{\gamma \in S} \gamma(U)\big) \geq \min\big(C\lambda(U),\ 1 - C^{-1}\big). \tag{2-1}$$

Since A and B have non-empty interior in the sphere, we can find a finite set of rotations T such that $\mathbb{S}^{n-1} = \cup_{\gamma \in T}\gamma(A) = \cup_{\gamma \in T}\gamma(B)$. Choose C so that $C \geq 2|T|$ and $C^{-1} \leq \lambda(A)/3$ and with the obtained $S = S(C)$ define

$$R = T^{-1}S \cup S^{-1}T = \{\tau^{-1}\gamma : \tau \in T, \gamma \in S\} \cup \{\gamma^{-1}\tau : \tau \in T, \gamma \in S\} \subset SO(n).$$

Then R is closed under taking inverses, $R^{-1} = R$.

As in Section 1.5, consider the bi-partite graph Γ whose set of vertices is $A \cup B$ and there is an edge between $x \in A$ and $y \in B$ if $y = \gamma(x)$ for some $\gamma \in R$. (We will also use the notation $xy \in E(\Gamma)$ in this case.) Then the following expansion property holds.

Claim 2.3. *Let $U \subset A \cup B$ and let $N(U)$ be the set of neighbours of U in G. Then*

$$\lambda(N(U)) \geq \min\big(2\lambda(U),\ \tfrac{2}{3}\lambda(A)\big). \tag{2-2}$$

Proof. It is enough to prove the claim when $U \subset A$ or $U \subset B$. We may assume $U \subset A$. Let $S.U$ denote $\cup\{\gamma(U) : \gamma \in S\}$. By (2-1), we have $\lambda(S.U) \geq 2|T|\lambda(U)$ or $\lambda(S.U) \geq 1 - \lambda(A)/3$. First assume the former. Since the sets $\tau(B)$ $(\tau \in T)$ cover the sphere, there is $\tau \in T$ such that

$$2\lambda(U) \leq \lambda(S.U \cap \tau(B)) = \lambda(\tau^{-1}(S.U) \cap B) \leq \lambda(N(U)).$$

Now assume that $\lambda(S.U) \geq 1 - \lambda(A)/3$. Then, for any $\tau \in T$,

$$\tfrac{2}{3}\lambda(B) \leq \lambda(S.U \cap \tau(B)) = \lambda(\tau^{-1}(S.U) \cap B) \leq \lambda(N(U)). \qquad \square$$

A matching $\mathfrak{M}$ in the graph Γ is called *Borel* if there exist disjoint Borel subsets $A_\gamma \subset A$ indexed by $\gamma \in R$ such that

$$\mathfrak{M} = \cup_{\gamma \in R}\big\{\{x, \gamma(x)\} : x \in A_\gamma\big\}. \tag{2-3}$$

Clearly, in order to finish the proof it is enough to find a Borel matching in Γ such that the set of unmatched vertices has measure zero. As noted in Lyons and Nazarov [2011, Remark 2.6], the expansion property (2-2) suffices for this. In the following, we just outline their strategy.

Recall that an *augmenting path* for a matching $\mathfrak{M}$ is a path which starts and ends at an unmatched vertex and such that every second edge belongs to $\mathfrak{M}$. A *Borel augmenting family* is a Borel subset $U \subset A \cup B$ and a finite sequence $\gamma_1, \dots, \gamma_l$ of elements of R such that (i) for every $x \in U$ the sequence $y_0, \dots, y_l$, where $y_0 = x$ and $y_j = \gamma_j(y_{j-1})$ for $j = 1, \dots, l$, forms an augmenting path and (ii) for every distinct $x, y \in U$ the corresponding augmenting paths are vertex-disjoint.

As shown by Elek and Lippner [2010], there exists a sequence $(\mathfrak{M}_i)_{i\in\mathbb{N}}$ of Borel matchings such that $\mathfrak{M}_i$ admits no augmenting path of length at most $2i - 1$ and $\mathfrak{M}_{i+1}$ can be obtained from $\mathfrak{M}_i$ by iterating the following at most countably many times: pick some Borel augmenting family $(U, \gamma_1, \dots, \gamma_l)$ with $l \le 2i + 1$ and flip (i.e. augment) the current matching along all paths given by the family. See Elek and Lippner [ibid.] for more details.

Our task now is to show, using Claim 2.3, that the measure of of vertices not matched by $\mathfrak{M}_i$ tends to zero as $i \to \infty$ and that the sequence $(\mathfrak{M}_i)_{i\in\mathbb{N}}$ stabilises almost everywhere (that is, for almost every vertex, the edge in $\mathfrak{M}_i$ containing the vertex stabilises as $i \to \infty$).

Lemma 2.4. *Let $i \ge 1$. Then the measure of vertices of A and B that are not covered by $\mathfrak{M}_i$ is at most*

$$2\lambda(A) \cdot 2^{-\lfloor (i-1)/2 \rfloor}.$$

Before proving this lemma, let us finish the proof of the theorem.

As we noted before, $\mathfrak{M}_{i+1}$ arises from $\mathfrak{M}_i$ by flipping augmenting paths of length at most $2i+1$ in a Borel way. When one such path is flipped, two vertices are removed from the current set of unmatched vertices. Using this observation and the fact that each rotation is measure-preserving, one can show that the set of vertices covered by the symmetric difference $\mathfrak{M}_{i+1} \triangle \mathfrak{M}_i$ has measure at most $(2i + 2)$ times the measure of unmatched vertices by $\mathfrak{M}_i$. We know from the lemma that this goes to 0 exponentially fast with i; in particular, it is summable over $i \in \mathbb{N}$. The Borel–Cantelli Lemma implies that the sequence of matchings $(\mathfrak{M}_i)_{i\in\mathbb{N}}$ stabilises almost everywhere.

Proof of Lemma 2.4. Let us fix $i \ge 1$ and let X_0 be the subset of A consisting of vertices that are not matched by $\mathfrak{M}_i$. An *alternating path of length l* is a sequence of distinct vertices $x_0, \dots, x_l$ such that (i) $x_0 \in X_0$, (ii) for odd j we have $x_j x_{j+1} \in \mathfrak{M}_i$, and (iii) for even j we have $x_j x_{j+1} \in E(\Gamma) \setminus \mathfrak{M}_i$. Let X_j consist of the end-vertices of alternating paths of length at most j. Clearly for all j we have $X_j \subset X_{j+1}$ and so, in particular, $\lambda(X_{j+1}) \geq \lambda(X_j)$. For $j \geq 1$, let $X'_j = X_j \setminus X_{j-1}$.

It's not difficult to show the following.

Claim 2.5. *For every odd $j \leq 2i-1$ we have $\lambda(X'_j) = \lambda(X'_{j+1})$ and $\lambda(X_j \cap B) \leq \lambda(X_{j+1} \cap A)$.*

Proof of Claim. All vertices in X'_j are covered by the matching $\mathfrak{M}_i$, for otherwise we would have an augmenting path of length j. It follows that $\mathfrak{M}_i$ gives a bijection between X'_j and X'_{j+1}. If we take the sets A_γ that represent $\mathfrak{M}_i$ as in Equation (2-3), then the partitions $\cup_{\gamma \in R} A_\gamma$ and $\cup_{\gamma \in R} \gamma(A_\gamma)$ induce a Borel equidecomposition between X'_j and X'_{j+1}, so these sets have the same measure, as required.

The second part (i.e. the inequality) follows analogously from the fact that $\mathfrak{M}_i$ gives an injection of $X_j \cap B$ into $X_{j+1} \cap A$ (with X_0 being the set of vertices missed by this injection). $\square$

Let k be even, with $2 \leq k \leq 2i-2$. Let $U = X_k \cap A$. We have that $N(U) = X_{k+1} \cap B$. By Claim 2.3,

$$\lambda(X_{k+1} \cap B) = \lambda(N(U)) \geq \min\left(\frac{2}{3}\lambda(A), 2\lambda(U)\right).$$

If $\lambda(X_{k+1} \cap B) \geq \frac{2}{3}\lambda(A)$ then, by Claim 2, $\lambda(X_{k+2} \cap A) \geq \lambda(X_{k+1} \cap B) \geq \frac{2}{3}\lambda(A)$ and thus

$$\lambda(X_{k+2}) = \lambda(X_{k+1} \cap B) + \lambda(X_{k+2} \cap A) \geq \frac{4}{3}\lambda(A).$$

Now, suppose that $\lambda(X_{k+1} \cap B) \geq 2\lambda(U)$. By applying Claim 2.5 for $j = k-1$ we obtain

$$\lambda(X'_{k+1}) = \lambda(X_{k+1} \cap B) - \lambda(X_{k-1} \cap B) \geq 2\lambda(U) - \lambda(U) = \lambda(U).$$

Again, by Claim 2, $\lambda(X'_{k+2}) = \lambda(X'_{k+1})$ and $\lambda(X_k) = \lambda(X_{k-1} \cap B) + \lambda(U) \leq 2\lambda(U)$. Thus

$$\lambda(X_{k+2}) = \lambda(X_k) + \lambda(X'_{k+1}) + \lambda(X'_{k+2}) \geq \lambda(X_k) + 2\lambda(U) \geq 2\lambda(X_k).$$

Thus the measure of X_k expands by factor at least 2 when we increase k by 2, unless $\lambda(X_{k+2}) \geq \frac{4}{3}\lambda(A)$. Also, this conclusion formally holds for $k = 0$, when $X_1 = N(X_0)$.

By using induction, we conclude that, for all even k with $0 \leq k \leq 2i$,

$$\lambda(X_k) \geq \min\left(\frac{4}{3}\lambda(A), 2^{k/2}\lambda(X_0)\right). \tag{2-4}$$

In the same fashion we define Y_0 to be the subset of B consisting of vertices not matched by $\mathfrak{M}_i$ and let Y_j consist of the end-vertices of alternating paths that start in Y_0 and have length at most j. As before, we obtain that the sets Y_j satisfy the analogue of Equation (2-4).

The sets X_{i-1} and Y_i are disjoint for otherwise we would find and augmenting path of length at most $2i-1$. It follows that they cannot each have measure more than $\lambda(A) = \frac{1}{2}\lambda(A \cup B)$. Since $\lambda(X_0) = \lambda(Y_0)$ we conclude that

$$\lambda(X_0 \cup Y_0) \leq 2\,\lambda(A) \cdot \left(\frac{1}{2}\right)^{\lfloor (i-1)/2 \rfloor}. \tag{2-5}$$

This proves the Lemma. □

This finishes the proof of Theorem 2.1. □

Theorem 2.6. *Let $n \geq 3$ and let $A, B \subset \mathbb{S}^{n-1}$ be measurable sets with non-empty interiors and of the equal Lebesgue measure. Then there are measurable sets $A' \subset A$ and $B' \subset B$ that are equidecomposable with measurable pieces such that $A \setminus A'$ and $B \setminus B'$ have measure zero.*

Sketch of proof. The argument that leads to Claim 2.3 can be adopted to show that $|N(X)| \geq 2\,|X|$ for every *finite* subset X of A (and of B). By a result of Rado [1949], this guarantees that Γ has a perfect matching. The (exact) measurable equidecomposition of A and B can be obtained by modifying the equidecomposition from Theorem 2.1 on suitably chosen sets of measure zero (containing $A \setminus A'$ and $B \setminus B'$) where we use Rado's theorem and the Axiom of Choice. This way we obtain a measurable equidecomposition between A and B. □

Remark 2.7. The spectral gap property as stated in Lemma 2.2 fails in $\mathbb{R}^n$. However, one can argue that a suitable reformulation of the expansion property still holds Grabowski, Máthé, and Pikhurko [2016], and that is enough to prove Theorem 1.11 for $\mathbb{R}^n$, $n \geq 3$. Alternatively, one can use the notion of a local spectral gap Boutonnet, Ioana, and Golsefidy [2017], see also Grabowski, Máthé, and Pikhurko [2016].

3 Laczkovich's circle-squaring

The aim of this section is to decribe the main steps in Laczkovich's proof of Tarski's circle-squaring problem, providing an equidecomposition between the disc and a square (using non-measurable pieces).

Instead of looking at the problem in $\mathbb{R}^n$, we may assume that A, B are subsets of the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. Indeed, we may assume that $A, B \subset [0, 1/3]^n$, and in this case, any equidecomposition of A to B on the torus yields an equidecomposition of A to B in $\mathbb{R}^n$, with the same (number of) pieces. Let λ denote the probability Lebesgue measure in $\mathbb{T}^n$.

Let d be a sufficiently large positive integer, and fix translation vectors $v_1, v_2, \ldots, v_d \in \mathbb{T}^n$ that are linearly independent over the rationals and generate a dense subgroup (isomorphic to $\mathbb{Z}^d$) of $\mathbb{T}^n$. (It will turn out later that a random choice of these vectors is what we need.) Considering how the subgroup's cosets intersect A and B we define, for $u \in \mathbb{T}^n$,

$$A_u^v = \Big\{(k_1, \ldots, k_d) \in \mathbb{Z}^d \ : \ u + k_1 v_1 + \ldots k_d v_d \in A\Big\},$$

$$B_u^v = \Big\{(k_1, \ldots, k_d) \in \mathbb{Z}^d \ : \ u + k_1 v_1 + \ldots k_d v_d \in B\Big\}.$$

The following lemma implies that if we can find bijections from A_u^v to B_u^v that move every point by at most a fixed distance, then A and B are equidecomposable using translation vectors that are integer linear combinations of the vectors v_j.

Lemma 3.1. *The following statements are equivalent for every constant C.*

(i) A and B are equidecomposable using the translation vectors

$$V_C = \left\{ \sum_{j=1}^{d} k_j v_j \in \mathbb{T}^n \ : \ |k_j| \le C \right\}.$$

(ii) For every $u \in \mathbb{T}^n$ there exist a bijection $f_u : A_u^v \to B_u^v$ such that

$$\| f_u(k) - k \|_\infty \le C \quad (k \in \mathbb{Z}^d).$$

Proof. Notice that A and B are equidecomposable using vectors $w_1, \ldots, w_m$ if and only if there is a bijection $f : A \to B$ such that for every $x \in A$, $f(x) - x \in \{w_1, \ldots, w_m\}$.

So (i) is equivalent to saying that there is a bijection $f : A \to B$ such that for every $x \in A$, $f(x) - x \in V_C$. On the other hand, it is easy to see that (ii) is equivalent to saying that for every coset C_u of the group generated by the vectors v_j there is a bijection $f : A \cap C_u \to B \cap C_u$ with $f(x) - x \in V_C$ for $x \in A \cap C_u$.

Clearly, (i) implies (ii). For the opposite direction, observe that many choices of u determine the very same coset. Still, we can pick exactly one u from each coset by the Axiom of Choice. Therefore (ii) implies (i). □

In order to prove Theorem 1.15, it is enough to prove that (ii) of Lemma 3.1 holds for some choice of vectors $v_1, \ldots, v_d$. By the pointwise ergodic theorem, we can expect that the densities of the sets A_u^v and B_u^v in $\mathbb{Z}^d$ will be the same (and equal to $\lambda(A) = \lambda(B)$). This is obviously necessary for the existence of the bijections f_u^v. However, we need to know much more about these sets than just the densities. In particular, we need fine estimates on the number of points of A_u^v and B_u^v inside any cube of $\mathbb{Z}^d$.

First consider the case that $A = X \subset \mathbb{T}^n$ is a box, that is, the product of sub-intervals of $[0,1)$. An application of the Erdős–Turán–Koksma inequality yields the following.

Lemma 3.2. *Let $v_1, \dots, v_d \in \mathbb{T}^n$ be uniformly distributed independent random vectors. Then with probability 1, there is a constant c such that for every box $X \subset \mathbb{T}^n$, every $u \in \mathbb{T}^n$, every cube $Q \subset \mathbb{Z}^d$ (containing more than 1 point) we have that*

$$\Big||X_u^v \cap Q| - \lambda(X)\,|Q|\Big| \le c \log^{n+d+1} |Q|.$$

This lemma implies a similar discrepancy result (with weaker upper bounds) for sets A of (upper) Minkowski dimension less than n.

Lemma 3.3. *Let $A \subset \mathbb{T}^n$ satisfy $\overline{\dim}_M(A) < n$. If d is large enough, the following is true. Let $v_1, \dots, v_d \in \mathbb{T}^n$ be uniformly distributed independent random vectors. With probability 1, there is a constant c (depending on A and the vectors v_j) such that for every $u \in \mathbb{T}^n$, every cube $Q \subset \mathbb{Z}^d$ with side length N we have that*

$$\Big||A_u^v \cap Q| - \lambda(A)\,|Q|\Big| \le c N^{d-2}.$$

(Instead of the bound N^{d-2}, any exponent less than $d-1$ would be sufficient.)

Lemma 3.3 follows from Lemma 3.2 by a result of Niederreiter and Wills [1975]. The outline of the proof is the following. For some $\delta > 0$, cover A by grid cubes of side δ. Consider those cubes that are in the interior of A. If we apply Lemma 3.2 to all of these boxes, we cannot obtain a good bound on the discrepancy. Instead, we merge some of these cubes into boxes in the following way. If two cubes (in the interior of A) share an $(n-1)$-dimensional face and have the same projection to the first $n-1$ coordinates, we merge them into the same box. Applying Lemma 3.2 to these boxes and using trivial bounds for the grid cubes that intersect the boundary of A, we obtain a bound on the discrepancy. The last step in the proof is to choose δ so that we minimize this bound on the discrepancy.

The key and most difficult part in Laczkovich's proof is the following theorem.

Theorem 3.4 (Laczkovich [1992b]). *Let $A^*, B^* \subset \mathbb{Z}^d$ and suppose that there are $\alpha, \varepsilon, c > 0$ such that*

$$\Big||A^* \cap Q| - \alpha|Q|\Big| \le c N^{d-1-\varepsilon},$$

$$\Big||B^* \cap Q| - \alpha|Q|\Big| \le c N^{d-1-\varepsilon}$$

for every cube $Q \subset \mathbb{Z}^d$ of side length N. Then there is a bijection $f : A^* \to B^*$ such that $\|f(k) - k\|_\infty \le C$ for every $k \in A^*$, where the constant C depends only on α, ε, c and d.

By Lemma 1.19, the existence of this bijection $f : A^* \to B^*$ is equivalent to the existence of a perfect matching in the bi-partite graph $\Gamma(A^*, B^*)$ where $a \in A^*$ is connected to $b \in B^*$ if $\|a-b\|_\infty \leq C$. The main part of Laczkovich's proof is checking that Hall's condition is satisfied (for large enough C).

Finally, it is easy to see that Theorem 3.4 and Lemma 3.3 imply Theorem 1.15.

Remark 3.5. Laczkovich's estimate is that about 10^{40} pieces are enough to equidecompose the disc to the square.

4 Circle-squaring with measurable functions

The following result was proved independently by Laczkovich [1996] and Wehrung [1992]. See also Laczkovich [2002, Theorem 9.6].

Theorem 4.1. *Suppose A and B are Lebesgue measurable sets in $\mathbb{R}^n$. If A and B are equidecomposable under isometries g_1, ..., g_m from an amenable group G (for example, they are all translations), then there are non-negative Lebesgue measurable functions $f_1, \dots, f_m$ such that*

$$\begin{aligned} 1_A &= f_1 + \dots + f_m \\ 1_B &= f_1 \circ g_1^{-1} + \dots + f_m \circ g_m^{-1}. \end{aligned}$$

(In such cases we say that A and B are continuously equidecomposable with Lebesgue measurable functions $f_1, \dots, f_m$.)

The proof is very short and enlightening so we include it here. The idea is that one can approximate non-measurable sets by Lebesgue measurable functions by considering convolutions with Lebesgue measurable mollifiers where the integration is with respect to a (finitely additive) G-invariant measure defined on all subsets of $\mathbb{R}^n$.

Proof. Since the Lebesgue measure is isometry invariant, Theorem 1.7 gives us a G-invariant finitely additive measure μ defined on all subsets of $\mathbb{R}^n$ that extends Lebesgue measure. (For $n = 1, 2$, this measure μ can be taken to be the Banach measure.)

By the assumption on equidecomposability, there is a partition $A = A_1 \cup \dots \cup A_m$ such that $B = g_1(A_1) \cup \dots \cup g_m(A_m)$. Let $B(x, r)$ denote the open ball around x of radius r.

Consider the sequences of densities

$$f_i^k(x) = \frac{\mu(B(x, 1/k) \cap A_i)}{\mu(B(x, 1/k))}$$

where k is a positive integer, and the denominator does not depend on x. The functions f_i^k are measurable, because they are Lipschitz. Notice that

$$\sum_{i=1}^{m} f_i^k(x) = \frac{\mu(B(x,1/k) \cap A)}{\mu(B(x,1/k))}$$

and, since μ is invariant under the isometries g_i,

$$\sum_{i=1}^{m} f_i^k(g_i^{-1}(x)) = \frac{\mu(B(x,1/k) \cap B)}{\mu(B(x,1/k))}.$$

Take a subsequence (k_j) such that the weak-* limits exist:

$$f_i = \lim_{j\to\infty} f_i^{k_j}.$$

We obtain measurable functions f_i (defined almost everywhere) that satisfy

$$\sum_{i=1}^{m} f_i = 1_A \tag{4-1}$$

and

$$\sum_{i=1}^{m} f_i \circ g_i^{-1} = 1_B \tag{4-2}$$

almost everywhere (since A and B are Lebesgue measurable). We claim that one can modify these functions on a nullset such that the equalities hold everywhere, using the original equidecompositions. Indeed, consider a nullset that is closed under the countable group generated by the isometries g_i and contains the points where (4-1) or (4-2) fail or the functions f_i are not defined. On this nullset, redefine f_i to be the characteristic function of A_i. □

5 Measurable circle-squaring

Theorem 5.1 (Grabowski, Máthé, and Pikhurko [2017]). *Let A and B be bounded measurable sets in $\mathbb{R}^n$ with equal positive Lebesgue measure such that $\overline{\dim}_M(\partial A) < n$ and $\overline{\dim}_M(\partial B) < n$. Then A and B are equidecomposable using translations with measurable pieces; that is, there exist partitions $A = \cup_{i=1}^k A_i$, $B = \cup_{i=1}^k B_i$ and translation vectors $x_i \in \mathbb{R}^n$ such that $B_i = A_i + x_i$.*

(Note that under the same assumptions, A and B are equidecomposable with Borel pieces by the result of A. S. Marks and S. T. Unger [2017].)

Note that it is enough to prove that A and B are equidecomposable up to nullsets with measurable pieces. Indeed, having such an equidecomposition, we can extend it and modify it on a nullset (which is invariant under our translations) using Theorem 1.15 to obtain a measurable equidecomposition of A to B with measurable pieces.

As in Section 3, we may assume that $A, B \subset \mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. We may also assume that A and B are disjoint.

Definition 5.2. Given a finite set $V = \{v_1, \ldots, v_d\} \subset \mathbb{T}^n$ and a positive integer C, let

$$V_C = \left\{ \sum_{j=1}^{d} k_j v_j \ : \ |k_j| \leq C \right\}.$$

As in Section 1.5, consider the bi-partite graph

$$\Gamma_{V_C}(A, B) = \{(a, b) \ : \ a \in A, \ b \in B, \ b - a \in V_C\}.$$

We may simply write $\Gamma(A, B)$ when V_C is clear from the context.

Similarly to Lemma 3.1, we have the following.

Lemma 5.3. *For any finite set of vectors $V = \{v_1, \ldots, v_d\} \subset \mathbb{T}^n$ and any constant C, the following are equivalent.*

(i) A and B are equidecomposable with measurable pieces using translation vectors from V_C.

(ii) There exists a measurable bijection $f : A \to B$ such that

$$f(x) - x \in V_C \ \text{ for every } x \in A.$$

In other words, there is a measurable perfect matching in $\Gamma_{V_C}(A, B)$. □

Recall that Laczkovich's proof of Theorem 1.15 relied on Hall's and Rado's theorem to conclude the existence of a perfect matching in the graph Γ_{V_C}. This graph has continuum many connected components as every connected component is contained by a coset of the subgroup generated by V. If there was a measurable set that intersected every coset in exactly one point, then the same proof would yield a measurable equidecomposition. Of course, such measurable set does not exist.

One of the ideas of the proof of Theorem 5.1 is to consider small Borel sets in $\mathbb{T}^n$ that intersect every coset in a sufficiently sparse but non-empty set, and use the points of these sparse sets as the origins of (local) coordinate systems in the cosets (that are isomorphic to $\mathbb{Z}^d$).

The measurable perfect matching in Γ_{V_C} (to be precise, the measurable almost perfect matching) is obtained by taking a limit of a sequence of measurable matchings that stabilises almost everywhere. These matchings are provided by an algorithm that improves our matchings by augmenting paths. The measurability of the matchings is an immediate consequence of the fact that our algorithm is local: whether there is an edge (a, b) in the i^{th} matching only depends on how the sets A, B and the graph $\Gamma_{V_C}(A, B)$ look like in the R_i neighbourhood of a (or b). Of course, $R_i \to \infty$.

The existence of this algorithm and the matchings rely on sufficient conditions on the discrepancy of the sets A and B in the cosets of V, similar to those needed by Laczkovich's proof. An extra ingredient that we need is the existence of short augmenting paths. We summarise these tools below.

Definition 5.4. Given $v = (v_1, \ldots, v_d) \in (\mathbb{T}^n)^d$, for $p \in \mathbb{Z}^d$, let

$$\langle v, p \rangle = v_1 p_1 + \ldots + v_d p_d \in \mathbb{T}^n.$$

For a set $P \subset \mathbb{Z}^d$, let

$$\langle v, P \rangle = \{\langle v, p \rangle \,:\, p \in P\} \subset \mathbb{T}^n.$$

When P is a product of intervals (i.e. sets of consecutive integers), we may refer to both P and $\langle v, P \rangle$ as a rectangle.

Lemma 5.5. *Fix any $\varepsilon > 0$. Let d be sufficiently large and let $v_1, \ldots, v_d$ be random independent uniformly chosen vectors in $\mathbb{T}^n$. Then, with probability 1, there is a positive integer C and $c > 0$ such that the following statements hold.*

1. *For any $x \in \mathbb{T}^n$ and any rectangle $R \subset \mathbb{Z}^d$ with maximal side length N,*

$$\Big| |A \cap (x + \langle v, R \rangle)| - |B \cap (x + \langle v, R \rangle)| \Big| \le c N^{d-1-\varepsilon}.$$

2. *Let R be a rectangle in $\mathbb{Z}^d$ with maximal side length N. Then, for every $x \in \mathbb{T}^n$, there is a matching inside $x + \langle v, R \rangle$, that is, inside*

$$\Gamma_{V_C}(A \cap (x + \langle v, R \rangle),\ B \cap (x + \langle v, R \rangle))$$

such that at most $cN^{d-1-\varepsilon}$ points are unmatched.

3. *Let Q be any cube in $\mathbb{Z}^d$ of side length N. Let $\mathfrak{M}$ be a matching inside $x + \langle v, Q \rangle$. If there are points both in A and B that are unmatched by $\mathfrak{M}$ inside $x + \langle v, Q \rangle$ then there is an augmenting path connecting two unmatched points of length at most N.*

Proofs of these (or similar) statements can be found in Grabowski, Máthé, and Pikhurko [2017]. All these are generalizations and strengthenings of statements of Laczkovich [1992b]. The third statement uses the fact that not only Hall's condition

$$|N(X)| \geq |X| \quad (X \subset A \text{ finite})$$

holds in the graph, but it holds relative to a large enough cube, moreover, it can be replaced (essentially) by the stronger inequality

$$|N(X)| \geq |X| + c'|\partial X| \geq |X| + |X|^{\frac{d-1}{d}}. \tag{5-1}$$

Here ∂X can be understood as those points x of X for which $x + V_1 \not\subset X$. Note that the exponent $(d-1)/d$ is optimal by the isoperimetric inequality. (To be correct, ∂X in (5-1) should be replaced by boundary of a smoothened version of X.)

Corollary 5.6. *Let Q be any cube in $\mathbb{Z}^d$ of side length N. Let M be a matching inside $x + \langle v, Q\rangle$ such that the number of unmatched points is t. Using augmenting paths we can define a new matching M' such that only $cN^{d-1-\varepsilon}$ points will be unmatched and that $|M \triangle M'| \leq tN$.*

Proof of Corollary 5.6. We can improve the matching by augmenting paths; combining part Item 2. and Item 3. of Lemma 5.5 concludes the proof. □

6 Open problems

Question 6.1 (Borel Banach–Tarski)**.** *Let $n \geq 3$. Let $A, B \in \mathbb{R}^n$ be bounded Borel sets of non-empty interior of equal Lebesgue measure. Is A equidecomposable to B using Borel pieces?*

Note that the answer is affirmative if A and B have nice boundaries by the theorem of A. S. Marks and S. T. Unger [2017]; that is, if the boundaries have upper Minkowski dimension less than n. (On the other hand, for $n = 1, 2$ the answer is negative. Laczkovich [2003] gave examples of Jordan domains in the plane that are not even equidecomposable with arbitrary pieces.)

The disc and the square are equidecomposable using Borel pieces, in particular, with pieces that are both Lebesgue and Baire measurable. However, in some sense, the "true" combination of Lebesgue and Baire measurability is Jordan measurability. Recall that a bounded set is Jordan measurable if its boundary has Lebesgue measure zero.

Question 6.2. *Is it possible to equidecompose the disc to a square by Jordan measurable pieces?*

Note that the result of Dubins, Hirsch, and Karush [1963] says that the pieces cannot be Jordan domains.

Acknowledgments. I would like to thank my co-authors Łukasz Grabowski and Oleg Pikhurko that I could work with them on measurable equidecompositions. I would like to thank Gábor Kun and Miklós Laczkovich for inspiring discussions. I am grateful to Oleg Pikhurko for comments on earlier versions of this manuscript.

References

Stefan Banach (1923). "Sur le problème de la mesure". *Fundamenta Mathematicae* 4.1, pp. 7–33 (cit. on p. 1732).

Stefan Banach and Alfred Tarski (1924). "Sur la décomposition des ensembles de points en parties respectivement congruentes". *Fund. math* 6.1, pp. 244–277 (cit. on p. 1732).

Bachir Bekka, Pierre de la Harpe, and Alain Valette (2008). *Kazhdan's property (T)*. Vol. 11. New Mathematical Monographs. Cambridge University Press, Cambridge, pp. xiv+472. MR: 2415834.

Rémi Boutonnet, Adrian Ioana, and Alireza Salehi Golsefidy (2017). "Local spectral gap in simple Lie groups and applications". *Invent. Math.* 208.3, pp. 715–802. MR: 3648974 (cit. on p. 1741).

Randall Dougherty and Matthew Foreman (1992). "Banach-Tarski paradox using pieces with the property of Baire". *Proc. Nat. Acad. Sci. U.S.A.* 89.22, pp. 10726–10728. MR: 1190902 (cit. on p. 1733).

– (1994). "Banach-Tarski decompositions using sets with the property of Baire". *J. Amer. Math. Soc.* 7.1, pp. 75–124. MR: 1227475 (cit. on p. 1733).

V. G. Drinfel'd (1984). "Finitely-additive measures on S^2 and S^3, invariant with respect to rotations". *Funktsional. Anal. i Prilozhen.* 18.3, p. 77. MR: 757256 (cit. on pp. 1734, 1738).

Lester Dubins, Morris W. Hirsch, and Jack Karush (1963). "Scissor congruence". *Israel J. Math.* 1, pp. 239–247. MR: 0165424 (cit. on pp. 1735, 1748).

Gábor Elek and Gábor Lippner (2010). "Borel oracles. An analytical approach to constant-time algorithms". *Proc. Amer. Math. Soc.* 138.8, pp. 2939–2947. MR: 2644905 (cit. on pp. 1734, 1739).

R. J. Gardner (1985). "Convex bodies equidecomposable by locally discrete groups of isometries". *Mathematika* 32.1, pp. 1–9. MR: 817100 (cit. on p. 1735).

Łukasz Grabowski, András Máthé, and Oleg Pikhurko (Aug. 2014). "Measurable equidecompositions via combinatorics and group theory". arXiv: 1408.1988 (cit. on p. 1737).

– (Jan. 2016). "Measurable equidecompositions for group actions with an expansion property". arXiv: 1601.02958 (cit. on pp. 1734, 1737, 1741).

– (2017). "Measurable circle squaring". *Ann. of Math. (2)* 185.2, pp. 671–710. MR: 3612006 (cit. on pp. 1736, 1745, 1748).

M. Laczkovich (1990). "Equidecomposability and discrepancy; a solution of Tarski's circle-squaring problem". *J. Reine Angew. Math.* 404, pp. 77–117. MR: 1037431 (cit. on p. 1735).

– (2003). "Equidecomposability of Jordan domains under groups of isometries". *Fund. Math.* 177.2, pp. 151–173. MR: 1992530 (cit. on p. 1748).

Miklós Laczkovich (1992a). "Decomposition of sets with small boundary". *J. London Math. Soc. (2)* 46.1, pp. 58–64. MR: 1180882 (cit. on p. 1735).

– (1992b). "Uniformly spread discrete sets in $\mathbf{R}^d$". *J. London Math. Soc. (2)* 46.1, pp. 39–57. MR: 1180881 (cit. on pp. 1743, 1748).

– (1993). "Decomposition of sets of small or large boundary". *Mathematika* 40.2, pp. 290–304. MR: 1260893 (cit. on p. 1735).

– (1996). "Decomposition using measurable functions". *C. R. Acad. Sci. Paris Sér. I Math.* 323.6, pp. 583–586. MR: 1411046 (cit. on p. 1744).

– (2002). "Paradoxes in measure theory". In: *Handbook of measure theory, Vol. I, II*. North-Holland, Amsterdam, pp. 83–123. MR: 1953492 (cit. on pp. 1737, 1744).

Russell Lyons and Fedor Nazarov (2011). "Perfect matchings as IID factors on non-amenable groups". *European J. Combin.* 32.7, pp. 1115–1125. MR: 2825538 (cit. on pp. 1734, 1739).

G. A. Margulis (1980). "Some remarks on invariant means". *Monatsh. Math.* 90.3, pp. 233–235. MR: 596890 (cit. on pp. 1734, 1738).

Andrew S. Marks and Spencer T. Unger (2017). "Borel circle squaring". *Ann. of Math. (2)* 186.2, pp. 581–605. MR: 3702673 (cit. on pp. 1736, 1746, 1748).

Andrew Marks and Spencer Unger (2016). "Baire measurable paradoxical decompositions via matchings". *Adv. Math.* 289, pp. 397–410. MR: 3439691.

Jan Mycielski (1979). "Finitely additive invariant measures. I". *Colloq. Math.* 42, pp. 309–318. MR: 567569 (cit. on p. 1733).

Heiner Niederreiter and Jörg M. Wills (1975). "Diskrepanz und Distanz von Maßen bezüglich konvexer und Jordanscher Mengen". *Math. Z.* 144.2, pp. 125–134. MR: 0376588 (cit. on p. 1743).

R. Rado (1949). "Factorization of even graphs". *Quart. J. Math., Oxford Ser.* 20, pp. 95–104. MR: 0030202 (cit. on pp. 1737, 1741).

Dennis Sullivan (1981). "For $n > 3$ there is only one finitely additive rotationally invariant measure on the n-sphere defined on all Lebesgue measurable subsets". *Bull. Amer. Math. Soc. (N.S.)* 4.1, pp. 121–123. MR: 590825 (cit. on pp. 1734, 1738).

Alfred Tarski (1925). "Probléme 38". *Fund. Math* 7, p. 381.

Grzegorz Tomkowicz and Stan Wagon (2016). *The Banach-Tarski paradox*. Second. Vol. 163. Encyclopedia of Mathematics and its Applications. With a foreword by Jan Mycielski. Cambridge University Press, New York, pp. xviii+348. MR: 3616119 (cit. on p. 1737).

Stan Wagon (1985). *The Banach-Tarski paradox*. Vol. 24. Encyclopedia of Mathematics and its Applications. With a foreword by Jan Mycielski. Cambridge University Press, Cambridge, pp. xvi+251. MR: 803509 (cit. on p. 1737).

Friedrich Wehrung (1992). "Injective positively ordered monoids. I, II". *J. Pure Appl. Algebra* 83.1, pp. 43–82, 83–100. MR: 1190444 (cit. on p. 1744).

Received 2018-03-05.

András Máthé
Mathematics Insitute
University of Warwick
Coventry CV4 7AL
United Kingdom
a.mathe@warwick.ac.uk

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1751–1770)

HIGHER ORDER COMMUTATORS AND MULTI-PARAMETER BMO

Stefanie Petermichl

Abstract

In this article we highlight the interplay of multi-parameter BMO spaces and boundedness of corresponding commutators. In a variety of settings, we discuss two-sided norm estimates for commutators of classical singular operators with a symbol function. In its classical form, this concerns a theorem by Nehari, factorisation of Hardy space, Hankel and Toeplitz forms. We highlight recent results in which a characterization of L^p boundedness of iterated commutators of multiplication by a symbol function and tensor products of Riesz and Hilbert transforms is obtained, completing a theory on characterisation of BMO spaces begun by Cotlar, Ferguson and Sadosky. In the light of real analysis, we discuss results in a more intricate situation; commutators of multiplication by a symbol function and Calderón-Zygmund or Journé operators. We show that the boundedness of these commutators is also determined by the inclusion of their symbol function in the same multi-parameter BMO class. In this sense the Hilbert or Riesz transforms or their tensor products are a representative testing class for Calderón-Zygmund or Journé operators.

1 Introduction

A classical result of Nehari Nehari [1957] studies L^2 boundedness of Hankel operators with anti-analytic symbol b mapping analytic functions into the space of anti-analytic functions by

$$H_b : f \mapsto P_{-}bf.$$

A BMO condition on the symbol characterises boundedness. This theorem has an equivalent formulation in terms of the boundedness of the commutator of the multiplication operator with symbol function b and the Hilbert transform

$$[H, b] = Hb - bH.$$

MSC2010: primary 42B35; secondary 47B35, 47B38, 30H35.
Keywords: Iterated commutator, Journé operator, multi-parameter BMO, Hankel operator, Toeplitz operator.

To see this correspondence one uses that up to a constant $H = P_+ - P_-$ and rewrites the commutator as a sum of Hankel operators with orthogonal ranges. One writes the two-sided inequality on the operator norm

$$\|b\|_{BMO} \lesssim \|[H,b]\|_{L^2 \to L^2} \lesssim \|b\|_{BMO}.$$

This two-sided estimate uses the classical factorisation into inner and outer functions. Notably, the lower commutator estimate relies heavily on its corollary, a factorisation theorem of functions in the complex Hardy space H^1 into a product of two H^2 functions. Here is a sketch of the argument showing necessity and sufficiency of a BMO condition for the boundedness of H_b.

$$\begin{aligned}
\|H_b\| &= \sup_{\|g\|_{H^2_-}=1} \sup_{\|f\|_{H^2}=1} |(H_b f, g)| \\
&= \sup_{\|g\|_{H^2_-}=1} \sup_{\|f\|_{H^2}=1} |(P_-(((P_- + P_+)b)f), g)| \\
&= \sup_{\|g\|_{H^2_-}=1} \sup_{\|f\|_{H^2}=1} |(P_-((P_- b)f), g)| \\
&= \sup_{\|g\|_{H^2_-}=1} \sup_{\|f\|_{H^2_+}=1} |((P_- b) f, g)| \\
&= \sup_{\|g\|_{H^2}=1} \sup_{\|f\|_{H^2}=1} |(P_- b, \bar{f} g)| \\
&= \sup_{\|g\|_{H^2}=1} \sup_{\|f\|_{H^2}=1} |(b, \bar{f} g)|.
\end{aligned}$$

Using $H^1 - BMO$ duality and factorisation of Hardy space for the necessity, we get the characterisation of BMO.

Let $H^2(\mathbb{T}^2)$ denote the Banach space of analytic functions in $L^2(\mathbb{T}^2)$. In Ferguson and Sadosky [2000], Ferguson and Sadosky study the symbols of bounded 'big' and 'little' Hankel operators on the bidisk. Big Hankel operators are those which project on to a 'big' subspace of $L^2(\mathbb{T}^2)$ - the orthogonal complement of $H^2(\mathbb{T}^2)$; while little Hankel operators project onto the smaller subspace of complex conjugates of functions in $H^2(\mathbb{T}^2)$ - or anti-analytic functions. The corresponding commutators are

$$[H_1 H_2, b],$$

and

$$[H_1, [H_2, b]]$$

where $b = b(x_1, x_2)$ and H_k are the Hilbert transforms acting in the k^{th} variable. Ferguson and Sadosky show that the first commutator is bounded if and only if the symbol b belongs to the so called little BMO class, consisting of those functions that are uniformly in BMO in each variable separately. Their argument is based on a classical fact on Toeplitz operators. They also show that if b belongs to the product BMO space, as identified by Chang and Fefferman Chang and Fefferman [1985], Chang and Fefferman [1980] then the second commutator is bounded. The fact that boundedness of the second commutator implies that b is in product BMO was shown in the groundbreaking paper of Ferguson and Lacey Ferguson and Lacey [2002]. The absence of factorisation theorems in this multi-parameter setting lead the authors to study two-sided commutator estimates - a very difficult task, considering the complicated structure of the product BMO space. The set up still has Hankel operators at heart, but the techniques to tackle this question in several parameters are very different and have brought valuable new insight and use to existing theories, for example in the interpretation of Journé's lemma Journé [1986] in combination with Carleson's example Carleson [1974]. Lacey and Terwilliger extended this result to an arbitrary number of iterates in Lacey and Terwilleger [2009], requiring thus, among others, a refinement of Pipher's iterated multi-parameter version Pipher [1986] of Journé's lemma. One can then deduce a weak factorisation theorem on the bi-disk. Commutators of the mixed type whose base case is for example

$$[H_1, [H_2 H_3, b]]$$

were considered by Ou, Strouse and the author in Yumeng, Petermichl, and Strouse [2016]. One classifies boundedness of these commutators by a little product BMO class: those functions $b = b(x_1, x_2, x_3)$ so that $b(\cdot, x_2, \cdot)$ and $b(\cdot, \cdot, x_3)$ are uniformly in product BMO. Similar results can be obtained for any finite iteration of any finite tensor product of Hilbert transforms. The proof for this Hilbert transform case is a simple application of Toeplitz operators, if one admits the work by Ferguson, Lacey and Terwilliger.

The main focus in this note however, is in the setting of real analysis, where Hankel and Toeplitz operators cannot be used as a tool.

When leaving the notion of Hankel operators behind, their interpretation as commutators allow for natural generalizations. Through the use of completely different real variable methods, Coifman, Rochberg and Weiss Coifman, Rochberg, and Weiss [1976] extended Nehari's one-parameter theory to real analysis in the sense that the Hilbert transforms were replaced by Riesz transforms. The missing features of the Riesz transforms include analytic projection on one hand as well as strong factorisation theorems of analytic function spaces.

The authors in Coifman, Rochberg, and Weiss [1976] obtained sufficiency, i.e. that a BMO symbol b yields an $L^2(\mathbb{R}^d)$ bounded commutator for certain more general, convolution type singular integral operators. For necessity, they showed that the collection of Riesz transforms was representative enough:

$$\|b\|_{\mathrm{BMO}} \lesssim \sup_{1\le j\le d} \|[R_j, b]\|_{2\to 2}.$$

Notably this lower bound was obtained somewhat indirectly through use of spherical harmonics in combination with the mean oscillation characterisation of BMO in one parameter.

These one-parameter results in Coifman, Rochberg, and Weiss [ibid.] were extended to the multi-parameter setting in the work by Lacey, Pipher, Wick and the author Lacey, Petermichl, Pipher, and Wick [2009]. Both the upper and lower estimate have proofs very different from those in one parameter. For the lower estimate, the methods in Ferguson and Lacey [2002] or Lacey and Terwilleger [2009] find an extension to real variables through operators closer to the Hilbert transform than the Riesz transforms (cone operators) and an indirect passage on the Fourier transform side.

In a recent paper Dalenc and Ou [2014] it is shown that iterated commutators formed with any arbitrary Calderón-Zygmund operators are bounded if the symbol belongs to product BMO.

Ou, Strouse and the author considered in Yumeng, Petermichl, and Strouse [2016] all generalisations of the base case

$$[R_{1,j_1}, [R_{2,j_2} R_{3,j_3}, b]],$$

where R_{k,j_k} are Riesz transforms of direction j_k acting in the k^{th} variable. We show necessity and sufficiency of the little product BMO condition when the R_{k,j_k} are allowed to run through all Riesz transforms by means of a two-sided estimate. While in the Hilbert transform case, Toeplitz operators with operator symbol arise naturally, using Riesz transforms in $\mathbb{R}^d$ as a replacement, there is an absence of analytic structure and tools relying on analytic projection or orthogonal spaces are not readily available. We again overcome part of this difficulty through the use of Calderón-Zygmund operators whose Fourier multiplier symbols are adapted to cones. In this situation, the Toeplitz forms create an additional difficulty which is overcome through an intermediate passage and the construction of a multi-parameter cone operator not of tensor product type.

Further it was shown in work by Holmes, Ou, Strouse, Wick and the author Yumeng, Petermichl, and Strouse [ibid.], Holmes, Petermichl, and Wick [2018] that the tensor products of Riesz transforms in the upper estimate can be replaced by Journé operators, these are singular integral operators of the product type.

Much like discussed in the base cases of the results Coifman, Rochberg, and Weiss [1976], Lacey, Petermichl, Pipher, and Wick [2009], boundedness of commutators involving Hilbert or Riesz transforms are a testing condition. If these commutators are bounded, the symbol necessarily belongs to a BMO, little BMO, product BMO or little product BMO. Then, iterated commutators using a much more general class than that of tensor products of Riesz transforms are also bounded: commutators withCalderón-Zygmund or Journé operators.

2 Aspects of Multi-Parameter Theory

This section contains some review on Hardy spaces in several parameters as well as some definitions and lemmas relevant to us.

2.1 Chang-Fefferman BMO. We describe the elements of product Hardy space theory, as developed by Chang and Fefferman as well as Journé. By this we mean the Hardy spaces associated with domains like the poly-disk or $\mathbb{R}^{\vec{d}} := \bigotimes_{s=1}^{t} \mathbb{R}^{d_s}$ for $\vec{d} = (d_1, \ldots, d_t)$. While doing so, we typically do not distinguish whether we are working on $\mathbb{R}^d$ or $\mathbb{T}^d$. In higher dimensions, the Hilbert transform is usually replaced by the collections of Riesz transforms.

The (real) one-parameter Hardy space $H^1_{\mathrm{Re}}(\mathbb{R}^d)$ denotes the class of functions with the norm

$$\sum_{j=0}^{d} \|R_j f\|_1$$

where R_j denotes the j^{th} Riesz transform or the Hilbert transform if the dimension is one. Here and below we adopt the convention that R_0, the 0^{th} Riesz transform, is the identity. This space is invariant under the one-parameter family of isotropic dilations, while the product Hardy space $H^1_{\mathrm{Re}}(\mathbb{R}^{\vec{d}})$ is invariant under dilations of each coordinate separately. That is, it is invariant under a t parameter family of dilations, hence the terminology 'multi-parameter' theory. One way to define a norm on $H^1_{\mathrm{Re}}(\mathbb{R}^{\vec{d}})$ is

$$\|f\|_{H^1} \sim \sum_{0 \le j_l \le d_l} \|\prod_{l=1}^{t} \mathrm{R}_{l,j_l} f\|_1 .$$

R_{l,j_l} is the Riesz transform in the j_l^{th} direction of the l^{th} variable, and the 0^{th} Riesz transform is the identity operator.

The dual of the real Hardy space $H^1_{\mathrm{Re}}(\mathbb{R}^{\vec{d}})^*$ is $\mathrm{BMO}(\mathbb{R}^{\vec{d}})$, the t-fold product BMO space. It is a theorem of S.-Y. Chang and R. Fefferman Chang and Fefferman [1985], Chang and Fefferman [1980] that this space has a characterization in terms of a product Carleson measure.

Define

$$\|b\|_{\mathrm{BMO}(\mathbb{R}^{\vec{d}})} := \sup_{U \subset \mathbb{R}^{\vec{d}}} \left(|U|^{-1} \sum_{R \subset U} \sum_{\vec{\varepsilon} \in \mathrm{sig}_{\vec{d}}} |(b, w_R^{\vec{\varepsilon}})|^2 \right)^{1/2} . \tag{1}$$

Here the supremum is taken over all open subsets $U \subset \mathbb{R}^{\vec{d}}$ with finite measure, and we use a wavelet basis $w_R^{\vec{\varepsilon}}$ adapted to rectangles $R = Q_1 \times \cdots \times Q_t$, where each Q_l is a cube. The superscript $\vec{\varepsilon}$ reflects the fact that multiple wavelets are associated to any dyadic cube, see Lacey, Petermichl, Pipher, and Wick [2009] for details. In this note most often we use the well known Haar wavelet basis. The fact that the supremum admits all open sets of finite measure cannot be omitted, as Carleson's example shows Carleson [1974]. This fact is responsible for some of the difficulties encountered when working with this space.

Theorem 2.1 (Chang, Fefferman). *We have the equivalence of norms*

$$\|b\|_{(H^1_{Re}(\mathbb{R}^{\vec{d}}))^*} \sim \|b\|_{BMO(\mathbb{R}^{\vec{d}})}.$$

That is, $BMO(\mathbb{R}^{\vec{d}})$ is the dual to $H^1_{Re}(\mathbb{R}^{\vec{d}})$.

This BMO norm is invariant under a t-parameter family of dilations. Here the dilations are isotropic in each parameter separately. See also Fefferman [1979] and Fefferman [1987].

2.2 Little BMO. Following Cotlar and Sadosky [1990] and Ferguson and Sadosky [2000], we review the space little BMO, often written as 'bmo'. A locally integrable function $b : \mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_s} \to \mathbb{C}$ is in bmo if and only if

$$\|b\|_{\mathrm{bmo}} = \sup_{\vec{Q} = Q_1 \times \cdots \times Q_s} |\vec{Q}|^{-1} \int_{\vec{Q}} |b(\vec{x}) - b_{\vec{Q}}| < \infty$$

Here the Q_k are d_k-dimensional cubes and $b_{\vec{Q}}$ denotes the average of b over $\vec{Q}$.

It is easy to see that this space consists of all functions that are uniformly in BMO in each variable separately. Let $\vec{x}_{\hat{v}} = (x_1, \ldots, x_{v-1}, \cdot, x_{v+1}, \ldots, x_s)$. Then $b(\vec{x}_{\hat{v}})$ is a function in x_v only with the other variables fixed. Its BMO norm in x_v is

$$\|b(\vec{x}_{\hat{v}})\|_{\mathrm{BMO}} = \sup_{Q_v} |Q_v|^{-1} \int_{Q_v} |b(\vec{x}) - b(\vec{x}_{\hat{v}})_{Q_v}| dx_v$$

and the little BMO norm becomes

$$\|b\|_{\mathrm{bmo}} = \max_v \{\sup_{\vec{x}_{\hat{v}}} \|b(\vec{x}_{\hat{v}})\|_{\mathrm{BMO}}\}.$$

On the bi-disk, this becomes

$$\|b\|_{\mathrm{bmo}} = \max\{\sup_{x_1} \|b(x_1, \cdot)\|_{\mathrm{BMO}}, \sup_{x_2} \|b(\cdot, x_2)\|_{\mathrm{BMO}}\},$$

the space discussed in Ferguson and Sadosky [ibid.]. All other cases are an obvious generalisation, at the cost of notational inconvenience.

2.3 Little product BMO. In this section we define a BMO space which is in between little BMO and product BMO. As mentioned in the introduction, we aim at characterising BMO spaces consisting for example of those functions $b(x_1, x_2, x_3)$ such that $b(x_1, \cdot, \cdot)$ and $b(\cdot, \cdot, x_3)$ are uniformly in product BMO in the remaining two variables.

Definition 2.2. *Let $b : \mathbb{R}^{\vec{d}} \to \mathbb{C}$ with $\vec{d} = (d_1, \cdots, d_t)$. Take a partition $\mathfrak{I} = \{I_s : 1 \le s \le l\}$ of $\{1, 2, ..., t\}$ so that $\dot{\cup}_{1 \le s \le l} I_s = \{1, 2, ..., t\}$. We say that $b \in BMO_{\mathfrak{I}}(\mathbb{R}^{\vec{d}})$ if for any choices $\mathbf{v} = (v_s)$, $v_s \in I_s$, b is uniformly in product BMO in the variables indexed by v_s. We call a BMO space of this type a 'little product BMO'. If for any $\vec{x} = (x_1, ..., x_t) \in \mathbb{R}^{\vec{d}}$, we define $\vec{x}_{\hat{\mathbf{v}}}$ by removing those variables indexed by v_s, the little product BMO norm becomes*

$$\|b\|_{BMO_{\mathfrak{I}}} = \max_{\mathbf{v}} \{\sup_{\vec{x}_{\hat{\mathbf{v}}}} \|b(\vec{x}_{\hat{\mathbf{v}}})\|_{BMO}\}$$

where the BMO norm is product BMO in the variables indexed by v_s.

When $\vec{d}$ and $\vec{s}$ have dimension one, the definition recovers that of little BMO. When $\vec{d}$ and $\vec{s}$ have dimension $t > 1$ and $\vec{s} = \vec{1}$, then we recover the t-parameter product BMO space in $\mathbb{R}^{\vec{d}}$. The following simple example captures the essence of the intermediary spaces: $\mathrm{BMO}_{(1,1)\cdot(2,1)}$ is a class of functions defined on $(\mathbb{R}^1 \times \mathbb{R}^1) \times (\mathbb{R}^1)$ and is uniformly in two-parameter product BMO in variables 1 and 3 as well as 2 and 3.

3 Upper Bounds

In this section we describe upper norm estimates for commutators in terms of BMO norms of their symbol.

3.1 Hilbert transform. The easiest such estimate is

$$\|[H,b]\|_{2\to 2} \le C\|b\|_{BMO}.$$

There are very simple proofs of this fact, using the projection structure of the Hilbert transform. Let us revisit a different proof using the seminal idea of Haar shift, a strategy started by the author in Petermichl [2000] to address a question by Pisier on the dimensional growth of Hankel operators with matrix symbol. We will see that this proof restricted to the scalar case enjoys the generalisations we are seeking. For historic reasons we detail the object in its original form.

We will be using a variety of dyadic grids in $\mathbb{R}$. The standard dyadic grid, starting at 0 with intervals of length $1\cdot 2^n$, will be called $\mathfrak{D}^{0,1}$.

$$\mathfrak{D}^{0,1} = \{2^{-k}([0,1)+m) : k,m \in \mathbb{Z}\}.$$

Then $h_J^{0,1}$ is the Haar function for $J \in \mathfrak{D}^{0,1}$, namely

$$h_J^{0,1} = 1/\sqrt{|J|}\,(\chi_{J_-} - \chi_{J_+})$$

where J_- is the left half of J and J_+ the right half of J. We obtain a variation of $\mathfrak{D}^{0,1}$ by first shifting the starting point 0 to $\alpha \in \mathbb{R}$ and secondly choosing intervals of length $r\cdot 2^n$ for positive r. The resulting grid is called $\mathfrak{D}^{\alpha,r}$, and the corresponding Haar functions $h^{\alpha,r}$ are chosen so that they are still normalized in L^2. We often omit the indices α, r in our notations for the Haar functions. For $f \in L^2(\mathbb{R})$ we have

$$f(x) = \sum_{I\in\mathfrak{D}^{\alpha,r}} (f,h_I)h_I(x) \quad \forall \alpha \in \mathbb{R},\ r>0.$$

We define for such α, r a dyadic shift operator $S^{\alpha,r}$ by

$$(S^{\alpha,r}f)(x) = \sum_{I\in\mathfrak{D}^{\alpha,r}} (f,h_I)(h_{I_-}(x) - h_{I_+}(x)).$$

Its L^2 operator norm is $\sqrt{2}$ and its representing kernel is

$$K^{\alpha,r}(t,x) = \sum_{I \in \mathfrak{D}^{\alpha,r}} h_I(t)(h_{I_-}(x) - h_{I_+}(x)). \tag{2}$$

Through elementary methods one can show (see Petermichl [ibid.])

Lemma 3.1. *For $x \neq t$ let*

$$K(t,x) = \lim_{L\to\infty} \frac{1}{2\log L} \int_{1/L}^{L} \lim_{R\to\infty} \frac{1}{2R} \int_{-R}^{R} K^{\alpha,r}(t,x)\, d\alpha \frac{dr}{r}.$$

The limits exist pointwise and the convergence is bounded for $|x-t| \geq \delta$ for every $\delta > 0$ and $K(t,x) = c_0/(t-x)$ for some $c_0 > 0$.

Notice that $1/(t-x)$ is, up to a constant, the kernel of the Hibert transform. This fundamental lemma allows one to estimate commutators with Haar shifts instead of the Hilbert transform. The latter is the correct tool to capture the cancellation of the commutator.

We show that for all $\alpha \in \mathbb{R}$ and for all $r > 0$

$$\|S^{\alpha,r}b - bS^{\alpha,r}\|_{L^2\to L^2} \leq C\|b\|_{BMO}. \tag{3}$$

In the following α, r will be omitted because all estimates do not depend on the dyadic grid. Consider formally

$$b(x) = \sum_{I\in\mathfrak{D}} (b, h_I)h_I(x) \tag{4}$$

and

$$f(x) = \sum_{I\in\mathfrak{D}} (f, h_I)h_I(x). \tag{5}$$

By multiplying the sums (4) and (5) formally one gets $bf = A_b(f) + \Pi_b(f) + R_b(f)$, where

$$A_b(f) = \sum_{I\in\mathfrak{D}} (b, h_I)(f, h_I)h_I^2$$

$$\Pi_b(f) = \sum_{I \in \mathfrak{D}} (b, h_I)\langle f \rangle_I \, h_I$$
$$R_b(f) = \sum_{I \in \mathfrak{D}} \langle b \rangle_I (f, h_I) h_I$$

The expressions can be made meaningful in a standard way. Hence

$$Sb - bS = SA_b - A_b S + S\Pi_b - \Pi_b S + SR_b - R_b S$$

and we can estimate the terms separately.

The term Π_b is a paraproduct with symbol b and $\|\Pi_b\|_{L^2 \to L^2} \leq C\|b\|_{BMO}$. Also $A_B^* = \Pi_{B^*}$, so this term is bounded. We estimate the last term as commutator, noting that

$$SR_b f - R_b Sf = 1/2 \sum_I (\langle b \rangle_{I_+} - \langle b \rangle_{I_-})(f, h_I)(h_{I_-} - h_{I_+}).$$

Observe that $|(\langle b \rangle_{I_+} - \langle b \rangle_{I_-})| \lesssim \|b\|_{BMO}$. We have therefore shown that $\|[H, b]\|_{2 \to 2} \lesssim \|b\|_{BMO}$.

3.2 Calderón-Zygmund operators. The idea of Haar shift and representation theorems for singular operators has found deep generalisations. To obtain a proof of the estimate

$$\|[T, b]\|_{2 \to 2} \leq C\|b\|_{BMO}$$

with T a Calderón-Zygmund operator that will generalise to the iterated case, we use a famous theorem by Hytönen Hytönen [2012]. The original argument in Coifman, Rochberg, and Weiss [1976] does not generalise to the multi-parameter case.

Recall that a Calderón-Zygmund operator T acts on test functions and has a kernel representation for $x \notin \operatorname{supp} f$

$$Tf(x) = \int_{\mathbb{R}^d} K(x, y) f(y) dy.$$

Here the kernel K satisfies the standard estimates such as for example

$$|K(x, y)| \leq \frac{c_0}{|x - y|^d}$$

$$|K(x, y) - K(x', y)| + |K(y, x) - K(y, x')| \leq \frac{c_1}{|x - y|^d} \left(\frac{|x - x'|}{|x - y|} \right)^\delta$$

for all x, x' with $|x-y| > 2[x-x'|$ for some $0 < \delta \leq 1$. We say that T is bounded if in addition it acts boundedly in L^2.

To obtain a representation formula for T, consider instead of simple translates and dilates of the dyadic grid as in the Hilbert transform case, the randomised grid due to Nazarov, Treil, and Volberg [2003] with parameter $\omega \in (\{0,1\}^d)^{\mathbb{Z}}$ and

$$I \dot{+} \omega = I + \sum_{j:2^{-j}<\ell(I)} 2^{-j}\omega_j$$

where I belongs to the standard dyadic grid and $\ell(I)$ is the side length. The space is endowed with the natural probability measure.

A dyadic shift with parameters $i, j \in \mathbb{N}$ is an operator $Sf = \sum_{K\in\mathfrak{D}} A_K f$ where

$$A_K f = \sum_{I,J\in\mathfrak{D},I,J\subset K,\ell(I)=2^{-i}\ell(K),\ell(J)=2^{-j}\ell(K)} a_{IJK}(f,h_I)h_J$$

with coefficients $|a_{IJK}| \leq \frac{(|I||J|)^{1/2}}{|K|}$. It is called cancellative if all Haar functions in the representation are cancellative, otherwise non-cancellative.

Let T be a bounded Calderón-Zygmund operator. Then it was proved by Hytönen that it has an expansion for test functions f, g

$$(g, Tf) = c_T \mathbb{E}_\omega \sum_{i,j=0}^{\infty} \tau(i,j)(g, S_\omega^{i,j} f)$$

where $S_\omega^{i,j}$ is a dyadic shift of parameters i, j on the dyadic system $\mathfrak{D}^\omega$. Except possibly $S_\omega^{0,0}$ all are cancellative. τ has exponential decay with respect to the complexity parameters i, j with some dependence on the characteristics of the operator T.

This representation, along with careful consideration allows one to obtain the upper estimate

$$\|[T,b]\|_{2\to 2} \leq C\|b\|_{BMO}$$

through the use of paraproducts. The specificity of this proof is its applicability to the more difficult multi-parameter situation. One obtains the theorem below

Theorem 3.2. *(*Dalenc-Ou*) Let us consider $\mathbb{R}^{\vec{d}}$ with $\vec{d} = (d_1, \ldots, d_t)$. Let $b \in BMO$ and let T_s denote a Calderón-Zygmund operator acting on function defined on $\mathbb{R}^{d_s}$. Then we have the estimate*

$$\|[T_1, \ldots [T_l, b]\ldots]\|_{L^2(\mathbb{R}^{\vec{d}})\hookleftarrow} \lesssim \|b\|_{BMO},$$

where on the right hand side the product BMO norm stands.

This estimate was first proved under a restriction on the kernel by Coifman, Rochberg, and Weiss [1976] in the one-parameter case and by Lacey, Petermichl, Pipher, and Wick [2009] and the author in the multi-parameter case. Through the use of Haar shift this last proof was simplified considerably and restrictions on the kernel were removed by Dalenc and Ou [2014]. It is now known that this estimate also holds in L^p for $1 < p < \infty$. These recent proofs make use of multi-parameter paraproducts and estimates at their endpoint, considered by Journé Journé [1985] and Muscalu, Pipher, Tao, and Thiele [2004] and Muscalu, Pipher, Tao, and Thiele [2006].

For $j = 1, 2, 3$ let $\{\varphi_{j,R} \mid R \in \mathfrak{D}_{\vec{d}}\}$ be three families of functions adapted to the dyadic rectangles in $\mathfrak{D}_{\vec{d}}$ (we consider here products of cancellative or non-cancellative Haar functions, the actual theorems hold in greater generality). We say $\varphi_{j,R}$ has zero in a coordinate if the corresponding Haar function in that coordinate is cancellative. Then define

$$\mathrm{B}(f_1, f_2) := \sum_{R \in \mathfrak{D}_{\vec{d}}} \frac{(f_1, \varphi_{1,R})}{|R|^{1/2}} (f_2, \varphi_{2,R}) \varphi_{3,R}.$$

Theorem 3.3. *Assume that the family $\{\varphi_{1,R}\}$ has zeros in all coordinates. For every other coordinate s, assume that there is a choice of $j = 2, 3$ for which the the family $\{\varphi_{j,R}\}$ has zeros in the sth coordinate. Then the operator* B *enjoys the property*

$$\mathrm{B} : \mathrm{BMO} \times L^p \longrightarrow L^p, \qquad 1 < p < \infty.$$

3.3 Journé operators. To pass to the little BMO case, we observe that the generality of the upper estimate holds for Calderón-Zygmund operators of the multi-parameter type (or Journé operators).

The first generation of multi-parameter singular integrals that are not of tensor product type goes back to Fefferman [1981] and was generalised by Journé in Journé [1985] to the non-convolution type in the framework of his $T(1)$ theorem in this setting. We restrict ourselves for clarity to the bi-parameter case.

The class of bi-parameter singular integral operators treated in this section is that of any Journé type operator (not necessarily a tensor product and not necessarily of convolution type) satisfying a certain weak boundedness property, which we define as follows:

Definition 3.4. *A continuous linear mapping $T : C_0^\infty(\mathbb{R}^n) \otimes C_0^\infty(\mathbb{R}^m) \to [C_0^\infty(\mathbb{R}^n) \otimes C_0^\infty(\mathbb{R}^m)]'$ is called a bi-parameter Calderón-Zygmund operator if the following conditions are satisfied:*

1. T is a Journé type bi-parameter δ-singular integral operator, i.e. there exists a pair (K_1, K_2) of δCZ-δ-standard kernels so that, for all $f_1, g_1 \in C_0^\infty(\mathbb{R}^n)$ and $f_2, g_2 \in$

$C_0^\infty(\mathbb{R}^m)$,

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \int f_1(y_1)\langle K_1(x_1, y_1) f_2, g_2\rangle g_1(x_1)\, dx_1 dy_1$$

when $spt f_1 \cap spt g_1 = \varnothing$;

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \int f_2(y_2)\langle K_2(x_2, y_2) f_1, g_1\rangle g_2(x_2)\, dx_2 dy_2$$

when $spt f_2 \cap spt g_2 = \varnothing$.

2. T satisfies the weak boundedness property $|\langle T(\chi_I \otimes \chi_J), \chi_I \otimes \chi_J\rangle| \lesssim |I||J|$, for any cubes $I \subset \mathbb{R}^n$, $J \in \mathbb{R}^m$.

T is called paraproduct free if $T(1 \otimes \cdot) = T(\cdot \otimes 1) = T^*(1 \otimes \cdot) = T^*(\cdot \otimes 1) = 0$.

Recall that δCZ-δ-standard kernel is a vector valued standard kernel taking values in the Banach space consisting of all Calderón-Zygmund operators. It is easy to see that an operator defined as above satisfies all the characterizing conditions in Martikainen's paper Martikainen [2012], hence is L^2 bounded and can be represented as an average of bi-parameter dyadic shift operators together with dyadic paraproducts. This is the generalisation of Hytönen's theorem to the bi-parameter case. See also higher order Journé operators treated by Ou in Ou [2014]. To be precise, for test functions f, g, one has the following representation:

$$\langle Tf, g\rangle = C\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2} \sum_{i_1, j_1 = 0}^{\infty} \sum_{i_2, j_2 = 0}^{\infty} 2^{-\max(i_1, j_1)} 2^{-\max(i_2, j_2)} \langle S^{i_1 j_1 i_2 j_2} f, g\rangle. \tag{6}$$

where expectation is with respect to a certain parameter of the dyadic grids. The dyadic shifts $S^{i_1 j_1 i_2 j_2}$ are defined as

$$S^{i_1 j_1 i_2 j_2} f := \sum_{K_1 \in \mathfrak{D}_1} \sum_{\substack{I_1, J_1 \subset K_1, I_1, J_1 \in \mathfrak{D}_1 \\ \ell(I_1) = 2^{-i_1}\ell(K_1) \\ \ell(J_1) = 2^{-j_1}\ell(K_1)}} \sum_{K_2 \in \mathfrak{D}_2} \sum_{\substack{I_2, J_2 \subset K_2, I_2, J_2 \in \mathfrak{D}_2 \\ \ell(I_2) = 2^{-i_2}\ell(K_2) \\ \ell(J_2) = 2^{-j_2}\ell(K_2)}} a_{I_1 J_1 K_1 I_2 J_2 K_2} \langle f, h_{I_1} \otimes h_{I_2}\rangle h_{J_1} \otimes h_{J_2}$$

$$=: \sum_{K_1} \sum_{I_1, J_1 \subset K_1}^{(i_1, j_1)} \sum_{K_2} \sum_{I_2, J_2 \subset K_2}^{(i_2, j_2)} a_{I_1 J_1 K_1 I_2 J_2 K_2} \langle f, h_{I_1} \otimes h_{I_2}\rangle h_{J_1} \otimes h_{J_2}.$$

The coefficients above satisfy $a_{I_1 J_1 K_1 I_2 J_2 K_2} \le \frac{\sqrt{|I_1||J_1||I_2||J_2|}}{|K_1||K_2|}$, which also guarantees that $\|S^{i_1 j_1 i_2 j_2}\|_{L^2 \to L^2} \le 1$. Moreover, if T is paraproduct free, all the Haar functions appearing above are cancellative. The theorem below was proved partially in the author's work with Ou and Strouse Yumeng, Petermichl, and Strouse [2016] for the paraproduct-free case and in full generality as part of the author's work with Holmes and Wick in Holmes, Petermichl, and Wick [2018]. We obtained the estimate below.

Theorem 3.5. *Let us consider $\mathbb{R}^{\vec{d}}$, $\vec{d} = (d_1, \ldots, d_t)$ with a partition $\mathcal{I} = (I_s)_{1\leq s\leq l}$ of $\{1, \ldots, t\}$ as discussed before. Let $b \in BMO_{\mathcal{I}}(\mathbb{R}^{\vec{d}})$ and let T_s denote a multi-parameter Journé operator acting on function defined on $\bigotimes_{k\in I_s} \mathbb{R}^{d_k}$. Then we have the estimate*

$$\|[T_1, \ldots [T_l, b] \ldots]\|_{L^2(\mathbb{R}^{\vec{d}})\leftarrow} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^{\vec{d}})}.$$

The same estimate holds in L^p for $1 < p < \infty$.

This last estimate is more general than all previously mentioned commutator estimates.

4 Lower Bounds

In this section we bring a list of notable theorems under one roof. The theorem by Nehari in its formulation through a Hilbert transform commutator:

Theorem 4.1. *(*Nehari) There holds

$$\|b\|_{BMO} \lesssim \|[H, b]\|_{2\to 2} \lesssim \|b\|_{BMO}$$

as well as Ferguson and Sadosky's theorem on the commutator with the double Hilbert using the little BMO space:

Theorem 4.2. *(*Ferguson-Sadosky) There holds

$$\|b\|_{bmo} \lesssim \|[H_1 H_2, b]\|_{2\to 2} \lesssim \|b\|_{bmo}$$

as well as the iterated Hilbert commutators by Ferguson, Lacey, Terwilleger using product BMO:

Theorem 4.3. *(*Ferguson, Lacey, Terwilleger)There holds

$$\|b\|_{BMO} \lesssim \|[H_1, \ldots [H_t, b] \ldots]\|_{2\to 2} \lesssim \|b\|_{BMO}.$$

. In the real variable situation it includes the characterisation of Coifman, Rochberg and Weiss:

Theorem 4.4. *(*Coifman, Rochberg, Weiss) There holds

$$\|b\|_{BMO} \lesssim \sup_j \|[R_j, b]\|_{2\to 2} \lesssim \|b\|_{BMO}$$

It also includes the characterisation of Lacey, Pipher, Wick and the author using product BMO:

Theorem 4.5. *(*Lacey, Pipher, Petermichl, Wick) There holds

$$\|b\|_{BMO} \lesssim \sup_{\vec{j}} \|[R_{1,j_1}, \ldots [R_{t,j_t}, b] \ldots]\|_{2\to 2} \lesssim \|b\|_{BMO}.$$

To be precise, we prove a characterisation theorem of the space $\mathrm{BMO}_{\mathfrak{l}}(\mathbb{R}^{\vec{d}})$. We model the exposition after the formulation of the result by Ferguson and Sadosky.

Theorem 4.6. *(*Ferguson-Sadosky*) For $b \in L^1(\mathbb{T}^2)$ the following are equivalent with linear relations of their norms:*

(1) $b \in bmo$

(2) The commutators $[H_1, b]$ and $[H_2, b]$ are bounded on $L^2(\mathbb{T}^2)$

(3) The commutator $[H_2 H_1, b]$ is bounded on $L^2(\mathbb{T}^2)$.

Corollary 4.7. *(*Ferguson-Sadosky*) There is the equivalence of norms*

$$\|b\|_{bmo} \lesssim \|[H_1 H_2, b]\|_{2\to 2} \lesssim \|b\|_{bmo}.$$

The punch line in their beautiful argument is the use of Toeplitz forms. Indeed, typical terms of simple commutators, say with H_1 in this setting are of the form $P_{1,-} b P_{1,+}$ while the double commutator has typical terms of the form $P_{2,+} P_{1,-} b P_{1,+} P_{2,+}$. The norms of these are equal, when regarded as a Toeplitz operator with Hankel symbol. Further, the $L^\infty(BMO)$ characterisation arises naturally, admitting Nehari's theorem as a base.

This theorem in the iterated real variable setting and in its most general form reads as follows. See Yumeng, Petermichl, and Strouse [2016].

Theorem 4.8. *The following are equivalent with linear dependence in the respective norms.*

(1) $b \in BMO_{\mathfrak{l}}(\mathbb{R}^{\vec{d}})$

(2) All commutators of the form $[R_{k_1,j_{k_1}}, \ldots, [R_{k_l,j_{k_l}}, b] \ldots]$ are bounded in $L^2(\mathbb{R}^{\vec{d}})$ where $k_s \in I_s$ and $R_{k_s,j_{k_s}}$ is the one-parameter Riesz transform in direction j_{k_s}.

(3) All commutators of the form $[\vec{R}_{1,\vec{j}^{(1)}}, \ldots, [\vec{R}_{l,\vec{j}^{(l)}}, b] \ldots]$ are bounded in $L^2(\mathbb{R}^{\vec{d}})$ where $\vec{j}^{(s)} = (j_k)_{k \in I_s}$, $1 \le j_k \le d_k$ and the operators $\vec{R}_{s,\vec{j}^{(s)}}$ are a tensor product of Riesz transforms $\vec{R}_{s,\vec{j}^{(s)}} = \bigotimes_{k \in I_s} R_{k,j_k}$.

Corollary 4.9. *Let $\vec{j} = (j_1, \dots, j_t)$ with $1 \leq j_k \leq d_k$ and let for each $1 \leq s \leq l$, $\vec{j}^{(s)} = (j_k)_{k \in I_s}$ be associated a tensor product of Riesz transforms $\vec{R}_{s,\vec{j}^{(s)}} = \bigotimes_{k \in I_s} R_{k,j_k}$; here R_{k,j_k} are j_k^{th} Riesz transforms acting on functions defined on the k^{th} variable. We have the two-sided estimate*

$$\|b\|_{BMO_{\mathfrak{l}}(\mathbb{R}^{\vec{d}})} \lesssim \sup_{\vec{j}} \|[\vec{R}_{1,\vec{j}^{(1)}}, \dots, [\vec{R}_{t,\vec{j}^{(t)}}, b] \dots]\|_{L^2(\mathbb{R}^{\vec{d}}) \hookleftarrow} \lesssim \|b\|_{BMO_{\mathfrak{l}}(\mathbb{R}^{\vec{d}})}.$$

Such two-sided estimates also hold in L^p for $1 < p < \infty$.

We make some remarks about the strategy of the proof.

In the Hilbert transform case, Toeplitz operators with operator symbol arise naturally.

While Riesz transforms in $\mathbb{R}^d$ are a good generalisation of the Hilbert transform, there is absence of analytic structure and tools relying on analytic projection or orthogonal spaces are not readily available. We overcome this difficulty through a first intermediate passage via tensor products of Calderón-Zygmund operators whose Fourier multiplier symbols are adapted to cones. This idea is inspired by Lacey, Petermichl, Pipher, and Wick [2009].

A class of operators of this type classifies little product BMO through two-sided commutator estimates, but it does not allow the passage to a classification through iterated commutators with tensor products of Riesz transforms. In a second step, we find it necessary to consider upper and lower commutator estimates using a well-chosen family of Journé operators that are not of tensor product type. These operators are constructed to resemble the multiple Hilbert transform. A two-sided estimate of iterated commutators involving operators of this family facilitates a passage to iterated commutators with tensor products of Riesz transforms. There is an increase in difficulty when the arising tensor products involve more than two Riesz transforms and when the dimension is greater than two.

The actual passage to the Riesz transforms requires for us to prove a stability estimate in commutator norms for the multi-parameter singular integrals in terms of the mixed BMO class (see the section on upper bounds). In this context, we prove a qualitative upper estimate for iterated commutators using Journé operators.

To give a flavour of the argument, let us focus on $\mathbb{R}^2$ for simplicity. Riesz transforms and cone operators are homogeneous and their Fourier symbols are determined through their values on $\mathbb{S}^1$. Riesz transforms have the symbols of the coordinates, while cone operators have value 1 on an interval on the sphere covering less than half of the sphere and 0 else. The cone multipliers have to be mollified to result in Calderón-Zygmund operators, a fact we will omit. Through polynomial approximation, the symbols of (mollified) cone operators can be expressed via Riesz transforms symbols. One uses the simple fact

$$[AB, b] = A[B, b] + [A, b]B$$

to pass from lower cone transform estimates to lower Riesz transform estimates. This was one of the essential points in Lacey, Petermichl, Pipher, and Wick [ibid.]. The Toeplitz forms that arise in the tensor product case create an additional difficulty. Most polynomial representations, such as obtained when using tensor products of cone operators, are no longer enough. Other cone operators have to be considered that we try to describe.

Cone functions based on the two oblique strips containing $\vec{\xi}$ are averaged as illustrated below. The cone multiplier is 1 where the two oblique strips containing $\vec{\xi}$ intersect, it is 1/2 in sections with just one of the two strips and 0 else.

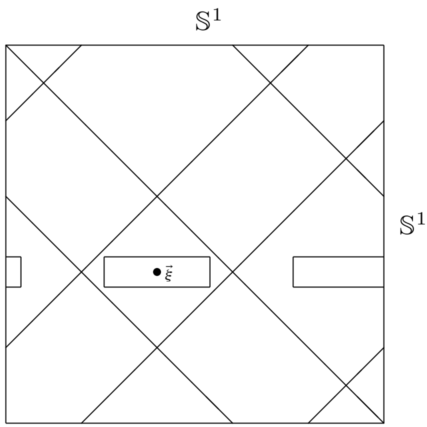


Figure 1: $\mathbb{R}^2 \times \mathbb{R}^2$

The rectangle around $\vec{\xi}$ with sides parallel to the axes illustrates the support of the tensor product of cone operators with direction $\vec{\xi}$. The longer side is the aperture that arises from the Hankel part Lacey, Petermichl, Pipher, and Wick [ibid.]. The short sides can be chosen freely as they arise from the Toeplitz part and is chosen small so that the rectangle fits into the oblique square. The other small rectangle corresponds to the Fourier support of the test function f.

This picture generalises to multiple copies of higher order spheres through the use of zonal harmonics and their identities. An averaging technique on products of spheres comes into play.

Using this intermediate tool, one can obtain lower commutator estimates with tensor products of Riesz transforms in accordance to the model of Ferguson and Sadosky in the Hilbert transform case.

5 Weak Factorization

It is well known that two-sided commutator estimates have an equivalent formulation in terms of *weak factorization* of Hardy space; indeed, this equivalence was important to the part of the proof of the two sided estimates of the iterated commutator. Let us recall the theorem of Lacey, Pipher, Wick and the author Lacey, Petermichl, Pipher, and Wick [2009].

Theorem 5.1. *We have the two-sided estimate*

$$\|b\|_{BMO} \lesssim \sup_{\vec{j}} \|[R_{1,j_1}, \dots [R_{t,j_t}, b] \dots]\|_{2\to 2} \lesssim \|b\|_{BMO}.$$

For $\vec{j}$ the vector above with $1 \le j_s \le d_s$, and $s = 1, \dots, t$, let $\Pi_{\vec{j}}$ be the bilinear operator defined by the following equation

$$\langle C_{\vec{j}}(b, f), g\rangle := \langle b, \Pi_{\vec{j}}(f, g)\rangle.$$

One can express $\Pi_{\vec{j}}$ as a linear combination of products of iterates of Riesz transforms, R_{s,j_s}, applied to the f and g. It follows immediately by duality from the two sided estimate for iterated Riesz commutators Lacey, Petermichl, Pipher, and Wick [ibid.] that for sequences $f_k^{\vec{j}}, g_k^{\vec{j}} \in L^2(\mathbb{R}^{\vec{d}})$ with $\sum_{\vec{j}} \sum_{k=1}^{\infty} \|f_k^{\vec{j}}\|_2 \|g_k^{\vec{j}}\|_2 < \infty$ we have

$$\sum_{\vec{j}} \sum_{k=1}^{\infty} \Pi_{\vec{j}}(f_k^{\vec{j}}, g_k^{\vec{j}}) \in H^1(\mathbb{R}^{\vec{d}}).$$

With this observation, we define

$$L^2(\mathbb{R}^{\vec{d}}) \widehat{\odot} L^2(\mathbb{R}^{\vec{d}}) := \left\{ f \in L^1(\mathbb{R}^{\vec{d}}) : f = \sum_{\vec{j}} \sum_{k=1}^{\infty} \Pi_{\vec{j}}(f_k^{\vec{j}}, g_k^{\vec{j}}) \right\}. \tag{7}$$

This is the projective tensor product given by

$$\|f\|_{L^2(\mathbb{R}^{\vec{d}}) \widehat{\odot} L^2(\mathbb{R}^{\vec{d}})} := \inf\left\{ \sum_{\vec{j}} \sum_{k} \|f_k^{\vec{j}}\|_2 \|g_k^{\vec{j}}\|_2 \right\}$$

where the infimum is taken over all decompositions of f as in (7). We have the following corollary.

We have $H^1(\mathbb{R}^{\vec{d}}) = L^2(\mathbb{R}^{\vec{d}}) \widehat{\odot} L^2(\mathbb{R}^{\vec{d}})$. Namely, for any $f \in H^1(\mathbb{R}^{\vec{d}})$ there exist sequences $f_k^{\vec{j}} \in L^2(\mathbb{R}^{\vec{d}})$ and $g_k^{\vec{j}} \in L^2(\mathbb{R}^{\vec{d}})$ such that

$$f = \sum_{\vec{j}} \sum_{k=1}^{\infty} \Pi_{\vec{j}}(f_k^{\vec{j}}, g_k^{\vec{j}})$$

with

$$\|f\|_{H^1} \simeq \sum_{\vec{j}} \sum_k \|f_k^{\vec{j}}\|_2 \|g_k^{\vec{j}}\|_2.$$

Similar results hold when replacing the exponent 2 by $1 < p < \infty$.

6 Div-Curl Lemma

Suppose $E, B \in L^2(\mathbb{R}^n, \mathbb{R}^n)$ are vector fields. It is immediate that their dot product $E \cdot B \in L^1(\mathbb{R}^n)$ with

$$\|E \cdot B\|_{L^1(\mathbb{R}^n)} \le \|E\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)} \|B\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)}.$$

If in addition they satisfy

$$\operatorname{div} E(x) = 0 \text{ and } \operatorname{curl} B(x) = 0,$$

then we have more cancellation: $E \cdot B \in H^1(\mathbb{R}^n)$ with

$$\|E \cdot B\|_{H^1(\mathbb{R}^n)} \lesssim \|E\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)} \|B\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)}.$$

Indeed, this fact is included in the paper by Coifman, Lions, Meyer, and Semmes [1993]. We sketch their elegant proof, using an upper Riesz commutator estimate. There exists a function ϕ such that $B_j = R_j \phi$ with $\|B\|_{L^2(\mathbb{R}^n, \mathbb{R}^n)} \sim \|\phi\|_{L^2(\mathbb{R}^n)}$. We then have point wise

$$E \cdot B = \sum_j E_j B_j = \sum_j E_j R_j \phi + \phi R_j E_j - \phi R_j E_j = \sum_j E_j R_j \phi + \phi R_j E_j.$$

The last equality is due to E being divergence free and $\sum_j R_j E_j = 0$. Now test this equality over $b \in BMO$ and obtain

$$(E \cdot B, b) = \sum_j (b, E_j R_j \phi + \phi R_j E_j) = \sum_j ([b, R_j](E_j), \phi).$$

Thanks to the BMO condition we know that $[b, R_j]$ is bounded. Thus

$$|(E \cdot B, b)| \lesssim \|b\|_{BMO} \|E\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)} \|\phi\|_{L^2(\mathbb{R}^n)} \sim \|b\|_{BMO} \|E\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)} \|B\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)}.$$

The duality between H^1 and BMO then yields

$$\|E \cdot B\|_{H^1(\mathbb{R}^n)} \lesssim \|E\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)} \|B\|_{L^2(\mathbb{R}^n;\mathbb{R}^n)}.$$

There are several possible generalisations of this result to the multi-parameter case. See Lacey, Pipher, Wick and the author in Lacey, Petermichl, Pipher, and Wick [2012]. We state one possible generalisation, that uses the upper estimate of the iterated Riesz commutator in terms of product BMO.

Theorem 6.1. *Suppose $E \in L^2(\mathbb{R}^n \times \mathbb{R}^n; \mathbb{R}^n)$ and $B \in L^2(\mathbb{R}^n \times \mathbb{R}^n; \mathbb{R}^n)$ have*

$$div_x E(x, y) = 0 \quad curl_x B(x, y) = 0 \;\; \forall y \in \mathbb{R}^n$$

and

$$div_y E(x, y) = 0 \quad curl_y B(x, y) = 0 \;\; \forall x \in \mathbb{R}^n.$$

Then

$$\int_{\mathbb{R}^n} \|E(x, \cdot) \cdot B(x, \cdot)\|_{H^1} dx \lesssim \|E\|_{L^2} \|B\|_{L^2}$$

and

$$\int_{\mathbb{R}^n} \|E(\cdot, y) \cdot B(\cdot, y)\|_{H^1} dy \lesssim \|E\|_{L^2} \|B\|_{L^2}.$$

References

Lennart Carleson (1974). "A counterexample for measures bounded on H^p spaces for the bidisk". *Mittag-Leffler Rep. No. 7, Inst. Mittag-Leffler* (cit. on pp. 1753, 1756).

Sun-Yung Chang and Robert Fefferman (1980). "A continuous version of duality of H^1 with BMO on thebidisc". *Ann. of Math. (2)* 112.1, pp. 179–201 (cit. on pp. 1753, 1756).

– (1985). "Some recent developments in Fourier analysis and H^p-theoryon product domains". *Bull. Amer. Math. Soc. (N.S.)* 12.1, pp. 1–43 (cit. on pp. 1753, 1756).

R. Coifman, P.L. Lions, Y Meyer, and S Semmes (1993). "Compensated Compactness and Hardy Spaces". *J. Math. Pures Appl. (9)* 72.3, pp. 247–286 (cit. on p. 1769).

Ronald Coifman, Richard Rochberg, and Guido Weiss (1976). "Factorization theorems for Hardy spaces in several variables". *Ann. of Math. (2)* 103.3, pp. 611–635 (cit. on pp. 1753–1755, 1760, 1762).

Misha Cotlar and Cora Sadosky (1990). "The Helson-Szegö theorem in L^p of the bidimensional torus". *Contemp. Math.* 107, pp. 19–37 (cit. on p. 1756).

Laurent Dalenc and Yumeng Ou (2014). "Upper bound for multi-parameter iterated commutators". *Preprint* (cit. on pp. 1754, 1762).

Robert Fefferman (1979). "Bounded mean oscillation on the polydisk". *Ann. of Math. (2)* 110.2, pp. 395–406 (cit. on p. 1756).

– (1981). "Singular integrals on product domains". *Bull. Amer. Math. Soc. (N.S.)* 4.2, pp. 195–201 (cit. on p. 1762).

– (1987). "Harmonic analysis on product spaces". *Ann. of Math. (2)* 126.1, pp. 109–130 (cit. on p. 1756).

Sarah Ferguson and Michael Lacey (2002). "A characterisation of product BMO by commutators". *Acta Math.* 189.2, pp. 143–160 (cit. on pp. 1753, 1754).

Sarah Ferguson and Cora Sadosky (2000). "Characterisations of bounded mean oscillation on the polydisk interms of Hankel operators and Carleson measures". *J. Anal. Math.* 81, pp. 239–267 (cit. on pp. 1752, 1756, 1757).

Irina Holmes, Stefanie Petermichl, and Brett Wick (2018). "Weighted little bmo and two-weight inequalities for Journé commutators". *Analysis and PDE* (cit. on pp. 1754, 1764).

Tuomas Hytönen (2012). "The sharp weighted bound for general Calderón-Zygmund operators". *Ann. of Math.* 175.3, pp. 1473–1506 (cit. on p. 1760).

Jean-Lin Journé (1985). "Calderón-Zygmund operators on product spaces". *Rev. Mat. Iberoamericana* 1.3, pp. 55–91 (cit. on p. 1762).

– (1986). "A covering lemma for product spaces". *Proc. Amer. Math. Soc.* 96.4, pp. 593–598 (cit. on p. 1753).

Michael Lacey, Stefanie Petermichl, Jill Pipher, and Brett Wick (2009). "Multiparameter Riesz commutators". *Amer. J. Math.* 131.3, pp. 731–769 (cit. on pp. 1754–1756, 1762, 1766–1768).

– (2012). "Multi-parameter Div-Curl Lemmas". *Bull. Lond. Math. Soc.* 10, pp. 1123–1131 (cit. on p. 1770).

Michael Lacey and Erin Terwilleger (2009). "Hankel operators in several complex variables and productBMO". *Houston J. Math.* 35.1, pp. 159–183 (cit. on pp. 1753, 1754).

Henri Martikainen (2012). "Representation of bi-parameter singular integrals by dyadic operators". *Adv. Math.* 229.3, pp. 1734–1761 (cit. on p. 1763).

Camil Muscalu, Jill Pipher, Terence Tao, and Christoph Thiele (2004). "Bi-parameter paraproducts". *Acta Math.* 193.2, pp. 269–296 (cit. on p. 1762).

– (2006). "Multi-parameter paraproducts". *Rev. Mat. Iberoam.* 22.3, pp. 963–976 (cit. on p. 1762).

Fedor Nazarov, Sergei Treil, and Alexander Volberg (2003). "The Tb theorem on non-homogenous spaces". *Acta Math.* 190.2, pp. 151–239 (cit. on p. 1761).

Zeev Nehari (1957). "On bounded bilinear forms". *Ann. of Math. (2)* 65, pp. 153–162 (cit. on p. 1751).

Yumeng Ou (2014). "Multi-parameter singular integral operators and representation theorem". arXiv: 1410.8055 (cit. on p. 1763).

Stefanie Petermichl (2000). "Dyadic Shifts and a Logarithmic Estimate for Hankel Operators with Matrix Symbol". *Comptes Rendus Acad. Sci. Paris* 1.1, pp. 455–460 (cit. on pp. 1758, 1759).

Jill Pipher (1986). "Journé's covering lemma and its extension to higherdimensions". *Duke Math. J.* 53.3, pp. 683–690 (cit. on p. 1753).

Ou Yumeng, Stefanie Petermichl, and Elizabeth Strouse (2016). "Higher Order Journe Commutators and multi-parameter BMO". *Adv. Math.* 291, pp. 24–58 (cit. on pp. 1753, 1754, 1764, 1765).

Received 2017-11-30.

STEFANIE PETERMICHL
stefanie.petermichl@gmail.com

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1771–1796)

TOEPLITZ METHODS IN COMPLETENESS AND SPECTRAL PROBLEMS

Alexei Poltoratski

Abstract

We survey recent progress in the gap and type problems of Fourier analysis obtained via the use of Toeplitz operators in spaces of holomorphic functions. We discuss applications of such methods to spectral problems for differential operators.

1 Introduction

This note is devoted to a discussion of several well-known problems from the area of the Uncertainty Principle in Harmonic Analysis (UP) and their reformulations in terms of kernels of Toeplitz operators. The Toeplitz approach first appeared in a series of papers by Hruščёv, Nikolskii, and Pavlov [1981], Nikolskii [1986], and Pavlov [1979], where it was applied to problems on Riesz bases and sequences of reproducing kernels in model spaces. It was extended to problems on completeness and spectral analysis of differential operators in Makarov and Poltoratski [2005, 2010] and used in several recent papers in the area.

The Beurling-Malliavin (BM) theory, created in the early 1960s to solve the problem on completeness of exponential functions in $L^2([0,1])$ Beurling and Malliavin [1962], Beurling and Malliavin [1967], and Beurling [1989], remains one of the deepest ingredients of UP. Although the theory did solve the classical problem it aimed to solve, the need for expansions to broader classes of function spaces, other systems of functions and related problems immediately appeared. At present, most of such problems remain open and suitable extensions of BM theory are yet to be found.

A search for such an extension in the settings of completeness problems and spectral problems for differential operators served as initial motivation for Makarov and Poltoratski [2005, 2010]. The first paper contained a list of problems of UP which could be translated into problems on injectivity of Toeplitz operators. It included problems on completeness

The author is partially supported by NSF Grant DMS-1665264.
MSC2010: 42A38.

of Airy and Bessel functions and spectral problems for 1D Schrödinger equation. The main result of the second paper gives a condition for the triviality of a kernel of a Toeplitz operator with a unimodular symbol, which can be viewed as an extension of the BM theorem.

In this note we will discuss the BM problem as well as the so-called gap and type problems of Fourier analysis, along with their reformulations in the Toeplitz language. We will show how such reformulations become a part of a new circle of problems on the partial ordering relations for the set of meromorphic inner functions (MIF) induced by Toeplitz operators. At the end of the paper we include applications to inverse spectral problems for Krein's canonical systems.

2 Completeness problems and spectral gaps

2.1 Beurling-Malliavin, gap and type problems. One of the canonical questions of Harmonic Analysis is whether any function from a given space can be approximated by linear combinations of functions from a selected collection of harmonics. The most common choices for the space are weighted L^p-spaces or spaces with weighted uniform norm of Bernstein's type. The role of harmonics can be played by polynomials, complex exponentials, solutions to various differential equations or special functions, such as Airy or Bessel functions, etc. A system of functions is called complete if finite linear combinations of its functions are dense in the space.

Let $\Lambda = \{\lambda_n\}$ be a discrete (without finite accumulation points) sequence of complex numbers. Denote by $\mathcal{E}_\Lambda$ the system of complex exponentials with frequencies from Λ:

$$\mathcal{E}_\Lambda = \{e^{i\lambda_n z},\ \lambda_n \in \Lambda\}.$$

The main question answered by BM theory is for what Λ will $\mathcal{E}_\Lambda$ be complete in $L^2([0,a])$. More precisely, let us define *the radius of completeness* of Λ as

$$R(\Lambda) = \sup\{a \mid \mathcal{E}_\Lambda \text{ is complete in } L^2([0,a])\}$$

and as 0 if the set is empty. BM theory provided a formula for $R(\Lambda)$ in terms of the exterior density of the sequence, defined as follows.

A sequence of disjoint intervals I_n on the real line is called long if

$$\sum \frac{|I_n|^2}{1+\operatorname{dist}^2(0,I_n)} = \infty, \tag{2-1}$$

and short otherwise. For a discrete sequence of real points Λ we define its exterior (BM) density as

$$D(\Lambda) = \sup\{d \mid \exists \text{ long sequence } \{I_n\} \text{ such that },\ \forall n,\ \#(\Lambda \cup I_n) > d|I_n|\}.$$

If Λ is a complex sequence, assuming without loss of generality that it has no purely imaginary points, we define $D(\Lambda)$ to be the density of the real sequence $1/(\Re\frac{1}{\lambda_n})$.

Theorem 1 (Beurling and Malliavin [1962] and Beurling and Malliavin [1967])**.**

$$R(\Lambda) = 2\pi D(\Lambda).$$

Since its appearance in the early 1960s, the BM theorem above and several ingredients of its proof had major impact on Harmonic Analysis. At present, new applications of the theorem continue to emerge in adjacent fields including Fourier analysis and spectral theory. At the same time, the search for generalizations of the BM theorem to other function spaces, several variables, other families of functions, etc., still continues with most of such problems remaining open.

A similar, and in some sense 'dual' completeness problem, is the so-called type problem which can be formulated as follows. For $a > 0$ denote by $\mathcal{E}_a$ the system of exponentials

$$\mathcal{E}_a = \{e^{isz},\ s \in [-a, a]\}.$$

For a finite positive Borel measure μ on $\mathbb{R}$ the exponential type of μ is defined as

$$\mathbf{T}_\mu = \inf\{a \mid \mathcal{E}_a \text{ is complete in } L^2(\mu)\}.$$

The problem of finding $\mathbf{T}_\mu$ in terms of μ appears in several adjacent areas of analysis and was studied by N. Wiener, A. Kolmogorov and M. Krein in connection with prediction theory and spectral problems for differential operators. For further information on such connections and problem's history see for instance Borichev and Sodin [2011] and Poltoratski [2015b]. We will return to the type problem in Section 2.5.

If the system $\mathcal{E}_a$ is incomplete in $L^p(\mu),\ p > 1$ then there exists $f \in L^q(\mu),\ \frac{1}{p}+\frac{1}{q} = 1$, annihilating all functions from $\mathcal{E}_a$. Equivalently for the Fourier transform $\widehat{f\mu}$ of the measure $f\mu$ we have

$$\widehat{f\mu}(s) = \frac{1}{\sqrt{2\pi}}\int e^{-ist} f(t)d\mu(t) = 0$$

for all $s \in [-a, a]$. Hence, the problem translates into finding an L^q-density f such that $f\mu$ has a gap in the Fourier spectrum (spectral gap) containing the interval $[-a, a]$. It turns out that before solving the type problem one needs to solve this version of the gap problem for $q = 1$, which is no longer equivalent to L^p-completeness. This is the so-called gap problem, whose recent solution we discuss in Section 2.3.

2.2 Toeplitz kernels. Recall that the Toeplitz operator T_U with a symbol $U \in L^\infty(\mathbb{R})$ is the map

$$T_U : H^2 \to H^2, \qquad F \mapsto P_+(UF),$$

where P_+ is the Riesz projection, i.e. the orthogonal projection from $L^2(\mathbb{R})$ onto the Hardy space H^2 in the upper half-plane $\mathbb{C}_+$. Passing from a function in H^2 to its non-tangential boundary values on $\mathbb{R}$, H^2 can be identified with a closed subspace of $L^2(\mathbb{R})$ which makes the Riesz projection correctly defined.

We will use the following notation for kernels of Toeplitz operators (or *Toeplitz kernels*) in H^2:

$$N[U] = \ker T_U.$$

A bounded analytic function in $\mathbb{C}_+$ is called inner if its boundary values on $\mathbb{R}$ have absolute value 1 almost everywhere. Meromorphic inner functions (MIFs) are those inner functions which can be extended as meromorphic functions in the whole complex plane. MIFs play a significant role in applications to spectral problems for differential operators, see for instance Makarov and Poltoratski [2005].

If θ is an inner function we denote by K_θ the so-called model space of analytic functions in the upper half-plane defined as the orthogonal complement in H^2 to the subspace θH^2, $K_\theta = H^2 \ominus \theta H^2$. An important observation is that for the Toeplitz kernel with the symbol $\bar\theta$ we have $N[\bar\theta] = K_\theta$ for any inner θ.

Along with H^2-kernels of Toeplitz operators, one may consider kernels $N^p[U]$ in other Hardy classes H^p, the kernel $N^{1,\infty}[U]$ in the 'weak' space $H^{1,\infty} = H^p \cap L^{1,\infty}$, $0 < p < 1$, or the kernel in the Smirnov class $\mathfrak{N}^+(\mathbb{C}_+)$, defined as

$$N^+[U] = \{f \in \mathfrak{N}^+ \cap L^1_{loc}(\mathbb{R}) : \bar U \bar f \in \mathfrak{N}^+\}$$

for $\mathfrak{N}^+$ and similarly for other spaces. If θ is a meromorphic inner function, $K^+_\theta = N^+[\bar\theta]$ can also be considered. For more on such kernels see Makarov and Poltoratski [2005] and Poltoratski [2015b].

Now let us discuss reformulations of the BM, gap and type problems mentioned above in the language of Toeplitz kernels, starting with the gap problem. One of the ways to state the gap problem is as follows.

Denote by M the set of all finite complex measures on $\mathbb{R}$. For a closed subset of real line X define its gap characteristic $\mathbf{G}_X$ as

$$\mathbf{G}_X = \sup\{a \mid \exists\, \mu \in M,\ \mu \not\equiv 0,\ \operatorname{supp}\mu \subset X,\ \text{ such that } \hat\mu = 0 \text{ on } [0,a]\}. \tag{2-2}$$

The problem is to find a formula for $\mathbf{G}_X$ in terms of X. Notice that the version of the problem discussed in the previous section, where for a fixed measure μ one looks for the supremum of the size of the spectral gap of $f\mu$ taken over all $f \in L^1(\mu)$ is equivalent to the last version of the problem with $X = \operatorname{supp}\mu$, see Proposition 1 in Poltoratski [2012] or Poltoratski [2015b].

While we postpone the formula for $\mathbf{G}_X$ until Section 2.3, here is the connection with the problem on injectivity of Toeplitz operators.

Here and throughout the paper for $a > 0$ we denote by S^a the exponential MIF, $S^a(z) = e^{iaz}$. For an inner function θ in the upper half-plane we denote by $\operatorname{spec}_\theta$ the closed subset of $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ of points at which the non-tangential limits of θ are equal to 1. Note that θ is a MIF if and only if $\operatorname{spec}_\theta$ is a discrete set.

If once again $X \subset \mathbb{R}$ is a closed set, denote

$$N_X = \sup\{a \mid N[\bar{\theta} S^a] \neq 0 \text{ for some meromorphic inner } \theta,\ \operatorname{spec}_\theta \subset X\}.$$

Theorem 2 (Makarov and Poltoratski [2005], Section 2.1)**.**

$$\mathbf{G}_X = N_X.$$

Similar translations can be given for the BM and type problems. If $\Lambda \subset \mathbb{C}$ is a complex sequence of frequencies, denote by $\Lambda' \subset \mathbb{C}_+$ the sequence in the upper half-plane obtained from Λ by replacing all points from the lower half-plane with their complex conjugates and replacing every real point $\lambda \in \Lambda$ with $\lambda + i$. Note that if Λ' does not satisfy the Blaschke condition in $\mathbb{C}_+$ then the radius of completeness of Λ, $R(\Lambda)$ defined in the last section, is infinite. If Λ' does satisfy the Blaschke condition, denote by B_Λ the Blaschke product with zeros at Λ'. Then the radius of completeness of Λ satisfies

$$R(\Lambda) = \sup\{a \mid N[S^a B_\Lambda] = 0\}.$$

This formula provides a reformulation of the BM problem in the language of injectivity of Toeplitz operators and can be used to translate the BM theorem into a result in this area. Such a translation can then be used in applications and point to further generalizations of BM theory, see Makarov and Poltoratski [2005, 2010].

Let μ be a finite positive singular measure on $\mathbb{R}$ and let us denote by $K\mu$ its Cauchy integral

$$K\mu(z) = \frac{1}{2\pi i} \int \frac{d\mu(t)}{t - z}.$$

Let $\theta = \theta_\mu$ denote the Clark inner function corresponding to μ, i.e., the inner function in $\mathbb{C}_+$ defined as

$$\theta(z) = \frac{K\mu(z) - 1}{K\mu(z) + 1}.$$

The Toeplitz version of the type problem is obtained via the following formula for the type of μ (defined in the last section):

$$\mathbf{T}_\mu = \sup\{a \mid N[\bar{\theta} S^a] \neq 0\}.$$

Comparing this equation with the Toeplitz version of the formula for the radius of completeness above, one can see the 'duality' relation between the BM and type problems, which translate into problems on injectivity of Toeplitz operators with complex conjugate symbols. Such a connection between the two problems was known to the experts on the intuitive level for a long time, but now can be expressed in precise mathematical terms using the Toeplitz language.

For a more detailed discussion of the results mentioned in this section and further references see Poltoratski [2015b].

2.3 A formula for the gap characteristic of a set. To give a formula for $\mathbf{G}_X$ defined in Section 2.2 we need to start with the following definition.

Let

$$... < a_{-2} < a_{-1} < a_0 = 0 < a_1 < a_2 < ...$$

be a two-sided sequence of real points. We say that the intervals $I_n = (a_n, a_{n+1}]$ form a short partition of $\mathbb{R}$ if $|I_n| \to \infty$ as $n \to \pm\infty$ and the sequence $\{I_n\}$ is short, i.e. the sum in (2-1) is finite.

Let $\Lambda = \{\lambda_n\}$ be a discrete sequence of distinct real points and let d be a positive number. We say that Λ is a d-uniform sequence if there exists a short partition $\{I_n\}$ such that

(2-3) $$\Delta_n = \#(\Lambda \cap I_n) = d|I_n| + o(|I_n|) \text{ as } n \to \pm\infty \text{ (density condition)}$$

and

(2-4) $$\sum_n \frac{\Delta_n^2 \log|I_n| - E_n}{1 + \operatorname{dist}^2(0, I_n)} < \infty \qquad \text{(energy condition)}$$

where

$$E_n = E(\Lambda \cap I_n) = \sum_{\lambda_k, \lambda_l \in I_n,\ \lambda_k \neq \lambda_l} \log|\lambda_k - \lambda_l|.$$

Notice that the series in the energy condition is positive since every term in the sum defining E_n is at most $\log|I_n|$ and there are less than Δ_n^2 terms. Convergence of positive series is usually easier to analyze.

The quantity E_n admits a physical interpretation as the potential energy of a system of 'flat electrons' placed at points of $\Lambda \cap I_n$. The term $\Delta_n^2 \log |I_n|$ corresponds to the energy of Δ_n electrons spread uniformly over I_n, up to a $O(|I_n|^2)$-term, which is negligible in (2-4) due to the shortness of $\{I_n\}$. Hence (2-4) can be viewed as the condition of finite work, needed to transform our sequence into an arithmetical progression.

In regard to the gap problem, d-uniform sequences have the property that any such sequence can support a measure with a spectral gap of the size $d - \varepsilon$ for any $\varepsilon > 0$. Conversely, any discrete sequence with this property must contain a d-uniform sequence. Moreover, for a general closed set we have

Theorem 3 (Poltoratski [2012, 2015b]).

$$\mathbf{G}_X = \sup\{d \mid X \textit{ contains a } d\textit{-uniform sequence}\},$$

if the set on the right is non-empty and $\mathbf{G}_X = 0$ *otherwise.*

One of the main ingredients of the proof is the Toeplitz approach to the gap problem discussed in the last section and used earlier in Mitkovski and Poltoratski [2010]. De Branges' "Theorem 66" (Theorem 66, de Branges [1968]) in Toeplitz form, which in its turn uses the Krein-Milman result on the existence of extreme points in a convex set, provides a key step of the proof allowing to discretize the problem. Another key component is the idea by Beurling and Malliavin to set up an extremal problem in the real Dirichlet space in $\mathbb{C}_+$ to construct an extremal measure with the desired spectral gap, see Poltoratski [2012, 2015b].

2.4 Bernstein's weighted uniform approximation. We say that a function $W \geqslant 1$ on $\mathbb{R}$ is a *weight* if W is lower semi-continuous and $W(x) \to \infty$ as $|x| \to \infty$. Denote by C_W the space of all continuous functions f on $\mathbb{R}$ such that $f/W \to 0$ as $x \to \pm\infty$ with the norm

$$||f||_W = \sup_{\mathbb{R}} \frac{|f|}{W}. \tag{2-5}$$

The weighted approximation problem posted by Sergei Bernstein in 1924 Bernstein [1924] asks to describe the weights W such that polynomials are dense in C_W. Similar questions can be formulated for exponentials and other families of functions in place of polynomials. Further information and references on the history of Bernstein's problem can be found in two classical surveys by Akhiezer Ahiezer [1956] and Mergelyan [1956], a recent one by Lubinsky [2007], or in the first volume of Koosis' book Koosis [1988].

It turns out that to obtain a formula for the type of a finite positive measure it makes sense to follow the historic path and first consider the problem on completeness of exponentials in Bernstein's settings. As a byproduct, for the original question on completeness of polynomials one obtains the following formula.

If Λ is a discrete real sequence we will assume that it is enumerated in the natural order, i.e. $\lambda_n < \lambda_{n+1}$, non-negative elements are indexed with non-negative integers and negative elements with negative integers.

We say that a sequence $\Lambda = \{\lambda_n\}$ has (two-sided) upper density d if

$$\limsup_{A\to\infty} \frac{\#[\Lambda \cap (-A, A)]}{2A} = d.$$

If $d = 0$ we say that the sequence has zero density.

A discrete sequence $\Lambda = \{\lambda_n\}$ is called *balanced* if the limit

$$\lim_{N\to\infty} \sum_{|n|<N} \frac{\lambda_n}{1+\lambda_n^2} \tag{2-6}$$

exists.

Observe that any even sequence (any sequence Λ satisfying $-\Lambda = \Lambda$) is balanced. So is any two-sided sequence sufficiently close to even. At the same time, a one-sided sequence has to tend to infinity fast enough to be balanced (the series $\sum \lambda_n^{-1}$ must converge).

Let $\Lambda = \{\lambda_n\}$ be a balanced sequence of finite upper density. For each n, $\lambda_n \in \Lambda$, put

$$p_n = \frac{1}{2}\left[\log(1+\lambda_n^2) + \sum_{n\neq k,\ \lambda_k\in\Lambda} \log\frac{1+\lambda_k^2}{(\lambda_k-\lambda_n)^2}\right],$$

where the sum is understood in the sense of principle value, i.e. as

$$\lim_{N\to\infty} \sum_{0<|n-k|<N} \log\frac{1+\lambda_k^2}{(\lambda_k-\lambda_n)^2}.$$

We will call the sequence of such numbers $P = \{p_n\}$ the *characteristic sequence* of Λ.

Note that for a sequence of finite upper density the last limit exists for every n if and only if it exists for some n, if and only if the sequence is balanced.

Theorem 4 (Poltoratski [2015a,b]). *Let W be a weight such that C_W contains all polynomials. Polynomials are not dense in C_W if and only if there exists a balanced sequence $\Lambda = \{\lambda_n\}$ of zero density such that Λ and its characteristic sequence $P = \{p_n\}$ satisfy*

$$\sum W(\lambda_n)\exp(p_n) < \infty. \tag{2-7}$$

The proof is elementary in nature and the result is similar to de Branges' theorem from de Branges [1959] where the condition of completeness of polynomials is formulated in terms of existence of a certain entire function. In the next section we will pass from Bernstein's problem to the type problem mentioned in the introduction by first replacing the polynomials with exponentials. Such a replacement complicates the problem significantly. In particular, its solution presented below requires advanced tools of BM theory, which are not required for the above result.

2.5 Type formulas. Continuing our discussion of Bernstein's weighted uniform approximation from the last section, for a weight W we define the type of W as

$$\mathbf{T}_W = \inf\{a \mid \mathcal{E}_a \text{ is complete in } C_W\}.$$

We put $\mathbf{T}_W = 0$ if the last set is empty.

Theorem 5 (Poltoratski [2015b]).

$$\mathbf{T}_W = \sup\left\{d \mid \sum \frac{\log W(\lambda_n)}{1+\lambda_n^2} < \infty \textit{ for some } d\textit{-uniform sequence } \Lambda\right\},$$

if the set is non-empty, and 0 *otherwise.*

Via the connection between Bernstein's and L^p-approximation problems found by A. Bakan in Bakan [2008], the last statement immediately yields the following L^p-statement. For $p > 1$ We define the L^p-type of a measure, $\mathbf{T}^p_\mu$, similarly to the definition of $\mathbf{T}_\mu$ given in Section 2.1, but with $L^2(\mu)$ replaced with $L^q(\mu)$. In these notations, $\mathbf{T}_\mu = \mathbf{T}^2_\mu$. We say that W is a μ-weight if $\int W d\mu < \infty$.

Corollary 1 (Poltoratski [2013]). *Let μ be a finite positive measure on the line. Let $1 < p \leqslant \infty$ and $a > 0$ be constants.*

Then $\mathbf{T}^p_\mu \geqslant a$ if and only if for any μ-weight W and any $0 < d < a$ there exists a d-uniform sequence $\Lambda = \{\lambda_n\} \subset \operatorname{supp}\mu$ such that

$$\sum \frac{\log W(\lambda_n)}{1+\lambda_n^2} < \infty. \tag{2-8}$$

As one can see from this statement, $\mathbf{T}_\mu = \mathbf{T}^p_\mu$ for any $p > 1$, which came as a surprise to some of the experts. In view of this property, it makes sense to return to the notation $\mathbf{T}_\mu$ in our future statements. Note that the case of $p = 1$ constitutes the gap problem discussed above with a different solution.

A more convenient L^p-statement was recently given in Poltoratski [n.d.]. If $\Lambda = \{\lambda_n\} \subset \mathbb{R}$ is a discrete sequence of distinct points we denote by Λ^* the sequence of intervals λ^*_n such that each λ^*_n is centered at λ_n and has the length equal to one-third of the distance from λ_n to the rest of Λ. Note that then the intervals λ^*_n are pairwise disjoint.

Theorem 6. *Let μ be a finite positive measure on the line. Then*

$$\mathbf{T}_\mu = \max\{d \,|\, \exists\, d\text{-}uniform\ \Lambda\ such\ that \sum \frac{\log \mu(\lambda_n^*)}{1+n^2} > -\infty\}$$

if the set is non-empty and $\mathbf{T}_\mu = 0$ otherwise.

2.6 Toeplitz order. Toeplitz operators provide universal language which can put many seemingly different problems from the area of UP into the same scale. Translations of known results and open problems from different areas into this universal language reveal surprising connections, point to correct generalizations and may indicate further directions for research. As was mentioned before, first such translations were found in the series of papers by Hruschev, Nikolski and Pavlov Hruščëv, Nikolskii, and Pavlov [1981] and Pavlov [1979] where problems on Riesz sequences and bases in model spaces were studied in terms of *invertibility* of related Toeplitz operators. In Makarov and Poltoratski [2005, 2010] connections between problems on completeness, uniqueness sets and spectral problems for differential operators with *injectivity* of Toeplitz operators were established. A recent attempt to systematize the problems on Toeplitz operators which emerge via this approach was undertaken in Poltoratski [2017]. The main idea is to define a partial order on the set of MIFs induced by Toeplitz operators and view several general questions of UP as questions on the properties of such an order. In this section we present a short overview of this approach.

Definition 1. *If θ is an inner function we define its (Toeplitz) dominance set $\mathfrak{D}(\theta)$ as*

$$\mathfrak{D}(\theta) = \{I\ inner \,|\, N[\bar{\theta} I] \neq 0\}.$$

Every collection of sets admits natural partial ordering by inclusion. In our case, we consider dominance sets $\mathfrak{D}(\theta)$ as subsets of the set of all inner functions in the upper half-plane and the partial order $\subset$ on this collection. This partial order induces a preorder on the set of all inner functions in $\mathbb{C}_+$. Proceeding in a standard way, we can modify this preorder into a partial order by introducing equivalence classes of inner functions. The details of this definition are as follows.

Definition 2. *We will say that two inner functions I and J are Toeplitz equivalent, writing $I \overset{\scriptscriptstyle T}{\sim} J$, if $\mathfrak{D}(I) = \mathfrak{D}(J)$. This equivalence relation divides the set of all inner functions in $\mathbb{C}_+$ into equivalence classes. We call this relation **Toeplitz equivalence** (TE).*

Further, we introduce a partial order on these equivalence classes defining it as follows.

Definition 3. *We write $I \overset{\scriptscriptstyle T}{\leqslant} J$ (meaning that the equivalence class of I is 'less or equal' than the equivalence class of J) if $\mathfrak{D}(I) \subset \mathfrak{D}(J)$. We call this partial order on the set of inner functions in $\mathbb{C}_+$ **Toeplitz order** (TO).*

Let B_n and B_k be Blaschke products of degree n and k correspondingly. Then $B_n \overset{T}{\sim} B_k$ iff $n = k$ and $B_n \overset{T}{<} B_k$ iff $n < k$. The relation becomes more interesting for infinite Blaschke products and singular inner functions. For instance, if J_μ and J_ν are two singular functions then $\nu - \mu \geq 0$ implies $J_\mu \overset{T}{\leqslant} J_\nu$ but not vice versa as follows from an example by A. Alexandrov.

Now let us present some of the translations of the known problems mentioned in previous sections in the language of TO.

Recall that any MIF I has the form $I = B_\Lambda S^a$ where B is a Blaschke product with a discrete sequence of zeros Λ and $S^a = e^{iaz}$ is the exponential function. Put $r(I) = D^*(\Lambda) + a$. The BM Theorem 1 discussed in Section 2.1 is equivalent to the following statement.

Theorem 7. *For any MIF I,*

$$I \overset{T}{\leqslant} S^b \;\Rightarrow\; r(I) \leq b$$

and

$$r(I) < b \;\Rightarrow\; I \overset{T}{<} S^b.$$

As we can see, the Beurling-Malliavin formula gives a metric condition for TO in the very specific case when one of the functions is the exponential function. Similar descriptions for more general classes of inner functions, especially those appearing in applications to completeness problems and spectral analysis remain mostly open. Below we present one of such extensions found in Makarov and Poltoratski [2010].

As was shown in Makarov and Poltoratski [2005] the class of MIFs with polynomially growing arguments appears naturally in a number of applications including completeness problems for Airy and Bessel functions, spectral problems for regular Schrödinger operators and Dirac systems, etc. An analog of Theorem 1 proved in Makarov and Poltoratski [2010] can be applied to some of such problems. Here we present an equivalent reformulation similar to Theorem 7.

Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\gamma(\mp\infty) = \pm\infty$. i.e.,

$$\lim_{x\to-\infty} \gamma(x) = +\infty, \qquad \lim_{x\to+\infty} \gamma(x) = -\infty.$$

Define γ^* to be the smallest non-increasing majorant of γ:

$$\gamma^*(x) = \max_{t\in[x,+\infty)} \gamma(t).$$

The family of intervals $BM(\gamma) = \{I_n\}$ is defined as the collection of the connected components of the open set

$$\{x \in \mathbb{R} |\ \gamma(x) \neq \gamma^*(x)\}.$$

Let $\kappa \geq 0$ be a constant. We say that γ is κ-*almost decreasing* if

$$\sum_{I_n \in BM(\gamma)} (\operatorname{dist}(I_n, 0) + 1)^{\kappa - 2} |I_n|^2 < \infty. \tag{2-9}$$

We define an argument of a MIF I on $\mathbb{R}$ is a real analytic function ψ such that $I = e^{i\psi}$.

Theorem 8 (Makarov and Poltoratski [2010] and Poltoratski [2017]). *Let U be a MIF with $|U'| \asymp x^\kappa$,*
$\kappa \geq 0$, $\gamma = \arg U$ on $\mathbb{R}$. Let J be another MIF, $\sigma = \arg J$ on $\mathbb{R}$.

I) If $\sigma - (1 - \varepsilon)\gamma$ is κ-almost decreasing, then $J \overset{\tau}{<} U$;
II) If $J \overset{\tau}{\leqslant} U$ then $\sigma - (1 + \varepsilon)\gamma$ is κ-almost decreasing.

Let us point out that even finding an analog for the above statement for $\kappa < 0$ presents an open problem. MIFs with $\kappa < 0$ appear in some of the applications mentioned in Makarov and Poltoratski [2010].

In regard to the type problem we have the following translation. We will denote by θ_μ the inner function with Clark measure μ as defined in Section 2.2.

Theorem 9 (Poltoratski [2017]).

$$\mathbf{T}_\mu = \sup\{a | S^a \overset{\tau}{\leqslant} \theta_\mu\}.$$

As one can see, the solutions to BM and type problems give formulas which can be used to compare MIFs with respect to TO in several particular situations. In both cases the functions are compared with the exponential function S^a. The extension found in Makarov and Poltoratski [2010] replaces S^a with a function with polynomially growing argument. Apart from the results mentioned here and a few elementary examples contained in Poltoratski [2017], giving metric conditions on MIFs I, J necessary or sufficient for $I \overset{\tau}{\sim} J$ present a collection of open problems with applications in several areas of UP.

3 Inverse spectral problems and truncated Toeplitz operators

3.1 Canonical systems. Consider a 2×2 differential system with a spectral parameter $z \in \mathbb{C}$:

$$\Omega \dot{X} = zH(t)X - Q(t)X, \quad -\infty < t_- < t < t_+ \tag{3-1}$$

where

$$X(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} \text{ and } \Omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We assume the (real-valued) coefficients to satisfy

$$H,\ Q \in L^1_{loc}((t_-,t_+) \to \mathbb{R}^{2\times 2}).$$

By definition, a solution $X = X_z(t)$ is a $C^2((t_-,t_+))$-function satisfying the equation. An initial value problem (IVP) for the system (3-1) can be given via an initial condition $X(t_-) = x,\ x \in \mathbb{C}^2$. Let us immediately point out the following well-known property.

Theorem 10. *Every IVP for* (3-1) *has a unique solution on* (t_-,t_+)*. For each fixed* t*, this solution presents an entire function* $E_t(z) = u_z(t) - i v_z(t)$ *of exponential type.*

Let us further assume that $H(t),\ Q(t)$ are real symmetric locally summable matrix-valued functions and that $H(t) \geq 0$. The Hilbert space $L^2(H)$ consists of (equivalence classes) of vector-functions with

$$||f||^2_H = \int_{t_-}^{t_+} < Hf, f > dt < \infty.$$

The system (3-1) is an eigenvalue equation $DX = zX$ for the (formal) differential operator

$$D = H^{-1}\left[\Omega \frac{d}{dt} + Q\right].$$

Many important classes of second order differential equations can be rewritten in the form of (3-1). Consider, for instance, the Schrödinger equation $-\ddot{u} = zu - qu$ on an interval. Put $v = -\dot{u}$ and $X = (u,v)^T$ to obtain

$$\Omega \dot{X} = z \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X - \begin{pmatrix} q & 0 \\ 0 & -1 \end{pmatrix} X.$$

Similarly one can use (3-1) to rewrite string, Dirac and several other known equations. A discrete version of (3-1) where the function H consists of so-called jump intervals can be used to express difference equations and Jacobi matrices. Well studied Dirac systems are usually written in the form of a 2×2 system. In that case $H \equiv I$ and the general form is

$$\Omega \dot{X} = z \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} X - \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} X,\ q_{12} = q_{21}.$$

The "standard form" of a Dirac system is with $Q = \begin{pmatrix} -q_2 & -q_1 \\ -q_1 & q_2 \end{pmatrix}$. In this case $f = q_1 + iq_2$ is the potential function.

Among all self-adjoint systems discussed above we single out a subclass of so-called canonical systems with $Q \equiv 0$:

(3-2) $$\Omega \dot{X} = zH(t)X.$$

This turns out to be the right object for the theory. The first key observation is that a general self-adjoint system can be reduced to canonical form. To perform such a reducion first solve $\Omega\dot{V} = -QV$ for a 2×2 matrix valued function V and make a substitution $X = VY$. Then (3-1) becomes

$$\Omega\dot{Y} = z\,[V^*HV]\,Y.$$

For instance, a Dirac system with real potential f becomes a canonical system with Hamiltonian

$$H^{CS} = \begin{pmatrix} e^{-2\int_0^t f} & 0 \\ 0 & e^{-2\int_0^t f} \end{pmatrix}.$$

When analyzing a second order equation one starts with a self-adjoint form of the system (3-1) into which the equation can be easily converted. Further conversion into the canonical form is necessary in the theory because only for canonical systems (after some additional normalizations) the theory provides a one-to-one correspondence with de Branges spaces of entire functions defined below. In the preliminary (3-1)-form two different systems may correspond to the same space.

3.2 De Branges spaces. We call an entire function E an an Hermite-Biehler (de Branges) function if it satisfies

$$|E(z)| > |E(\bar{z})|$$

for all $z \in \mathbb{C}_+$. As before we denote by H^2 the Hardy space in $\mathbb{C}_+$.

For an entire funcion $G(z)$ we denote $G^\#(z) = \bar{G}(\bar{z})$. For every Hermite-Biehler (HB) function E we define the space $B(E)$ to be the Hilbert space of entire functions F such that

$$F/E,\ F^\#/E \in H^2.$$

The Hilbert structure in the space is inherited from H^2, i.e. if $F, G \in B(E)$ then

$$< F, G >_{B(E)} = < F/E, G/E >_{H^2} = \int_{\mathbb{R}} F(x)\bar{G}(x)\frac{dx}{|E|^2}.$$

The well-known Paley-Wiener spaces of entire functions PW_a give a particular example of de Branges spaces $B(E)$ with $E(z) = e^{-iaz},\ a > 0$.

One of the fundamental properties of de Branges' spaces is that they admit an equivalent axiomatic definition. A similar definition with a slightly different set of axioms (for a slightly different space) was earlier found by Krein.

Theorem 11 (de Branges [1968]).

Suppose that H is a Hilbert space of entire functions that satisfies

(A1) $F \in H, F(\lambda) = 0 \Rightarrow F(z)(z-\bar{\lambda})/(z-\lambda) \in H$ *with the same norm*
(A2) $\forall \lambda \notin \mathbb{R}$, *the point evaluation is bounded*
(A3) $F \to F^{\#}$ *is an isometry*
Then $H = B(E)$ *for some* $E \in HB$.

Let E be an Hermite-Biehler function. Consider entire functions

$$A = (E + E^{\#})/2 \text{ and } B = (E - E^{\#})/2i.$$

The space $B(E)$ is a reproducing kernel space, i.e. any $\lambda \in \mathbb{C}$, there exists $K_\lambda \in B(E)$ such that $F(\lambda) =< F, K_\lambda >$ for any $F \in B(E)$. The kernels are given by the formula

$$K_\lambda(z) = \frac{1}{\pi}\frac{B(z)\bar{A}(\lambda) - A(z)\bar{B}(\lambda)}{z-\bar{\lambda}}.$$

Each HB-function E gives rise to a MIF $\theta_E(z) = \frac{E^{\#}(z)}{E(z)}$. Conversely, for each MIF θ there exists an HB-function E such that $\theta = \theta_E$. The model space K_θ defined in Section 2.2 is related to the de Branges space $B(E)$ via the simple identity $B(E) = EK_\theta$ with the map $f \mapsto Ef$ defining an isometric isomorphism between the spaces.

The following connection between canonical systems and de Branges spaces translates spectral problems for various classes of second order differential equations into the language of complex analysis.

For the sake of brevity here we will consider only canonical systems (3-2) without "jump intervals", i.e. without intervals where H is a constant matrix of rank 1. This assumption is made in many survey articles on the subject as it simplifies the main statements in the theory. At the same time, the case of jump intervals allows one to include discrete models, such us difference equations, Jacobi matrices, orthogonal polynomials, etc., into the scope of Krein - de Branges theory.

Consider a canonical system (without jump intervals) with any real initial condition at t_-. Denote the solution by $X_z(t) = (A_t(z), B_t(z))$. For each fixed t consider the entire function $E_t(z) = A_t(z) - iB_t(z)$. The following statement connects canonical systems with HB-functions and de Branges spaces, see de Branges [1968] and Makarov and Poltoratski [2005].

Theorem 12. *For any fixed* t, $E_t(z)$ *is a Hermit-Biehler entire function. The map* W *defined as* $WX_z = K^t_{\bar{z}}$ *extends unitarily to the map from* $L^2(H,(t_-,t))$ *to the de Branges space* $B(E_t)$ *(Weyl Transform).*

The formula for W*:*

$$Wf(z) =< Hf, X_{\bar{z}} >_{L^2(H,(t_-,t))} = \int_{t_-}^{t} < H(t)f(t), X_{\bar{z}}(t) > dt.$$

Here K_z^t denotes the reproducing kernel in $B(E_t)$. The Weyl transform can be viewed as a general form of Fourier transform that puts Krein's canonical systems into one-to-one correspondence with chains of de Branges' spaces $B(E_t)$, $t_- \leq t < t_+$. The case of free Dirac system ($Q = 0$) produces the standard Fourier transform and Payley-Wiener spaces PW_t.

One of the key results of the Krein-de Branges theory says that for any (regular) chain of de Branges spaces there exists a canonical system generating that chain as in the last statement. For some such chains the corresponding systems will have jump intervals, the case we do not discuss in this note. For instance, orthogonal polynomials in $L^2(\mu)$ satisfy difference equations that can be rewritten as a Krein system with jump intervals. In that case $B(E_t) = B_n$ will be a space of polynomials of degree less than n. The space will remain the same, as a set, on each jump interval.

We will write $B(E_1) \doteq B(E_2)$ if the two de Branges spaces $B(E_1)$ and $B(E_2)$ are equal as sets, with possibly different norms. Such relations occur in spectral theory when the difference between corresponding Hamiltonians is locally summable or similar. For instance, the following is an important observation in the theory of Dirac systems:

Theorem 13. *Let $B(E_t)$, $t \in (t_-, t_+)$ be the chain of de Branges' spaces corresponding to a Dirac system on (t_-, t_+) with an L^1_{loc}-potential. Then $B(E_t) \doteq PW_t$.*

A more general Gelfand-Levitan theory can be viewed a subset of Krein - de Branges theory in the case when the system corresponds to the chain of de Branges spaces that are equal to Payley-Wiener spaces as sets. Together with regular Schrödinger and Dirac operators such theory will contain a broader class of canonical systems, appearing in applications, see Makarov and Poltoratski [n.d.].

The following important question arises in connection with Gelfand-Levitan theory. For what functions E will the de Brnanges space $B(E)$ coincide with the Payley-Wiener space as a set (with a different but equivalent norm)? This question is equivalent to characterizing Riesz bases of reproducing kernels in Payley-Wiener spaces or Riesz bases of exponential functions in $L^2(a, b)$. It is related to many similar questions in harmonic analysis, such as problems on frames, sampling and interpolation.

A more general question, that can be similarly translated, is to describe pairs of HB-functions such that $B(E_1) \doteq B(E_2)$. Even in the Payley-Wiener case above, despite a large number of deep results (see for instance Ortega-Cerdà and Seip [2002] and Seip [2004] for such results and further references), the problem is not completely finished. Very little is known in the general case.

In the language of Toeplitz Order defined in Section 2.6 we have the following connection.

Theorem 14 (Poltoratski [2017]).

$$B(E_1) = B(E_2) \quad \Leftrightarrow \quad \theta_{E_1} \overset{r}{\sim} \theta_{E_2}.$$

3.3 Inverse spectral problems. Once again, let us consider a canonical system (3-2) with no jump intervals. Let us fix a boundary condition $X(t_-) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. We call a positive measure μ a spectral measure for the system (3-2) (corresponding to the chosen boundary condition at t_-) if for every $t \in [t_-, t_+)$ the space $B(E_t)$ is isometrically embedded into $L^2(\mu)$. Such a spectral measure may be unique (limit point case) or belong to a one-parameter family of measures with similar property (limit circle case). The limit point case occurs when the integral

$$\int_{t_-}^{t_+} \text{trace } H(t)dt$$

is infinite and the limit circle case corresponds to the finite integral.

Equivalently, a spectral measure can be defined with the condition that the Weyl transform is a unitary operator $L^2(H, (t_-, t)) \to L^2(\mu)$ for any $t \in [t_-, t_+)$.

An inverse spectral problem for the canonical system (3-2) asks to recover the system, i.e., the Hamiltonian H, from its spectral measure μ. Classical results by Borg, Marchenko (Schrödinger case) and de Branges (general canonical system) establish uniqueness of solution for the inverse spectral problem. In the case of canonical systems, there is also a remarkable existence result by de Branges de Branges [1968] which says that any Poisson-finite positive measure is a spectral measure of a canonical system. We call a measure μ on $\mathbb{R}$ Poisson-finite if

$$\int \frac{d|\mu|}{1+x^2} < \infty.$$

As it often happens, the existence theorem does not provide an algorithm for the recovery of H from μ. In fact, after many decades of research only several elementary examples of explicit solutions to the inverse spectral problem for canonical systems have been recorded in the literature. Our methods based on the use of truncated Toeplitz operators provide such an algorithm which yields a number of new interesting examples for the inverse problem. Our next goals is to describe our methods and present examples.

3.4 PW-measures and systems. This part of the note is based on our recent work with N. Makarov, Makarov and Poltoratski [n.d.].

Let μ be a Poisson-finite positive measure on $\mathbb{R}$. We say that μ is a sampling measure for the Paley-Wiener space PW_a if there exist constants $0 < c < C$ such that for any $f \in PW_a$,

$$c||f||_{PW_a} < ||f||_{L^2(\mu)} < C||f||_{PW_a}.$$

We say that μ is a Payley-Wiener (PW) measure if it is sampling for all spaces PW_a, $0 < a < \infty$.

Note that any PW-measure defines equivalent norms in all PW_a spaces. By verifying the axioms from Theorem 11 one can show the chain of PW-spaces with norms inherited from $L^2(\mu)$ is a chain of de Branges spaces $B(E_t)$ for some unknown HB-functions E_t, $B(E_t) \doteq PW_t$.

On the other hand, by a theorem from de Branges [1968], μ is a spectral measure of a canonical system (3-2) with a locally summable Hamiltonian H. By uniqueness of regular de Branges chains isometrically embedded in L^2-spaces of Poisson-finite measures, the chain induced by the canonical system must coincide with the chain $B(E_t)$ obtained above.

Conversely, if one starts with a canonical system on $[0, \infty)$ whose de Branges chain satisfies $B(E_t) \doteq PW_t$, its spectral measure μ is a PW-measure. We will call such systems PW-systems. This is a broad class of canonical systems which includes all of the equations considered in the classical Gelfand-Levitan theory and more.

As we saw, PW-systems and measures are in one-to-one correspondence with each other (after standard normalization of the time variable in the system). A study of spectral problems for PW-systems can be viewed as a generalization of the Gelfand-Levitan theory.

Let us start with the following analytic description of PW-measures. Let δ be a positive constant. We say that an interval $I \subset \mathbb{R}$ is a (μ, δ)-interval if

$$\mu(I) > \delta \text{ and } |I| > \delta.$$

Theorem 15 (Makarov and Poltoratski [n.d.]). *A positive Poisson-finite measure μ is a Paley-Wiener measure if and only if*

1) $\sup_{x\in\mathbb{R}} \mu((x, x+1)) < \infty$.

2) For any $d > 0$ there exists $\delta > 0$ such that for all sufficiently large intervals I there exist at least $d|I|$ disjoint (μ, δ)-intervals intersecting I.

3.5 Truncated Toeplitz operators. Let $\phi \in L^\infty(\mathbb{R})$. The truncated Toeplitz operator with symbol ϕ is defined as

$$L_\phi : PW_a \to PW_a, \; L_\phi f = P(\phi f),$$

where P denotes the orthogonal projection $L^2(\mathbb{R}) \to PW_a$. If μ is a measure on $\mathbb{R}$ one can define L_μ via quadratic forms with the operator $L_\mu : PW_a \to PW_a$ given by the relation

$$\int_{\mathbb{R}} f \bar{g} dx = \int f \bar{g} d\mu.$$

Notice that if $d\mu(x) = \phi(x)dx$ for $\phi \in L^\infty(\mathbb{R})$, then $L_\mu = L_\phi$.

Lemma 1 (Makarov and Poltoratski [ibid.]). *L_μ is a positive invertible operator in PW_a if and only if μ is a sampling measure for PW_a. Consequently, L_μ is a positive invertible operator in every PW_a, $0 < a < \infty$, if and only if μ is a PW-measure.*

Truncated Toeplitz operators corresponding to PW-measures appear in inverse spectral problems for canonical systems in the following way. For simplicity, let us consider the case of Krein's string, i.e., a canonical system (3-2) with a diagonal locally summable Hamiltonian

$$H(t) = \begin{pmatrix} h_{11}(t) & 0 \\ 0 & h_{22}(t) \end{pmatrix}.$$

Via a proper change of variable, one can normalize the problem so that $\det H = 1$ a.e. on (t_-, t_+). After performing such a normalization and fixing a boundary condition at t_-, such systems are put in one-to-one correspondence with even Poisson-finite positive measures on the real line (spectral measures).

Let now μ be a spectral measure of a PW-type Krein's string. Define the truncated Toeplitz operator L_μ and notice that by the last lemma L_μ is invertible in every PW_a. The key relation which solves the inverse spectral problem in this case is that the reproducing kernel in the de Branges space $B(E_t)$ corresponding to the system is the image of the sinc function, the reproducing kernel in PW_a, under L_μ^{-1}:

$$K_0^t(z) = L_\mu^{-1}\left(\frac{\sin tz}{z}\right).$$

After the reproducing kernel K_0^t is recovered, the Hamiltonian of the system can be found from

$$h_{11}(t) = \pi \frac{d}{dt} K_0^t(0) \tag{3-3}$$

and $h_{22} = 1/h_{11}$ (since $\det H = 1$).

In the case of general, non-diagonal Hamiltonians, the problem requires several additional steps, which can be completed via similar methods. The key ingredient in the general case, which we have no space to discuss here, is the so-called generalized Hilbert transform, which maps the de Branges chain $B(E_t)$ into the conjugate chain $B(\tilde{E}_t)$ corresponding to the Hamiltonian

$$\tilde{H} = \begin{pmatrix} h_{11}(t) & -h_{12}(t) \\ -h_{21}(t) & h_{22}(t) \end{pmatrix},$$

see Example 3 below. This operator reduces to the standard Hilbert transform in the free case. For more detailed account and the proofs see Makarov and Poltoratski [ibid.].

3.6 Examples of inverse spectral problems via truncated Toeplitz operators. We finish with the following examples of solutions to the inverse spectral problem.

Example 1. *Let $\mu = \sqrt{2\pi}\delta_0 + \frac{1}{\sqrt{2\pi}}m$, where m stands for the Lebesgue measure on $\mathbb{R}$ and δ_a is the unit mass at the point a. This is probably the simplest new example in our theory. The measure satisfies conditions of Theorem 15 and therefore is a PW-measure. The measure is even, hence H must be diagonal*

$$H = \begin{pmatrix} h(x) & 0 \\ 0 & \frac{1}{h}(x) \end{pmatrix}.$$

Our goal is to find h using the formula (3-3).

Denote by f_t the Fourier transform of $K_0^t(x)$. Then f_t is supported on $[-t,t]$ and satisfies

$$f_t * \widehat{\mu} = 1 \text{ on } [-t,t]$$

because $\frac{1}{\sqrt{2\pi}}\chi_{[-t,t]}$ is the Fourier transform of the sinc function, the reproducing kernel of PW_t, and

$$\widehat{K_0^t\mu} = \frac{1}{\sqrt{2\pi}} f_t * \widehat{\mu}.$$

Since $\widehat{\mu} = \delta_0 + m$, the Fourier transform of the reproducing kernel satisfies

$$1 = f_t(x) + \int_{-t}^{t} f_t(y)dy \text{ on } [-t,t].$$

It follows that $f_t(x) = c(t)\chi_{[-t,t]}(x)$ where

$$1 = c(t) + 2tc(t), \text{ i.e., } c(t) = 1/(1+2t).$$

Then

$$h(t) = \frac{d}{dt}K_0^t(0) = \frac{d}{dt}\int_{-t}^{t} c(t)dx = \frac{d}{dt}\frac{2t}{1+2t} = \frac{2}{(1+2t)^2}$$

and

$$H = \begin{pmatrix} \frac{2}{(1+2t)^2} & 0 \\ 0 & \frac{(1+2t)^2}{2} \end{pmatrix}.$$

for $0 \le t < \infty$.

Example 2. *Consider the Krein system with the spectral measure $\mu = \frac{1}{\sqrt{2\pi}}(1+\cos x)m$. Once again, the measure is a PW-measure. It is even and therefore H must be diagonal.*

As before, the Fourier transform of K_0^t, f_t, satisfies

$$f_t * \widehat{\mu} = 1 \text{ on } [-t,t].$$

Inserting the Fourier transform of μ and solving the last equation one obtains f_t as a step function described below.

For $\frac{n}{2} < t < \frac{n+1}{2}$, define f_t as follows. Consider the infinite Toeplitz matrix

$$J = \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 & 0 & \dots \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & \dots \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 & \dots \\ 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & \dots \\ 0 & 0 & 0 & \frac{1}{2} & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the diagonals are equal to the pointmasses of $\widehat{\mu}$ at n, $n \in \mathbb{Z}$. Denote by J_n the $n \times n$ sub-matrix in the upper left corner of J.

Consider the intervals $I_k = (a_k, b_k), 1 \le k \le n+1$, of length $t - \frac{n}{2}$ centered at

$$\frac{-n}{2}, \frac{-n+2}{2}, \dots, \frac{n-2}{2}, \frac{n}{2}$$

enumerated in the natural left-to-right order. Denote by J_k complementary intervals $J_k = (b_k, a_{k+1})$. On each I_k define $f_t = \frac{1}{2}a_k^{n+1}$ and on each J_k define $f_t = \frac{1}{2}a_k^n$, where

$$\begin{pmatrix} a_1^m \\ a_2^m \\ \vdots \\ a_m^m \end{pmatrix} = J_m^{-1} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Notice that

$$h(t) = \frac{d}{dt} K_0^t(0) = \frac{d}{dt} \int_{-t}^{t} f_t(x)dx = \Sigma(J_{n+1}^{-1}) - \Sigma(J_n^{-1}) \text{ on } \left(\frac{n}{2}, \frac{n+1}{2}\right],$$

where Σ denotes the sum of all elements of the matrix, with $\Sigma(J_0^{-1})$ defined as 0. Elementary calculations show the values of $h(t)$ on $(0, \frac{1}{2}], (\frac{1}{2}, 1], (1, \frac{3}{2}], \dots$ to be

$$1, \frac{1}{3}, \frac{2}{3}, \frac{2}{5}, \frac{3}{5}, \frac{3}{7}, \frac{4}{7}, \frac{4}{9}, \frac{5}{9}, \frac{5}{11}, \frac{6}{11}, \frac{6}{13},$$

$$\frac{7}{13}, \frac{7}{15}, \frac{8}{15}, \frac{8}{17}, \frac{9}{17}, \frac{9}{19}, \frac{10}{19}, \frac{10}{21}, \frac{11}{21}, \dots$$

$$\dots, \frac{n}{2n-1}, \frac{n}{2n+1}, \frac{n+1}{2n+1}, \frac{n+1}{2n+3}, \frac{n+2}{2n+3}, \dots$$

Example 3. *Let now* $\mu = \frac{1}{\sqrt{2\pi}}(1+\sin x)m$. *Note that* μ *is not even and hence the Hamiltonian has the general form*

$$H = \begin{pmatrix} \alpha(x) & \beta(x) \\ \beta(x) & \gamma(x) \end{pmatrix},$$

with non-zero β. *Clearly,* μ *is a* PW*-measure and hence this example can be treated within the extended Gelfand-Levitan theory. The entries* α *and* γ *can be calculated as in the last example. The Toeplitz matrix for the pointmasses of* $\widehat{\mu} = \delta_0 + \frac{i}{2}(\delta_{-1} - \delta_1)$ *is*

$$J = \begin{pmatrix} 1 & -\frac{i}{2} & 0 & 0 & 0 & \dots \\ \frac{i}{2} & 1 & -\frac{i}{2} & 0 & 0 & \dots \\ 0 & \frac{i}{2} & 1 & -\frac{i}{2} & 0 & \dots \\ 0 & 0 & \frac{i}{2} & 1 & -\frac{i}{2} & \dots \\ 0 & 0 & 0 & \frac{1}{2} & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Then

$$\alpha(t) = \Sigma(J_{n+1}^{-1}) - \Sigma(J_n^{-1}) \text{ on } \left(\frac{n}{2}, \frac{n+1}{2}\right].$$

Elementary calculations give the following values of α *on on* $(0, \frac{1}{2}], (\frac{1}{2}, 1], (1, \frac{3}{2}], \dots$:

$$i1, 5/3,\ 4/3,\ 4/5,\ 13/15,\ 25/21,\ 8/7,\ 8/9,\ 41/45,\ 61/55,\ 12/11,\ 12/13,$$

$$85/91,\ 113/105,\ 16/15,\ 16/17,\ 145/153,\ 181/171,\ 20/19,\ 20/21, \dots$$

For γ *we need to consider*

$$\sigma = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{2}m + \frac{1}{2}\sum \delta_{2\pi n - \frac{\pi}{2}}\right),$$

the so-called dual Clark measure for μ *(i.e.,* σ *is the Clark measure for* $-\theta$*, while* μ *is the Clark measure for* θ*, as defined is Section 2.2). After calculating*

$$\widehat{\sigma} = \frac{1}{2}\delta_0 + \frac{1}{2}\sum_{n\in\mathbb{Z}} (i)^n \delta_n,$$

the corresponding Toeplitz matrix is

$$L = \begin{pmatrix} 1 & \frac{i}{2} & -\frac{1}{2} & -\frac{i}{2} & \frac{1}{2} & \cdots \\ -\frac{i}{2} & 1 & \frac{i}{2} & -\frac{1}{2} & -\frac{i}{2} & \cdots \\ -\frac{1}{2} & -\frac{i}{2} & 1 & \frac{i}{2} & -\frac{1}{2} & \cdots \\ \frac{i}{2} & -\frac{1}{2} & -\frac{i}{2} & 1 & \frac{i}{2} & \cdots \\ \frac{1}{2} & \frac{i}{2} & -\frac{1}{2} & -\frac{i}{2} & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Then

$$\gamma(t) = \Sigma(L_{n+1}^{-1}) - \Sigma(L_n^{-1}) \text{ on } \left(\frac{n}{2}, \frac{n+1}{2}\right],$$

which produces the values

$$1, 5/3,\ 17/6,\ 5/2,\ 5/3,\ 37/21,\ 65/28,\ 9/4,\ 9/5,\ 101/55,\ 145/66,\ 13/6,$$

$$13/7,\ 197/105,\ 257/120,\ 17/8,\ 17/9,\ 325/171,\ 401/190,\ 21/10, \ldots$$

For the calculation of β we need to utilize the generalized Hilbert transform of the kernel K_0^t (at 0) mentioned in Section 3.5.

We define the Hilbert transform as

$$Hf = K(f\mu) - K(f\mu) + cf,$$

where

$$Kf\mu = \frac{1}{i\pi} \int \frac{1}{t-z} f(t) d\mu(t) = P(f\mu) + iQ(f\mu).$$

Here P and Q stand for the Poisson and the conjugate Poisson transforms correspondingly. The constant c is to be determined at the end of the calculation from the condition $\det H = 1$.

One of the main formulas of the theory gives the remaining coefficient $\beta(t)$ as

$$\beta(t) = \frac{d}{dt} H K_0^t(0).$$

We have

$$\widehat{PK_0^t\mu} = \widehat{K_0^t\mu} \text{ and } \widehat{QK_0^t\mu} = (-i)\operatorname{sign} x \cdot \widehat{K_0^t\mu},$$

and therefore

$$\widehat{KK_0^t\mu} = K_0^t\mu + \operatorname{sign} x \cdot \widehat{K_0^t\mu}$$

For $K_0^t K\mu$ we have

$$\widehat{K_0^t K\mu} = \widehat{K_0^t\mu} + \widehat{K_0^t} * (\operatorname{sign} x \cdot \widehat{\mu}).$$

Altogether we get

$$\beta(t) = \frac{d}{dt}\int_{\mathbb{R}}[\widehat{K_0^t\mu} + \text{sign } x \cdot \widehat{K_0^t\mu} - \widehat{K_0^t\mu} - \widehat{K}_0^t * (\text{sign } x \cdot \widehat{\mu}) + c\widehat{K}_0^t] =$$

$$\frac{d}{dt}\int_{\mathbb{R}}[\text{sign } x \cdot \widehat{K_0^t\mu} \ - \widehat{K}_0^t * (\text{sign } x \cdot \widehat{\mu}) + c\widehat{K}_0^t].$$

Put $n = [2t]$. *Let, like in the first section,*

$$\begin{pmatrix} a_1^m \\ a_2^m \\ \vdots \\ a_m^m \end{pmatrix} = J_m^{-1}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

By our construction, $\widehat{K_0^t\mu}$ *is equal to*

$$\widehat{K_0^t\mu} = \begin{cases} -\frac{i}{2}a_1^n \text{ on } (-t-1, -1-n/2) \\ -\frac{i}{2}a_1^{n-1} \text{ on } (-1-n/2, -t) \\ 1/2 \text{ on } (-t, t) \\ \frac{i}{2}a_{n-1}^{n-1} \text{ on } (t, n/2+1) \\ \frac{i}{2}a_n^n \text{ on } (n/2+1, t+1) \end{cases}.$$

Hence,

$$\frac{d}{dt}\int_{\mathbb{R}}\widehat{K_0^t\mu} = 1 + \frac{i}{2}(a_n^n - a_{n-1}^{n-1} - a_1^n + a_1^{n-1}).$$

Taking into account the relation $a_1^m = \bar{a}_m^m$, *we get*

$$\frac{d}{dt}\int_{\mathbb{R}}\widehat{K_0^t\mu} = 1 + \Im a_1^n - \Im a_1^{n-1}.$$

Since $\mu = (1+\sin x)m$,

$$\int_{\mathbb{R}}\widehat{K_0^t\mu} = \int_{\mathbb{R}}\widehat{K_0^t} + \frac{i}{2}\int_{\mathbb{R}}\widehat{K_0^t} - \frac{i}{2}\int_{\mathbb{R}}\widehat{K_0^t} = \int_{\mathbb{R}}\widehat{K_0^t}.$$

and we get

$$\frac{d}{dt}\int_{\mathbb{R}}\widehat{K_0^t} = 1 + \Im(a_1^n - a_1^{n-1}).$$

Also,

$$\frac{d}{dt}\int_{\mathbb{R}}\widehat{K_0^t\mu}\text{ sign } x = \frac{1}{2}(a_n^n - a_{n-1}^{n-1} + a_1^n - a_1^{n-1}) =$$

$$\Re(a_1^n - a_1^{n-1}).$$

Finally,

$$\int_{\mathbb{R}} \widehat{K}_0^t * (\operatorname{sign} x \cdot \widehat{\mu}) = \frac{i}{2}\int_{\mathbb{R}} \widehat{K_0^t} + \frac{i}{2}\int_{\mathbb{R}} \widehat{K_0^t} = i\int_{\mathbb{R}} \widehat{K_0^t}.$$

For β we get

$$\beta(t) = \Re(a_1^n - a_1^{n-1}) + i(1 + \Im(a_1^n - a_1^{n-1})) + c(1 + \Im(a_1^n - a_1^{n-1})).$$

Numerical calculation and the condition $\alpha\gamma - \beta^2 = 1$ suggest $c = -1 - i$ and

$$\beta = \Re(a_1^n - a_1^{n-1}) - \Im(a_1^n - a_1^{n-1}) - 1$$

on $((n-1)/2, n/2]$.

This produces the values of β on $(0, \frac{1}{2}], (\frac{1}{2}, 1], (1, \frac{3}{2}], ...$:

$$0, -4/3,\ -5/3,\ -1,\ -2/3,\ -22/21,\ -9/7,\ -1,\ -4/5,\ -56/55,\ -13/11,$$

$$-1,\ -6/7,\ -106/105,\ -17/15,\ -1,\ -8/9,\ -172/171,\ -21/19,$$

$$-1,\ -10/11,\ -254/253,\ -25/23, -1...$$

Note that the sequence for $-\beta$ displays the following pattern. Each entry number $4n$, $n = 1, 2, ...$, is equal to 1. The $4n+1$ entry, $n = 0, 1, 2, ...$ has denominator $2n+1$ and numerator $2n$. The $4n+3$ entry has the denominator $4n+3$ and numerator $4n+5$. Finally the $4n+2$ entry, between the last two, has the denominator equal to the product of their denominators, $(2n+1)(4n+3)$ and the numerator $(2n+1)(4n+3)+1$.

These and further examples of inverse spectral problems along with the necessary proofs can be found in Makarov and Poltoratski [n.d.].

References

N. I. Ahiezer (1956). "On weighted approximations of continuous functions by polynomials on the entire number axis". *Uspehi Mat. Nauk (N.S.)* 11.4(70), pp. 3–43. MR: 0084064 (cit. on p. 1777).

Andrew G. Bakan (2008). "Representation of measures with polynomial denseness in $L_p(\mathbb{R}, d\mu)$, $0 < p < \infty$, and its application to determinate moment problems". *Proc. Amer. Math. Soc.* 136.10, pp. 3579–3589. MR: 2415042 (cit. on p. 1779).

Michael Benedicks (1985). "On Fourier transforms of functions supported on sets of finite Lebesgue measure". *J. Math. Anal. Appl.* 106.1, pp. 180–183. MR: 780328.

S. Bernstein (1924). “La problème de l’approximation des fonctions continues sur tout l’axe réel et l’une de ses applications”. *Bull. Soc. Math. France* 52, pp. 399–410. MR: 1504852 (cit. on p. 1777).

A Beurling (1989). *Collected works, vols I-II*. Birkhauser, Boston (cit. on p. 1771).

A. Beurling and P. Malliavin (1962). “On Fourier transforms of measures with compact support”. *Acta Math.* 107, pp. 291–309. MR: 0147848 (cit. on pp. 1771, 1773).

Arne Beurling and Paul Malliavin (1967). “On the closure of characters and the zeros of entire functions”. *Acta Math.* 118, pp. 79–93. MR: 0209758 (cit. on pp. 1771, 1773).

Alexander Borichev and Mikhail Sodin (2011). “Weighted exponential approximation and non-classical orthogonal spectral measures”. *Adv. Math.* 226.3, pp. 2503–2545. MR: 2739783 (cit. on p. 1773).

Louis de Branges (1959). “The Bernstein problem”. *Proc. Amer. Math. Soc.* 10, pp. 825–832. MR: 0114080 (cit. on p. 1778).

– (1968). *Hilbert spaces of entire functions*. Prentice-Hall, Inc., Englewood Cliffs, N.J., pp. ix+326. MR: 0229011 (cit. on pp. 1777, 1784, 1785, 1787, 1788).

Douglas N. Clark (1972). “One dimensional perturbations of restricted shifts”. *J. Analyse Math.* 25, pp. 169–191. MR: 0301534.

H. Dym and H. P. McKean (1976). *Gaussian processes, function theory, and the inverse spectral problem*. Probability and Mathematical Statistics, Vol. 31. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, pp. xi+335. MR: 0448523.

Harry Dym (1970). “An introduction to de Branges spaces of entire functions with applications to differential equations of the Sturm-Liouville type”. *Advances in Math.* 5, pp. 395–471. MR: 0276741.

I. M. Gelfand and B. M. Levitan (1955). “On the determination of a differential equation from its spectral function”. *Amer. Math. Soc. Transl. (2)* 1, pp. 253–304. MR: 0073805.

Victor Havin and Burglind Jöricke (1994). *The uncertainty principle in harmonic analysis*. Vol. 28. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, pp. xii+543. MR: 1303780.

VP Havin (2001). “On the uncertainty principle in harmonic analysis”. In: Lecture notes.

S. V. Hruščëv, N. K. Nikolskii, and B. S. Pavlov (1981). “Unconditional bases of exponentials and of reproducing kernels”. In: *Complex analysis and spectral theory (Leningrad, 1979/1980)*. Vol. 864. Lecture Notes in Math. Springer, Berlin-New York, pp. 214–335. MR: 643384 (cit. on pp. 1771, 1780).

Paul Koosis (1988). *The logarithmic integral. I*. Vol. 12. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, pp. xvi+606. MR: 961844 (cit. on p. 1777).

M. G. KreÄn (1954). "On a basic approximation problem of the theory of extrapolation and filtration of stationary random processes". *Doklady Akad. Nauk SSSR (N.S.)* 94, pp. 13–16. MR: 0062980.

M. Krein (1945). "On a problem of extrapolation of A. N. Kolmogoroff". *C. R. (Doklady) Acad. Sci. URSS (N. S.)* 46, pp. 306–309. MR: 0012700.

Norman Levinson (1940). *Gap and Density Theorems*. American Mathematical Society Colloquium Publications, v. 26. American Mathematical Society, New York, pp. viii+246. MR: 0003208.

B. M. Levitan and I. S. Sargsjan (1991). *Sturm-Liouville and Dirac operators*. Vol. 59. Mathematics and its Applications (Soviet Series). Translated from the Russian. Kluwer Academic Publishers Group, Dordrecht, pp. xii+350. MR: 1136037.

D. S. Lubinsky (2007). "A survey of weighted polynomial approximation with exponential weights". *Surv. Approx. Theory* 3, pp. 1–105. MR: 2276420 (cit. on p. 1777).

Nikolai Makarov and Alexei Poltoratski (n.d.). "Études for the inverse spectral problem". Preprint (cit. on pp. 1786–1789, 1795).

– (2005). "Meromorphic inner functions, Toeplitz kernels and the uncertainty principle". In: *Perspectives in analysis*. Vol. 27. Math. Phys. Stud. Springer, Berlin, pp. 185–252. MR: 2215727 (cit. on pp. 1771, 1774, 1775, 1780, 1781, 1785).

– (2010). "Beurling-Malliavin theory for Toeplitz kernels". *Invent. Math.* 180.3, pp. 443–480. MR: 2609247 (cit. on pp. 1771, 1775, 1780–1782).

S. N. Mergelyan (1956). "Weighted approximations by polynomials". *Uspehi Mat. Nauk (N.S.)* 11.5(71). Translation to English in Mergelyan [1958], pp. 107–152. MR: 0083614 (cit. on pp. 1777, 1797).

– (1958). "Weighted approximations by polynomials". *Amer. Math. Soc. Translations*. 2nd ser. 10. English translation of Mergelyan [1956], pp. 59–106 (cit. on p. 1797).

Mishko Mitkovski and Alexei Poltoratski (2010). "Pólya sequences, Toeplitz kernels and gap theorems". *Adv. Math.* 224.3, pp. 1057–1070. MR: 2628803 (cit. on p. 1777).

Nikolai K. Nikolski (2002). *Operators, functions, and systems: an easy reading. Vol. 2*. Vol. 93. Mathematical Surveys and Monographs. Model operators and systems, Translated from the French by Andreas Hartmann and revised by the author. American Mathematical Society, Providence, RI, pp. xiv+439. MR: 1892647.

N. K. Nikolskii (1986). *Treatise on the shift operator*. Vol. 273. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Spectral function theory, With an appendix by S. V. Hruščev [S. V. Khrushchëv] and V. V. Peller, Translated from the Russian by Jaak Peetre. Springer-Verlag, Berlin, pp. xii+491. MR: 827223 (cit. on p. 1771).

Joaquim Ortega-Cerdà and Kristian Seip (2002). "Fourier frames". *Ann. of Math. (2)* 155.3, pp. 789–806. MR: 1923965 (cit. on p. 1786).

B. S. Pavlov (1979). “The basis property of a system of exponentials and the condition of Muckenhoupt”. *Dokl. Akad. Nauk SSSR* 247.1, pp. 37–40. MR: 545940 (cit. on pp. 1771, 1780).

Alexei Poltoratski (n.d.). “Type alternative for Frostman measures”. Preprint (cit. on p. 1779).

– (2012). “Spectral gaps for sets and measures”. *Acta Math.* 208.1, pp. 151–209. MR: 2910798 (cit. on pp. 1775, 1777).

– (2013). “A problem on completeness of exponentials”. *Ann. of Math. (2)* 178.3, pp. 983–1016. MR: 3092474 (cit. on p. 1779).

– (2015a). “Bernstein’s problem on weighted polynomial approximation”. In: *Operator-related function theory and time-frequency analysis*. Vol. 9. Abel Symp. Springer, Cham, pp. 147–171. MR: 3561432 (cit. on p. 1778).

– (2015b). *Toeplitz approach to problems of the uncertainty principle*. Vol. 121. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, pp. viii+216. MR: 3309830 (cit. on pp. 1773–1779).

– (2017). “Toeplitz Order”. arXiv: 1706.08002 (cit. on pp. 1780, 1782, 1787).

Alexei Poltoratski and Donald Sarason (2006). “Aleksandrov-Clark measures”. In: *Recent advances in operator-related function theory*. Vol. 393. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 1–14. MR: 2198367.

Kristian Seip (2004). *Interpolation and sampling in spaces of analytic functions*. Vol. 33. University Lecture Series. American Mathematical Society, Providence, RI, pp. xii+139. MR: 2040080 (cit. on p. 1786).

Georges Valiron (1925). “Sur la formule d’interpolation de Lagrange”. *Bull. Sci. Math* 49.2, pp. 181–192.

Hermann Weyl (1977). *Gruppentheorie und Quantenmechanik*. second. Wissenschaftliche Buchgesellschaft, Darmstadt, pp. xi+366. MR: 0450450.

Norbert Wiener (1924). “The quadratic variation of a function and its Fourier coefficients”. *Studies in Applied Mathematics* 3.2, pp. 72–94.

– (1964). *I Am A Mathematician*.

Norbert Wiener, Norbert Wiener, Cyberneticist Mathematician, Norbert Wiener, Norbert Wiener, and Cybernéticien Mathématicien (1949). *Extrapolation, interpolation, and smoothing of stationary time series*. MIT press Cambridge.

Received 2017-11-29.

ALEXEI POLTORATSKI

Texas A&M University
Department of Mathematics
College Station, TX 77843
USA

and

Department of Mathematics and Mechanics
St. Petersburg State University
St. Petersburg
Russia
apoltora@gmail.com
alexeip@math.tamu.edu

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1797–1818)

FINITARY APPROXIMATIONS OF GROUPS AND THEIR APPLICATIONS

ANDREAS THOM

Abstract

In these notes we will survey recent results on various finitary approximation properties of infinite groups. We will discuss various restrictions on groups that are approximated for example by finite solvable groups or finite-dimensional unitary groups with the Frobenius metric. Towards the end, we also briefly discuss various applications of those approximation properties to the understanding of the equational theory of a group.

1 Introduction

1.1 The setup. Let Γ be a finitely presented group, given by a finite generating set $X := \{x_1, \dots, x_k\}$ and a finite set $R \subset \mathbf{F}_k = \langle X \rangle$ of relations, i.e. $\Gamma := \langle X | R \rangle = \mathbf{F}_k / N$, where $\mathbf{F}_k$ denotes the free group on X and $N = \langle\langle R \rangle\rangle$ the normal subgroup generated by R. Throughout the article, X and R will be fixed.

Very basic questions about Γ are usually hard to answer unless Γ can be realized as a group of symmetries of a sufficiently concrete object, such as a finite set, a finite-dimensional vector space or a metric space with suitable properties. In the easiest situation, maybe Γ is residually finite, i.e. for any finite subset $F \subset \Gamma$, there exists a homomorphism to a finite group $\varphi \colon \Gamma \to H$, such that the restriction of φ to F is injective. In order to overcome the algebraic and arithmetic obstruction to the existence of finite quotients and finite-dimensional unitary representations it is worthwhile to relax these notions. Informally speaking, we will seek for asymptotic homomorphisms from Γ with values in a family of (typically compact) metric groups. Let (G, d) be a metric group and assume throughout the entire article that $d : G \times G \to [0, \infty)$ is bi-invariant, i.e.

$$d(gh, gk) = d(h, k) = d(hg, kg), \quad \forall g, h, k \in G.$$

This research was supported by ERC Consolidator Grant No. 681207.
MSC2010: 20F69.

Note that any bi-invariant metric is uniquely determined by the associated *invariant length function* $\ell(g) = d(1, g)$, which is a subadditive, symmetric and conjugation invariant, $[0, \infty)$-valued function on G that takes the value 0 only at 1_G.

Well-known invariant length functions in this context include the normalized Hamming distance on symmetric groups or various length functions induced by unitarily invariant norms on groups of unitary matrices. Somewhat similarly, there is a rank length function, which is defined if $G \leq \mathrm{GL}_n(q)$ for some $n \in \mathbb{N}$ and q a prime power: $\ell_G^{\mathrm{rank}}(g) := n^{-1} \cdot \mathrm{rk}(1-g)$. Another important example is the conjugacy length function, which is defined by

$$\ell^{\mathrm{conj}}(g) := \log_{|G|}|g^G|$$

for $g \in G$, where g^G is the conjugacy class of $g \in G$, and G is a finite and centerless group. If G is the alternating group or a finite simple group of Lie type, then the conjugacy metric is comparable to the more geometrically defined metrics above. A fundamental result in the work of Liebeck and Shalev [2001] says that the conjugacy length is intrinsically tied to the algebraic properties of G if G is simple: indeed, there exists an absolut constant $c > 0$, such that $(g^G)^{ck} = G$ if $k > \ell^{\mathrm{conj}}(G)^{-1}$. Thus, the conjugacy length is also comparable to the normalized word metric w.r.t. any sufficiently small conjugacy class. It turns out that there is essentially just one invariant length function up to a suitable notion of equivalence on a finite simple group.

1.2 Approximation and stability.

Now we can define more precisely what we mean by metric approximation of an abstract group by a class of metric groups $\mathcal{C}$.

Definition 1.1. *A group* Γ *is called* $\mathcal{C}$-approximated *if there is a length function* $\delta\colon \Gamma \to [0, \infty)$ *such that for any finite subset* $S \subseteq \Gamma$ *and* $\varepsilon > 0$ *there exist a group* $(G, d) \in \mathcal{C}$ *and a map* $\varphi\colon \Gamma \to G$*, such that*

(*i*) *if* $g, h, gh \in S$*, then* $d(\varphi(g)\varphi(h), \varphi(gh)) < \varepsilon$ *and*

(*ii*) *for* $g \in S$ *we have* $d(1_H, \varphi(g)) \geq \delta(g)$.

In some situations, we will only fix a class $\mathcal{C}$ of groups and let the choice of bi-invariant metrics be arbitrary – we will also speak about $\mathcal{C}$-approximability in this context. The previous definition has emerged from various contexts, including an influential work of Ulam [1960], the work of Connes [1976], Gromov [1999], Weiss [2000] and later work Thom [2012], Holt and Rees [2016], and Glebsky [2016].

Let us also introduce the closely related notion of asymptotic homomorphisms. Note that any map $\varphi\colon X \to G$, for some $(G, d) \in \mathcal{C}$, uniquely determines a homomorphism $\mathbf{F}_k \to G$ which we will also denote by φ.

Definition 1.2. *Let $(G, d) \in \mathcal{C}$ and let $\varphi, \psi \colon X \to G$ be maps. The* defect *of φ is defined by*

$$\operatorname{def}(\varphi) := \max_{r \in R} d(\varphi(r), 1_G).$$

The distance *between φ and ψ is defined by*

$$\operatorname{dist}(\varphi, \psi) = \max_{1 \le i \le k} d(\varphi(x_i), \psi(x_i)).$$

The homomorphism distance *of φ is defined by*

$$\operatorname{HomDist}(\varphi) := \inf_{\pi \in \operatorname{Hom}(\Gamma, G)} \operatorname{dist}(\varphi, \pi|_X).$$

Definition 1.3. *A sequence of maps $\varphi_n \colon X \to G_n$, for $(G_n, d_n) \in \mathcal{C}$, is called an* asymptotic homomorphism *with values in $\mathcal{C}$ if*

$$\lim_{n \to \infty} \operatorname{def}(\varphi_n) = 0.$$

Definition 1.4. *Let $G_n \in \mathcal{C}, n \in \mathbb{N}$. Two sequences $\varphi_n, \psi_n \colon X \to G_n$ are called equivalent if*

$$\lim_{n \to \infty} \operatorname{dist}(\varphi_n, \psi_n) = 0.$$

If an asymptotic homomorphism $(\varphi_n)_{n \in \mathbb{N}}$ is equivalent to a sequence of homomorphisms, we call $(\varphi_n)_{n \in \mathbb{N}}$ trivial.

It is easy to see that a group $\Gamma = \langle X | R \rangle$ is $\mathcal{C}$-approximated if and only if there is an asymptotic homomorphism with values in $\mathcal{C}$ that separates elements in a suitable sense. Note that the existence of a finite presentation is assumed mostly for convenience. Moreover, note that it is easy to see that the property of being $\mathcal{C}$-approximated depends only on Γ and not on the finite presentation.

Let us discuss some examples of $\mathcal{C}$-approximated groups. We denote by **Alt** (resp. **Fin**) the class of finite alternating groups (resp. the class of all finite groups). A group is called sofic (resp. weakly sofic) if and only if it is **Alt**-approximated (resp. **Fin**-approximated) as an abstract group, see Glebsky [2016]. The class of sofic groups is of central interest in group theory. Indeed, eversince the work of Gromov [1999] on Gottschalk's Surjunctivity Conjecture, the class of sofic groups has attracted much interest in various areas of mathematics. Major applications of this notion arose in the work of Elek and Szabó [2004] on Kaplansky's Direct Finiteness Conjecture, on Lück's Determinant Conjecture by Elek and Szabó [2005], and more recently in joint work of Klyachko and Thom [2017] on generalizations of the Kervaire-Laudenbach Conjecture and Howie's Conjecture. Despite considerable effort, no non-sofic group has been found so far – whether all groups are sofic is one of the outstanding open problems in group theory.

Question 1.5 (Gromov)**.** Are all groups sofic?

Examples of sofic groups which fail to be locally residually amenable are given in Cornulier [2011] and Kar and Nikolov [2014] (see also Thom [2010]), answering a question of Gromov [1999].

Groups approximated by certain classes of finite simple groups of Lie type have been studied in Arzhantseva and Păunescu [2017] and Thom and Wilson [2014, 2016]. We will discuss approximation by groups of unitary matrices, a central topic in the theory of operator algebras and free probability theory, at length in Section 3.

Non-approximation results are rare, however, in Thom [2012] it was proved that the so-called Higman group cannot be approximated by finite groups with commutator-contractive invariant length functions. In Howie [1984] Howie presented a group which (by a result of Glebsky and Rivera [2008]) turned out not to be approximated by finite nilpotent groups with arbitrary invariant length function. In Sections 2 and 3, we will survey more general results of this type that have been proved recently in Nikolov, Schneider, and Thom [2017].

A central definition in the present context is the notion of stability that was studied in De Chiffre, Glebsky, Lubotzky, and Thom [2017].

Definition 1.6. *The group* Γ *is called* $\mathcal{C}$-stable *if all asymptotic homomorphisms with values in* $\mathcal{C}$ *are trivial, that is: for all* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that* $\operatorname{def}(\varphi) < \delta$ *implies* $\operatorname{HomDist}(\varphi) < \varepsilon$ *for all* $\varphi \colon X \to (G, d)$ *with* $(G, d) \in \mathcal{C}$.

It is clear that any group Γ that is both $\mathcal{C}$-approximated and $\mathcal{C}$-stable is residually a subgroup of groups in $\mathcal{C}$. If $\mathcal{C}$ consists of finite groups or more generally compact groups this readily implies that Γ must be residually finite. This observation has been used in De Chiffre, Glebsky, Lubotzky, and Thom [ibid.] to prove non-approximation results by proving that certain groups are $\mathcal{C}$-stable but not residually $\mathcal{C}$, see Theorem 3.2.

1.3 Metric ultraproducts.

Throughout these notes, we fix a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$. Let $(G_n)_{n\in\mathbb{N}}$ be a family of groups, all equipped with bi-invariant metrics d_n. In this case, the subgroup

$$N = \left\{ (g_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} G_n \,\middle|\, \lim_{n\to\mathcal{U}} d_n(g_n, 1_{G_n}) = 0 \right\}$$

of the direct product $\prod_{n\in\mathbb{N}} G_n$ is normal, so that we can define the *metric ultraproduct*

$$\prod_{n\to\mathcal{U}} (G_n, d_n) := \prod_{n\in\mathbb{N}} G_n \Big/ N.$$

The relevance of metric ultraproducts becomes apparent in the following folklore result:

Proposition 1.7. *Let $\mathcal{C}$ is a class of metric groups. A group $\Gamma = \langle X | R \rangle$ is $\mathcal{C}$-approximated if and only if it is isomorphic to a subgroup of a metric ultraproduct of $\mathcal{C}$-groups.*

For more details on the algebraic and geometric structure of such ultraproducts see also Stolz and Thom [2014] and Thom and Wilson [2014, 2016] and Schneider and Thom [2017]. In view of Proposition 1.7 it is natural to generalize the notion of a $\mathcal{C}$-approximated group to topological groups using ultraproducts:

Definition 1.8. *A topological group is called* $\mathcal{C}$-approximated *if it is topologically isomorphic to a closed subgroup of a metric ultraproduct of $\mathcal{C}$-groups.*

We will constrain ourselves to Polish groups and countable ultraproducts, but that is just for convenience. Typically, for example in the context of sofic or weakly sofic groups, it is easy to see that an abstract $\mathcal{C}$-approximated group is also $\mathcal{C}$-approximated when viewed as a topological group with the discrete topology. It is clear that any **Fin**-approximated topological group must admit a bi-invariant metric that induces the topology. We will discuss various less obvious restrictions on **Fin**-approximability in Section 2.3.

In view of the definition of metric ultraproducts, any approximation property for a group Γ by a class of compact groups leads to an embedding into a quotient of a compact group.

Question 1.9. Is any group a sub-quotient of a compact group?

2 Weak soficity and the pro-finite topology

2.1 Connections with the pro-finite topology. In this section, we want to survey some recent results that were proved in joint work with Nikolov and Schneider, see Nikolov, Schneider, and Thom [2017]. The main insight that helped us was to combine the relationship between soficity and properties of the pro-finite topology that was established by Glebsky and Rivera [2008] with the deep work on finite groups by Segal [2009] and Nikolov and Segal [2012]. Let $\mathcal{C}$ be a class of finite groups. Adapting Theorem 4.3 of Glebsky and Rivera [2008], one can prove the following theorem relating $\mathcal{C}$-approximated groups to properties of the pro-$\mathcal{C}$ topology on a free group:

Theorem 2.1 (Nikolov, Schneider, and Thom [2017])**.** *Let $\mathbf{F}_k/N$ be a presentation of a group Γ. Then, if Γ is $\mathcal{C}$-approximated, for each finite sequence $n_1, \ldots, n_m \in N$ it holds that*

$$\overline{n_1^{\mathbf{F}} \cdots n_m^{\mathbf{F}}} \subseteq N,$$

where the closure is taken in the pro-$\mathcal{C}$ topology on $\mathbf{F}_k$. The converse holds under mild assumptions on $\mathcal{C}$.

The coarsest such topology on $\mathbf{F}_k$ is of course the pro-finite topology and at the time of writing of Glebsky and Rivera [2008] it was an open problem to decide whether a finite product of conjugacy classes in a non-abelian free group is always closed in this topology. As has been remarked in Nikolov, Schneider, and Thom [2017], it is a rather straightforward consequence of the work of Nikolov-Segal (see Nikolov and Segal [2012] or Theorem 2.7) that this is not the case. Indeed, one of their main results implies that in $\mathbf{F}_k$ the profinite closure of a finite product of conjugacy classes of $x_1^{-1}, x_1, \ldots, x_k^{-1}, x_k$ contains the entire commutator subgroup, while it is a well known fact (see Theorem 3.1.2 of Segal [2009]) that the commutator width of $\mathbf{F}_k$ is infinite if $k > 1$. This implication was first observed by Segal and independently discovered by Gismatullin. In view of this observation it seems unlikely that the pro-finite closure is always contained in the normal closure, but this remains an open problem.

Question 2.2 (Glebsky and Rivera [2008])**.** Let $n_1, \ldots, n_m \in \mathbf{F}_k$. Is it true, that

$$\overline{n_1^{\mathbf{F}} \cdots n_m^{\mathbf{F}}} \subseteq \langle\langle n_1, \ldots, n_m \rangle\rangle,$$

where the closure is taken in the pro-finite topology on $\mathbf{F}_k$?

In general, there are quite a number of mysteries that can be formulated in terms of closure properties of the pro-$\mathcal{C}$ topology for particular more restricted families of groups. Indeed, let us just mention a question from Herwig and Lascar [2000].

Question 2.3 (Herwig and Lascar [ibid.])**.** It is easy to see that if a finitely generated subgroup $H < \mathbf{F}_k$ is closed in the pro-odd topology, then it satisfies $a^2 \in H \Rightarrow a \in H$. Is the converse true?

2.2 Approximation by classes of finite groups.

For more restricted families of groups, the answer to Question 2.2 becomes negative. Indeed, let **Sol** (resp. **Nil**) be the class of finite solvable (resp. nilpotent) groups. In view of Theorem 2.1 this implies that there are groups which are not **Sol**-approximated. More precisely, we proved:

Theorem 2.4 (Nikolov, Schneider, and Thom [2017])**.** *Every finitely generated* **Sol***-approximated group has a non-trivial abelian quotient.*

As a consequence, perfect groups cannot be **Sol**-approximated and a finite group is **Sol**-approximated if and only if it is solvable. Indeed, any finite solvable group is clearly **Sol**-approximated and on the other hand, a non-solvable finite group contains a non-trivial perfect subgroup and hence cannot be **Sol**-approximated by Theorem 2.4.

Initially, Howie proved in Howie [1984] that the group $\langle x, y | x^{-2}y^{-3}, x^{-2}(xy)^5 \rangle$ is not **Nil**-approximated. We followed his proof for any non-trivial finitely generated perfect group and then extended it in Nikolov, Schneider, and Thom [2017] and established that these groups are not even **Sol**-approximated using techniques of Segal [2000, 2009].

Note that finite generation is crucial in the statement of Theorem 2.4. Indeed, there exist countably infinite locally finite-p groups which are perfect and even characteristically simple, see McLain [1954]. These groups are **Nil**-approximated, since finite p-groups are nilpotent. It is known that locally finite-solvable groups cannot be non-abelian simple, but it seems to be an open problem if there exist **Sol**-approximated simple groups.

Let us also remark, that the assumptions of Theorem 2.4 are not enough to conclude that a finitely generated **Sol**-approximated group has an infinite solvable quotient. Indeed, consider a suitable congruence subgroup of $\mathrm{SL}(3, \mathbb{Z})$, which is residually p-finite and thus even **Nil**-approximated and has Kazhdan's property (T) – thus all amenable quotients are finite. However, the following seems to be an open problem.

Question 2.5. Is every finitely presented and **Sol**-approximated group residually finite-solvable?

Even more, it could be that all finitely presented groups are **Sol**-stable (in a suitable sense). A positive answer to the previous question would be in sharp contrast to other forms of approximability. Indeed, for example the Baumslag-Solitar group $\mathrm{BS}(2, 3)$ is not residually finite, but residually solvable and hence sofic and in particular **Fin**-approximated.

Let us finish this section by mentioning some structure result on the class of **Fin**-approximated groups. Let **PSL** be the class of simple groups of type $\mathrm{PSL}(n, q)$, i.e. $n \in \mathbb{N}_{\geq 2}$ and q is a prime power and $(n, q) \neq (2, 2), (2, 3)$, and recall that **Fin** is the class of all finite groups. In Nikolov, Schneider, and Thom [2017] we prove the following result.

Theorem 2.6 (Nikolov, Schneider, and Thom [ibid.]). *Any non-trivial finitely generated* **Fin***-approximated group has a non-trivial* **PSL***-approximated quotient. In particular, every finitely generated simple and* **Fin***-approximated group is* **PSL***-approximated.*

The proof makes use of seminal results of Liebeck and Shalev [2001] and Nikolov and Segal [2012]. The previous result may be seen as a first step towards a proof that all **Fin**-approximated groups are sofic.

2.3 Approximability of Lie groups. Let us explain how a theorem of Nikolov-Segal allows us to deduce two results concerning the approximability of Lie groups by finite groups and one result on compactifications of pseudofinite groups.

Theorem 2.7 (Theorem 1.2 of Nikolov and Segal [2012]). *Let $g_1, \ldots, g_m$ be a symmetric generating set for the finite group G. If $K \trianglelefteq G$, then*

$$[K, G] = \left(\prod_{j=1}^{m} [K, g_j] \right)^e,$$

where e only depends on m.

The following result is an immediate corollary of Theorem 2.7.

Corollary 2.8 (Nikolov, Schneider, and Thom [2017]). *Let G be a finite group, then for $g, h \in G$ and $k \in \mathbb{N}$ we have*

$$[g^k, h^k] \in \big([G, g][G, g^{-1}][G, h][G, h^{-1}]\big)^e$$

for some fixed constant $e \in \mathbb{N}$ that is independent of G.

We deduce immediately that the same conclusion holds for any quotient of a product of finite groups and in particular, for any metric ultraproduct of finite groups. Combining the finitary approximation with the local geometry of Lie groups we obtain the following consequence.

Theorem 2.9 (Nikolov, Schneider, and Thom [ibid.]). *A connected Lie group is* **Fin**-*approximated as a topological group if and only if it is abelian.*

Indeed, this is a direct consequence of the following auxiliary result:

Lemma 2.10. *Let $\varphi, \psi : \mathbb{R} \to (H, \ell) = \prod_{\mathfrak{U}} (H_i, \ell_i)$ be continuous homomorphisms into a metric ultraproduct of finite metric groups with bi-invariant metrics. Then, the images of φ and ψ commute.*

Note that Theorem 2.9 provides an answer to Question 2.11 of Doucha [n.d.] whether there are groups with invariant length function that do not embed in a metric ultraproduct of finite groups with invariant length function. Note also that the topology matters a lot in this context. Indeed, it can be shown that any compact Lie group is a discrete subgroup of a countable metric ultraproduct of finite groups, see Nikolov, Schneider, and Thom [2017].

When one restricts the class of finite groups further and approximates with symmetric groups, one can not even map the real line $\mathbb{R}$ non-trivially and continuously to a metric ultraproduct of such groups with invariant length function. Indeed, for the symmetric group $\mathrm{Sym}(n)$, it can be shown that all invariant length functions ℓ on it satisfy $\ell(\sigma^k) \leq 3\ell(\sigma)$, for every $k \in \mathbb{Z}$ and $\sigma \in \mathrm{Sym}(n)$. Using this identity, it is simple to deduce that the only continuous homomorphism of $\mathbb{R}$ into a metric ultraproduct of finite symmetric groups with invariant length function is trivial.

Referring to a question of Zilber [2014, p. 17] (see also Question 1.1 of Pillay [2017]) whether a compact simple Lie group can be a quotient of the algebraic ultraproduct of finite groups, we obtained the following second application of Corollary 2.8:

Theorem 2.11 (Nikolov, Schneider, and Thom [2017]). *Let G be a Lie group equipped with an bi-invariant metric generating its topology. If G is an abstract quotient of a product of finite groups, then G has abelian identity component.*

The proof of this result is almost identical to the proof of Theorem 2.9. Theorem 2.11 implies that any compact simple Lie group, the simplest example being $\mathrm{SO}(3, \mathbb{R})$, is not a quotient of a product of finite groups, answering the questions of Zilber and Pillay. Note also that these results are vast generalizations of an ancient result of Turing [1938].

Moreover, Theorem 2.11 remains valid if we replace the product of finite groups by a *pseudofinite group*, i.e. a group which is a model of the theory of all finite groups. It then also provides a negative answer to Question 1.2 of Pillay [2017], whether there is a surjective homomorphism from a pseudofinite group to a compact simple Lie group.

By a *compactification* of an abstract group G, we mean a compact group C together with a homomorphism $\iota\colon G \to C$ with dense image. Pillay conjectured that the Bohr compactification (i.e. the universal compactification) of a pseudofinite group has abelian identity component (Conjecture 1.7 in Pillay [ibid.]). We answer this conjecture in the affirmative by the following result:

Theorem 2.12 (Nikolov, Schneider, and Thom [2017]). *Let G be a pseudofinite group. Then the identity component of any compactification $\iota\colon G \to C$ is abelian.*

Again, the proof is an application of Corollary 2.8.

3 Approximation by unitary matrices

3.1 The choice of the metric. We will now focus on approximation of groups by unitary matrices. Today, the theme knows many variations, ranging from operator-norm approximations that appeared in the theory of operator algebras Blackadar and Kirchberg [1997] and Carrion, Dadarlat, and Eckhardt [2013] to questions related to Connes' Embedding Problem, see Connes [1976] and Pestov [2008] for details. Several examples of this situation have been studied in the literature:

(1) $G_n = \mathrm{U}(n)$, where the metric d_n is induced by the Hilbert-Schmidt norm $\|T\|_{\mathrm{HS}} = \sqrt{n^{-1}\sum_{i,j=1}^{n}|T_{ij}|^2}$. In this case, approximated groups are sometimes called hyperlinear Pestov [2008], but we choose to call them Connes-embeddable.

(2) $G_n = \mathrm{U}(n)$, where the metric d_n is induced by the operator norm $\|T\|_{\mathrm{op}} = \sup_{\|v\|=1} \|Tv\|$. In this case, groups which are $(G_n, d_n)_{n=1}^{\infty}$-approximated are called MF, see Carrion, Dadarlat, and Eckhardt [2013].

(3) $G_n = \mathrm{U}(n)$, where the metric d_n is induced by the *unnormalized* Hilbert-Schmidt norm $\|T\|_{\mathrm{Frob}} = \sqrt{\sum_{i,j=1}^{n} |T_{ij}|^2}$, also called Frobenius norm. We will speak about Frobenius-approximated groups in this context, see De Chiffre, Glebsky, Lubotzky, and Thom [2017].

Let us emphasize that the approximation properties are *local* in the sense that only finitely many group elements and their relations have to be considered at a time. This is in stark contrast to the *uniform* situation, which – starting with the work of Grove, Karcher, and Ruh [1974] and Kazhdan [1982] – is much better understood, see Burger, Ozawa, and Thom [2013] and Ozawa, Thom, and De Chiffre [2017].

Again, there are longstanding problems that ask if *any* group exists which is not approximated in the sense of (1), a problem closely related to Connes' Embedding Problem Connes [1976] and Pestov [2008].

Question 3.1 (Connes [1976]). Is every discrete group Connes-embeddable?

Connes' Embedding Problem has many incarnations and we want to mention only a few of them, see Ozawa [2004] and Pestov [2008] for more details. The most striking alternative formulation is due to Kirchberg, who showed that Connes' Embedding Problem has an affirmative answer if and only if the group $\mathbf{F}_2 \times \mathbf{F}_2$ is residually finite dimensional, i.e. if the finite-dimensional unitary representations of this group are dense in the unitary dual equipped with the Fell topology.

Blackadar and Kirchberg [1997] conjectured that any stably finite C^*-algebra is embeddable into an norm-ultraproduct of matrix algebras, implying a positive answer to the approximation problem in the sense of (2) for any group. Recent breakthrough results imply that any amenable group is MF, i.e. approximated in the sense of (2), see Tikuisis, White, and Winter [2017].

Approximation in the sense of (3) is known to be more restrictive – as has been shown in De Chiffre, Glebsky, Lubotzky, and Thom [2017]. Indeed, in joint work with De Chiffre, Glebsky, Lubotzky, and Thom [ibid.], a conceptually new technique was introduced, that allowed to provide groups that are not approximated in the sense of (3) above. An analogous result for the normalized Frobenius norm would answer the Connes' Embedding Problem. Even though we had little to say about Connes' Embedding Problem, we believe that we provided a promising new angle of attack.

Theorem 3.2 (De Chiffre, Glebsky, Lubotzky, and Thom [ibid.]). *There exists a finitely presented group, which is not Frobenius-approximated. Specifically, we can take a certain* *central extension of a lattice in* $\mathrm{U}(2n) \cap \mathrm{Sp}(2n, \mathbb{Z}[i, 1/p])$ *for a large enough prime* p *and* $n \geq 3$.

The key insight was that there exists a cohomological obstruction (in the second cohomology with coefficients in a certain unitary representation) to the possibility of improving the asymptotic homomorphism. The use of cohomological obstructions goes in essence back to the pioneering work Kazhdan [1982] on stability of (uniform) approximate representations of amenable groups. The main result of De Chiffre, Glebsky, Lubotzky, and Thom [2017] is the following theorem.

Theorem 3.3 (De Chiffre, Glebsky, Lubotzky, and Thom [ibid.]). *Let* Γ *be a finitely presented group such that*

$$H^2(\Gamma, \mathcal{H}_\pi) = 0$$

for every unitary representation $\pi : \Gamma \to \mathrm{U}(\mathcal{H}_\pi)$. *Then, any asymptotic homomorphism* $\varphi_n : \Gamma \to \mathrm{U}(n)$ *w.r.t. the Frobenius norm is asymptotically close to a sequence of homomorphisms, i.e.* Γ *is Frobenius-stable.*

It is well-known that a discrete group has Kazhdan's property (T) if and only if $H^1(\Gamma, \mathcal{H}_\pi) = 0$ for all unitary representations. In De Chiffre, Glebsky, Lubotzky, and Thom [ibid.], the notion of a group to be n-Kazhdan was introduced as a vanishing condition of cohomology in dimension n with coefficients in arbitrary unitary representations, see Definition 3.8. Important work by Garland [1973] and Ballmann and Świątkowski [1997] provides first examples of 2-Kazhdan groups. However, those groups all act on Bruhat-Tits buildings of higher rank and thus are residually finite. The remaining delicate work was then to show that nevertheless there *do* exist finitely presented groups which are 2-Kazhdan and are *not* residually finite. The method in De Chiffre, Glebsky, Lubotzky, and Thom [2017] is based on Deligne's construction Deligne [1978] of a non-residually finite central extension of a $\mathrm{Sp}(2n, \mathbb{Z})$.

Before we outline the definition of the cohomological obstruction to the possibility of improving an asymptotic homomorphism and consider a few examples, let us mention a few open questions.

Question 3.4. Are all amenable (or even all nilpotent or solvable) groups Frobenius-approximated?

Question 3.5. Is the class of Frobenius-approximated groups closed under central quotients or crossed products by $\mathbb{Z}$, compare with Ozawa, Rørdam, and Sato [2015] and Thom [2010]?

The analogue of Theorem 3.3 also holds for approximation in the sense of (2) above. However, the corresponding cohomology vanishing results in order to apply the theorem in a non-trivial situation are not available. Note that Kirchberg's conjecture discussed above implies that MF-stable groups should be Ramanujan in the sense of Lubotzky and Shalom [2004].

3.2 Cohomological obstructions to stability. In this section, we want to outline how a cohomological obstruction to stability can be obtained. Consider the family of matrix algebras $\mathrm{M}_n(\mathbb{C})$ equipped with some unitarily invariant, submultiplicative norms $\|\cdot\|_n$, say the Frobenius norms. We consider the ultraproduct Banach space

$$\mathrm{M}_{\mathfrak{U}} := \prod_{n\to\mathfrak{U}} (\mathrm{M}_n(\mathbb{C}), \|\cdot\|_n),$$

and the metric ultraproduct

$$\mathrm{U}_{\mathfrak{U}} := \prod_{n\to\mathfrak{U}} (\mathrm{U}(n), d_{\|\cdot\|_n}).$$

We can associate an element $[\alpha] \in H^2(\Gamma, \prod_{n\to\mathfrak{U}}(\mathrm{M}_n(\mathbb{C}), \|\cdot\|))$ to an asymptotic representation $\varphi_n : X \to \mathrm{U}(n)$. This is done so that if $[\alpha] = 0$, then the defect can be diminished in the sense that there is an equivalent asymptotic representation φ'_n with effectively better defect, more precisely $\mathrm{def}(\varphi'_n) = o_{\mathfrak{U}}(\mathrm{def}(\varphi_n))$.

Note that an asymptotic representation as above induces a homomorphism $\varphi_{\mathfrak{U}} : \Gamma \to \mathrm{U}_{\mathfrak{U}}$ on the level of the group Γ. Thus Γ acts on $\mathrm{M}_{\mathfrak{U}}$ through $\varphi_{\mathfrak{U}}$. We consider a section $\sigma : \Gamma \to \mathbf{F}_k$ of the natural surjection $\mathbf{F}_k \to \Gamma$ and have $\sigma(g)\sigma(h)\sigma(gh)^{-1} \in \langle\langle R \rangle\rangle$ for all $g, h \in \Gamma$. We set $\tilde{\varphi}_n = \varphi_n \circ \sigma$.

Let us now outline how to define an element in $H^2(\Gamma, \mathrm{M}_{\mathfrak{U}})$ associated to φ_n. To this end we define $c_n := c_n(\varphi_n) : \Gamma \times \Gamma \to \mathrm{M}_n(\mathbb{C})$ by

$$c_n(g, h) = \frac{\tilde{\varphi}_n(g)\tilde{\varphi}_n(h) - \tilde{\varphi}_n(gh)}{\mathrm{def}(\varphi_n)},$$

for all $n \in \mathbb{N}$ such that $\mathrm{def}(\varphi_n) > 0$ and $c_n(g, h) = 0$ otherwise, for all $g, h \in \Gamma$.

Then, it follows that for every $g, h \in \Gamma$, $c_n(g, h)$ is a bounded sequence, so that the sequence defines a map

$$c = (c_n)_{n\in\mathbb{N}} : \Gamma \times \Gamma \to \mathrm{M}_{\mathfrak{U}}.$$

The map c is a Hochschild 2-cocycle with values in the Γ-module $\mathrm{M}_{\mathfrak{U}}$ and $\alpha(g, h) := c(g, h)\varphi_{\mathfrak{U}}(gh)^*$ is a 2-cocycle in the usual group cohomology. We call α the cocycle associated to the sequence $(\varphi_n)_{n\in\mathbb{N}}$.

Assume now that α represents the trivial cohomology class in $H^2(\Gamma, \mathrm{M}_{\mathfrak{U}})$, i.e. there exists a map $\beta : \Gamma \to \mathrm{M}_{\mathfrak{U}}$ satisfying

$$\alpha(g, h) = \varphi_{\mathfrak{U}}(g)\beta(h)\varphi_{\mathfrak{U}}(g)^* - \beta(gh) + \beta(g), \qquad g, h \in \Gamma.$$

Then, we have $\beta(1_\Gamma) = 0$, $\beta(g) = -\varphi_{\mathcal{U}}(g)\beta(g^{-1})\varphi_{\mathcal{U}}(g)^*$ and

$$c(g,h) = \varphi_{\mathcal{U}}(g)\beta(h)\varphi_{\mathcal{U}}(h) - \beta(gh)\varphi_{\mathcal{U}}(gh) + \beta(g)\varphi_{\mathcal{U}}(gh).$$

Furthermore, we can choose $\beta(g)$ to be skew-symmetric for all $g \in \Gamma$. Now let β be as above and let $\beta_n : \Gamma \to \mathrm{M}_n(\mathbb{C})$ be any bounded and skew-symmetric lift of β. Then $\exp(-\operatorname{def}(\varphi_n)\beta_n(g))$ is a unitary for every $g \in \Gamma$, so we can define a sequence of maps $\psi_n : \Gamma \to \mathrm{U}(n)$ by

$$\psi_n(g) = \exp(-\operatorname{def}(\varphi_n)\beta_n(g))\tilde{\varphi}_n(g).$$

Note that since $\tilde{\varphi}_n(1_\Gamma) = 1_n$ and $\beta_n(1_\Gamma) = 0$, we have $\psi_n(1_\Gamma) = 1_n$. It follows easily that $\psi_n|_X$ is an asymptotic representation with

$$\operatorname{def}(\psi_n|_X) = O_{\mathcal{U}}(\operatorname{def}(\varphi_n)),$$

but we prove that the defect of $\psi_n|_X$ is actually $o_{\mathcal{U}}(\operatorname{def}(\varphi_n))$. If we define the asymptotic representation $\varphi'_n : X \to \mathrm{U}(n)$ by $\varphi'_n = \psi_n|_X$, the conclusion can be summarized as follows:

Theorem 3.6 (De Chiffre, Glebsky, Lubotzky, and Thom [2017]). *Let $\Gamma = \langle X | R\rangle$ be a finitely presented group and let $\varphi_n : X \to \mathrm{U}(n)$ be an asymptotic representation with respect to a family of submultiplicative, unitarily invariant norms. Assume that the associated 2-cocycle $\alpha = \alpha(\varphi_n)$ is trivial in $H^2(\Gamma, \mathrm{M}_{\mathcal{U}})$. Then there exists an asymptotic representation $\varphi'_n : X \to \mathrm{U}(n)$ such that*

$$\operatorname{dist}(\varphi_n, \varphi'_n) = O_{\mathcal{U}}(\operatorname{def}(\varphi_n)) \quad \textit{and} \quad \operatorname{def}(\varphi'_n) = o_{\mathcal{U}}(\operatorname{def}(\varphi_n)).$$

The converse of Theorem 3.6 is also valid in the following sense.

Proposition 3.7 (De Chiffre, Glebsky, Lubotzky, and Thom [ibid.]). *Let $\Gamma = \langle X | R\rangle$ be a finitely presented group, let $\varphi_n, \psi_n : X \to \mathrm{U}(n)$ be asymptotic representations with respect to some family of submultiplicative, unitarily invariant norms and suppose*

- $\operatorname{dist}(\varphi_n, \psi_n) = O_{\mathcal{U}}(\operatorname{def}(\varphi_n))$ *and*
- $\operatorname{def}(\psi_n) = o_{\mathcal{U}}(\operatorname{def}(\varphi_n))$.

Then, the 2-cocycle α associated with $(\varphi_n)_{n\in\mathbb{N}}$ is trivial in $H^2(\Gamma, \mathrm{M}_{\mathcal{U}})$. In particular, if φ_n is sufficiently close to a homomorphism for n large enough, then α is trivial.

It remains to observe that in case of the Frobenius-norm, the ultraproduct $M_{\mathcal{U}}$ is a Hilbert space and the action of Γ is given by a unitary representation. Together with a somewhat subtle minimality argument this proves Theorem 3.3.

It is now clear that we are in need of large classes of groups for which general vanishing results for the second cohomology with Hilbert space coefficients can be proven. We will discuss some aspects of this problem in the next section.

3.3 Cohomology vanishing and examples of n-Kazhdan groups. Recall that if Γ is a finitely generated group, then Γ has Kazhdan's Property (T) if and only if the first cohomology $H^1(\Gamma, \mathcal{H}_\pi)$ vanishes for every unitary representation $\pi : \Gamma \to \mathrm{U}(\mathcal{H}_\pi)$ on a Hilbert space $\mathcal{H}_\pi$, see Bekka, de la Harpe, and Valette [2008] for a proof and more background information. We will consider groups for which the higher cohomology groups vanish. Higher dimensional vanishing phenomena have been studied in various articles, see for example Bader and Nowak [2015], Ballmann and Świątkowski [1997], Dymara and Januszkiewicz [2002], and Oppenheim [2015].

In De Chiffre, Glebsky, Lubotzky, and Thom [2017], we proposed the following terminology.

Definition 3.8. *Let $n \in \mathbb{N}$. A group Γ is called* n-Kazhdan *if $H^n(\Gamma, \mathcal{H}_\pi) = 0$ for all unitary representations $(\pi, \mathcal{H}_\pi)$ of Γ. We call Γ* strongly n-Kazhdan*, if Γ is k-Kazhdan for $k = 1, \dots, n$.*

So 1-Kazhdan is Kazhdan's classical property (T). See Bader and Nowak [2015] and Oppenheim [2015] for discussions of other related higher dimensional analogues of Property (T). Let's discuss briefly one source of n-Kazhdan groups for $n \geq 2$. Let K be a non-archimedean local field of residue class q, i.e. if $\mathcal{O} \subset K$ is the ring of integers and $\mathfrak{m} \subset \mathcal{O}$ is its unique maximal ideal, then $q = |\mathcal{O}/\mathfrak{m}|$. Let $\mathbf{G}$ be a simple K-algebraic group of K-rank r and assume that $r \geq 1$. The group $G := \mathbf{G}(K)$ acts on the associated Bruhat-Tits building $\mathcal{B}$. The latter is an infinite, contractible, pure simplicial complex of dimension r, on which G acts transitively on the chambers, i.e. the top-dimensional simplices. Let Γ be a uniform lattice in G, i.e. a discrete cocompact subgroup of G. When Γ is also torsion free, then the quotient $X := \Gamma \backslash \mathcal{B}$ is a finite r-dimensional simplicial complex and $\Gamma = \pi_1(X)$. In particular, the group Γ is finitely presented. We will use the following theorem which essentially appears in work of Ballmann and Świątkowski [1997] building on previous work of Garland [1973].

Theorem 3.9 (Garland, Ballmann–Świątkowski)**.** *For every natural number $r \geq 2$, there exists $q_0(r) \in \mathbb{N}$ such that the following holds. If $q \geq q_0(r)$ and G and Γ are as above, then Γ is strongly $(r-1)$-Kazhdan. In particular, if $r \geq 3$, then Γ is 2-Kazhdan.*

It is very natural to wonder what happens in the analogous real case. It is worth noting that already $H^5(\mathrm{SL}_n(\mathbb{Z}), \mathbb{R})$ is non-trivial for n large enough Borel [1974]; thus $\mathrm{SL}_n(\mathbb{Z})$ fails to be 5-Kazhdan for n large enough. Similarly, note that $H^2(\mathrm{Sp}(2n, \mathbb{Z}), \mathbb{R}) = \mathbb{R}$ for all $n \geq 2$ Borel [ibid.], so that the natural generalization to higher rank lattices in real Lie groups has to be formulated carefully; maybe just by excluding an explicit list of finite-dimensional unitary representations.

Question 3.10 (De Chiffre, Glebsky, Lubotzky, and Thom [2017])**.** Is it true that $\mathrm{SL}_n(\mathbb{Z})$ is 2-Kazhdan for $n \geq 4$?

It is worthwhile to return to the remark that Theorem 3.6 and the analogue of Theorem 3.3 is valid if one replaces $\|\cdot\|_{\mathrm{Frob}}$ with any submultiplicative norm $\|\cdot\|$ and the assumption that $H^2(\Gamma, V) = \{0\}$ whenever $V = \prod_{n\to\mathfrak{U}}(\mathrm{M}_n(\mathbb{C}), \|\cdot\|)$ equipped with some action of Γ. This, for instance, gives a sufficient condition for stability with respect to the operator norm, but it seems difficult to prove the existence of a group Γ with vanishing second cohomology in this case. The following question seems more approachable.

Question 3.11. Can the above strategy be used to prove stability results w.r.t. to the Schatten-p-norm?

The techniques rely on submultiplicativity of the norm and thus cannot be directly applied to the normalized Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}}$. However, it is worth noting, that since $\frac{1}{\sqrt{k}}\|A\|_{\mathrm{Frob}} = \|A\|_{\mathrm{HS}} \leq \|A\|_{\mathrm{op}} \leq \|A\|_{\mathrm{Frob}}$ for $A \in \mathrm{M}_k(\mathbb{C})$, we get the following immediate corollary to Theorem 3.3.

Corollary 3.12 (De Chiffre, Glebsky, Lubotzky, and Thom [ibid.]). *Let $\Gamma = \langle X | R\rangle$ be a finitely presented 2-Kazhdan group and let $\varphi_n \colon X \to \mathrm{U}(n)$ be a sequence of maps such that*

$$\mathrm{def}(\varphi_n) = o_{\mathfrak{U}}(n^{-1/2}),$$

where the defect is measured with respect to either $\|\cdot\|_{\mathrm{HS}}$ or $\|\cdot\|_{\mathrm{op}}$. Then φ_n is equivalent to a sequence of homomorphisms.

The preceding corollary provides some quantitative information on Connes' Embedding Problem. Indeed, if a finitely presented, non-residually finite, 2-Kazhdan group is Connes-embeddable, then there is some upper bound on the quality of the approximation in terms of the dimension of the unitary group. Needless to say it would be very interesting to decide if groups as above are Connes-embeddable. A positive answer to Question 3.11 for $p > 2$ should lead to improvements in Corollary 3.12.

4 Applications to group theory

4.1 The basic setup. For any group Γ, an element w in the free product $\Gamma * \mathbf{F}_k$ determines a word map $w\colon \Gamma^{\times n} \to \Gamma$ given by evaluation. Let us denote by $\varepsilon\colon \Gamma * \mathbf{F}_k \to \mathbf{F}_k$ the natural augmentation which sends Γ to the neutral element and call $\varepsilon(w)$ the *content* of w. We call $w \in \Gamma * \mathbf{F}_k$ a group word in k variables with coefficients in Γ. Every group word $w \in \Gamma * \mathbf{F}_k$ determines an equation $w(X) = 1$ in k variables with coefficients in Γ in an obvious way. We say that $w \in \Gamma * \mathbf{F}_k$ can be solved *over* Γ if there exists an overgroup $\Lambda \supseteq \Gamma$ and $g_1, \ldots, g_k \in \Lambda$ such that $w(g_1, \ldots, g_k) = 1$, where 1 denotes the neutral element in Λ. Similarly, we say that it can be solved *in* Γ if we can take $\Lambda = \Gamma$. It is clear that an equation $w \in \Gamma * \mathbf{F}_k$ can be solved over Γ if and only if the natural homomorphism $\Gamma \to \Gamma * \mathbf{F}_k/\langle\langle w \rangle\rangle$ is injective. Similarly, an equation can be solved in Γ if and only if the natural homomorphism $\Gamma \to \Gamma * \mathbf{F}_k/\langle\langle w \rangle\rangle$ is split-injective, i.e., it has a left inverse.

The study of equations over groups dates back to the work of Neumann [1943]. There is an extensive literature about equations over groups, including Gersten [1987], Gerstenhaber and Rothaus [1962], Howie [1981], Klyachko [1993], Levin [1962], Neumann [1943], Klyachko and Thom [2017], and Roman'kov [2012]. In this section, we plan to survey some observations and results that were obtained in joint work with Klyachko and Thom [2017].

Let us start with an observation. It is well-known that not all equations with coefficients in Γ are solvable over Γ. For example if $\Gamma = \langle a, b | a^2, b^3 \rangle$, then the equation $w(x) = xax^{-1}b$ with variable x is not solvable over Γ. Indeed, a and b cannot become conjugate in any overgroup of Γ. Another example involving only one kind of torsion is $\Gamma = \mathbb{Z}/p\mathbb{Z} = \langle a \rangle$ with the equation $w(x) = xax^{-1}axa^{-1}x^{-1}a^{-2}$. However, in both cases we have $\varepsilon(w) = 1 \in \mathbf{F}_k$. Indeed, the only known examples of equations which are not solvable over some Γ are equations whose content is trivial. We call an equation $w \in \Gamma * \mathbf{F}_k$ singular if its content is trivial, and non-singular otherwise. This leads to the following question:

Question 4.1 (Klyachko and Thom [ibid.])**.** Let Γ be a group and $w \in \Gamma * \mathbf{F}_k$ be an equation in n variables with coefficients in Γ. If w is non-singular, is it true, that it is solvable over Γ? In addition, if Γ is finite, can we find a solution in a finite extension?

The case $k = 1$ is the famous Kervaire-Laudenbach Conjecture. The one-variable case was studied in work by Gerstenhaber–Rothaus, see Gerstenhaber and Rothaus [1962]. They showed that if Γ is finite, then every non-singular equation in one variable can be solved in a finite extension of Γ. Their proof used computations in cohomology of the compact Lie groups $\mathrm{U}(n)$. Their strategy was to use homotopy theory to say that the associated word map $w \colon \mathrm{U}(n) \to \mathrm{U}(n)$ has a non-vanishing degree (as a map of oriented manifolds) and thus must be surjective. Any preimage of the neutral element provides a solution to the equation w. The key to the computation of the degree is to observe that the degree depends only on the homotopy class of w and thus – since $\mathrm{U}(n)$ is connected – does not change if w is replaced by $\varepsilon(w)$. The computation of the degree is now an easy consequence of classical computations of Hopf [1940]. We conclude that any non-singular equation in one variable with coefficients in $\mathrm{U}(n)$ can be solved $\mathrm{U}(n)$ – thus $\mathrm{U}(n)$ deserves to be called *algebraically closed*.

The property of being algebraically closed is easily seen to pass to arbitrary Cartesian products of groups and arbitrary quotients of groups. As a consequence, non-singular equations in one variable with coefficients in Γ as above can be solved over Γ if Γ is isomorphic to a subgroup of a quotient of the infinite product $\prod_n \mathrm{U}(n)$ – an observation that is due to Pestov [2008]:

Theorem 4.2 (Pestov [ibid.]). *The Kervaire-Laudenbach Conjecture holds for Connes-embeddable groups.*

Note that this covers all amenable groups, or more generally, all sofic groups Pestov [ibid.]. As we have discussed, the Connes' Embedding Conjecture predicts (among other things) that every countable group is Connes-embeddable and thus implies the Kervaire-Laudenbach Conjecture – this was also observed in Pestov [ibid.].

Actually, Gerstenhaber and Rothaus [1962] studied the more involved question whether m equations of the form $w_1, \ldots, w_m \in \Gamma * \mathbf{F}_k$ in k variables can be solved simultaneously over Γ. Their main result is that this is the case if Γ is finite (or more generally, locally residually finite) and the presentation two-complex

$$X := K\langle X | \varepsilon(w_1), \ldots, \varepsilon(w_m)\rangle$$

satisfies $H_2(X, \mathbb{Z}) = 0$, i.e., the second homology of X with integral coefficients vanishes.

Howie [1981] proved the same result for locally indicable groups and conjectured it to hold for all groups – we call that Howie's Conjecture. Again, Connes' Embedding Conjecture implies Howie's Conjecture – and more specifically, every Connes-embeddable group satisfies Howie's Conjecture.

4.2 Topological methods to prove existence of solutions. The main goal of Klyachko and Thom [2017] was to provide examples of singular equations in many variables which are solvable over every Connes-embeddable group, where the condition on the equation *only* depends on its content. Indeed, we gave a positive answer to Question 4.1 when $k = 2$ in particular cases. This should be compared for example with results of Gersten [1987], where the conditions on w depended on the unreduced word obtained by deleting the coefficients from w.

Theorem 4.3 (Klyachko and Thom [2017]). *Let Γ be a Connes-embeddable group. An equation in two variables with coefficients in Γ can be solved over Γ if its content does not lie in $[\mathbf{F}_2, [\mathbf{F}_2, \mathbf{F}_2]]$. Moreover, if Γ is finite, then a solution can be found in a finite extension of Γ.*

In order to prove our main result we had to refine the study of Gerstenhaber–Rothaus on the effect of word maps on cohomology of compact Lie groups. Again, the strategy is to show that such equations can be solved in $\mathrm{SU}(n)$ for sufficiently many $n \in \mathbb{N}$. More specifically, we proved:

Theorem 4.4 (Klyachko and Thom [2017]). *Let p be a prime number and let $w \in \mathrm{SU}(p) * \mathbf{F}_2$ be a group word. If*

$$\varepsilon(w) \notin [\mathbf{F}_2, \mathbf{F}_2]^p[\mathbf{F}_2, [\mathbf{F}_2, \mathbf{F}_2]],$$

then the equation $w(a,b) = 1$ can be solved in $\mathrm{SU}(p)$.

If $\varepsilon(w) \notin [\mathbf{F}_2, \mathbf{F}_2]$, then this theorem is a direct consequence of the work of Gerstenhaber–Rothaus. However, if $\varepsilon(w) \in [\mathbf{F}_2, \mathbf{F}_2]$, then a new idea is needed. We showed – under the conditions on p which are mentioned above – that the induced word map $w\colon \mathrm{PU}(p) \times \mathrm{PU}(p) \to \mathrm{SU}(p)$ is surjective, where $\mathrm{SU}(p)$ denotes the special unitary group and $\mathrm{PU}(p)$ its quotient by the center. Again, the strategy was to replace w by the much simpler and homotopic map induced by $\varepsilon(w)$ and study its effect on cohomology directly.

The proof is a *tour de force* in computing the effect of $\varepsilon(w)\colon \mathrm{PU}(p) \times \mathrm{PU}(p) \to \mathrm{SU}(p)$ in cohomology with coefficients in $\mathbb{Z}/p\mathbb{Z}$. Using fundamental results of Serre, Bott, and Baum-Browder on the p-local homotopy type of spheres, lens spaces and projective unitary groups and finally computations of Kishimoto and Kono [2009], we managed to show that the image of the top-dimensional cohomology class of $\mathrm{SU}(p)$ is non-trivial. This implies that no map that is homotopic to $\varepsilon(w)$ can be non-surjective. In particular, we can conclude as in the arguments of Gerstenhaber–Rothaus that $w\colon \mathrm{PU}(p) \times \mathrm{PU}(p) \to \mathrm{SU}(p)$ must be surjective.

In general, the assumption on $\varepsilon(w)$ cannot be omitted in the previous theorem. Indeed, in previous work the following result (independently obtained by Elon Lindenstrauss) was shown.

Theorem 4.5 (Thom [2013]). *For every $k \in \mathbb{N}$ and every $\varepsilon > 0$, there exists $w \in \mathbf{F}_2 \setminus \{e\}$, such that*

$$\|w(a,b) - 1_n\| < \varepsilon, \quad \forall a, b \in \mathrm{SU}(k).$$

In particular, the equation $w(a,b) = g$ is not solvable when $\|g - 1_n\| \geq \varepsilon$.

The construction that proves the preceding theorem yields words in $\mathbf{F}_2$ that lie deep in the derived series, so that there is no contradiction with Theorem 4.4.

The surjectivity of word maps without coefficients is an interesting subject in itself. Larsen conjectured that for each non-trivial $w \in \mathbf{F}_2$ and n high enough, the associated word map $w\colon \mathrm{PU}(n) \times \mathrm{PU}(n) \to \mathrm{PU}(n)$ is surjective. This was shown (with some divisibility restrictions on n) for words not in the second derived subgroup of $\mathbf{F}_2$ by Elkasapy and the author in Elkasapy and Thom [2014].

In a similar direction, we believe that for n high enough – or again, with some divisibility restrictions – the word map w should define a non-trivial homotopy class and be not even homotopic to a non-surjective map. In this context let us mention some questions that appear naturally at the interface between homotopy theory and the study of word maps. Given a topological group G, it is natural to study the group of words modulo those which are null-homotopic. Indeed, we set

$$N_{n,G} := \{w \in \mathbf{F}_k | w \colon G^n \to G \text{ is homotopically trivial}\}$$

and define $\mathcal{H}_{n,G} := \mathbf{F}_k / N_{n,G}$.

Question 4.6. Can we compute $\mathcal{H}_{2,\mathrm{SU}(n)}$?

See James and Thomas [1959] and Yagita [1993] for partial information about $\mathcal{H}_{n,G}$ in special cases. For example, it follows from classical results of Whitehead that $\mathcal{H}_G$ is k-step nilpotent for some $k \leq 2\dim(G)$.

Similarly, we call $w \in \mathbf{F}_k$ *homotopically surjective* with respect to G if every map in the homotopy class of $w \colon G^{\times n} \to G$ is surjective.

Question 4.7. Let $w \in \mathbf{F}_2$ be non-trivial. Is $w \colon \mathrm{PU}(n) \times \mathrm{PU}(n) \to \mathrm{PU}(n)$ homotopically surjective for large n?

In order to study words which lie deeper in the lower central series, we suspected in Klyachko and Thom [2017] that it might be helpful to observe that the induced word map $w \colon \mathrm{PU}(p) \times \mathrm{PU}(p) \to \mathrm{PU}(p)$ does not only lift to $\mathrm{SU}(p)$ – which is the simply connected cover of $\mathrm{PU}(p)$ – but lifts even to higher connected covers of $\mathrm{PU}(p)$. Indeed, for example one can show that if $w \in [\mathbf{F}_2, [\mathbf{F}_2, \mathbf{F}_2]]$, then the associated word map lifts to the complex analogue of the string group.

Acknowledgments. I thank Jakob Schneider for careful reading.

References

Goulnara Arzhantseva and Liviu Păunescu (2017). "Linear sofic groups and algebras". *Transactions of the American Mathematical Society* 369.4, pp. 2285–2310 (cit. on p. 1800).

Uri Bader and Piotr W. Nowak (2015). "Cohomology of deformations". *J. Topol. Anal.* 7.1, pp. 81–104. MR: 3284391 (cit. on p. 1810).

Werner Ballmann and Jacek Świątkowski (1997). "On L^2-cohomology and property (T) for automorphism groups of polyhedral cell complexes". *Geom. Funct. Anal.* 7.4, pp. 615–645. MR: 1465598 (cit. on pp. 1807, 1810).

Bachir Bekka, Pierre de la Harpe, and Alain Valette (2008). *Kazhdan's property (T)*. Vol. 11. New Mathematical Monographs. Cambridge University Press, Cambridge, pp. xiv+472 (cit. on p. 1810).

Bruce Blackadar and Eberhard Kirchberg (1997). "Generalized inductive limits of finite-dimensional C^*-algebras". *Math. Ann.* 307.3, pp. 343–380. MR: 1437044 (cit. on pp. 1805, 1806).

Armand Borel (1974). "Stable real cohomology of arithmetic groups". *Ann. Sci. École Norm. Sup. (4)* 7, 235–272 (1975) (cit. on p. 1810).

Marc Burger, Narutaka Ozawa, and Andreas Thom (2013). "On Ulam stability". *Israel J. Math.* 193.1, pp. 109–129. MR: 3038548 (cit. on p. 1806).

John Carrion, Marius Dadarlat, and Caleb Eckhardt (2013). "On groups with quasidiagonal C^*-algebras". *J. Funct. Anal.* 265.1, pp. 135–152 (cit. on pp. 1805, 1806).

Alain Connes (1976). "Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$". *Ann. of Math. (2)* 104.1, pp. 73–115. MR: 0454659 (cit. on pp. 1798, 1805, 1806).

Yves Cornulier (2011). "A sofic group away from amenable groups". *Mathematische Annalen* 350.2, pp. 269–275 (cit. on p. 1800).

Marcus De Chiffre, Lev Glebsky, Alex Lubotzky, and Andreas Thom (2017). "Stability of asymptotic representations and cohomology vanishing". arXiv: 1711.10238 (cit. on pp. 1800, 1806, 1807, 1809–1811).

Pierre Deligne (1978). "Extensions centrales non résiduellement finies de groupes arithmétiques". French, with English summary. *C. R. Acad. Sci. Paris Sér. A-B* 287.4, A203–A208. MR: 507760 (cit. on p. 1807).

Michal Doucha (n.d.). "Metric topological groups: their metric approximation and metric ultraproducts". To appear in Groups, Geometry, and Dynamics. arXiv: 1601.07449 (cit. on p. 1804).

Jan Dymara and Tadeusz Januszkiewicz (2002). "Cohomology of buildings and their automorphism groups". *Invent. Math.* 150.3, pp. 579–627. MR: 1946553 (cit. on p. 1810).

Gábor Elek and Endre Szabó (2004). "Sofic groups and direct finiteness". *Journal of Algebra* 280.2, pp. 426–434 (cit. on p. 1799).

– (2005). "Hyperlinearity, essentially free actions and L^2-invariants. The sofic property". *Mathematische Annalen* 332.2, pp. 421–441 (cit. on p. 1799).

Abdelrhman Elkasapy and Andreas Thom (2014). "About Gotô's method showing surjectivity of word maps". *Indiana Univ. Math. J.* 63.5, pp. 1553–1565 (cit. on p. 1814).

Howard Garland (1973). "p-adic curvature and the cohomology of discrete subgroups of p-adic groups". *Ann. of Math. (2)* 97, pp. 375–423. MR: 0320180 (cit. on pp. 1807, 1810).

Steve M. Gersten (1987). "Reducible diagrams and equations over groups". In: *Essays in group theory*. Vol. 8. Math. Sci. Res. Inst. Publ. New York: Springer, pp. 15–73 (cit. on pp. 1812, 1813).

Murray Gerstenhaber and Oscar S. Rothaus (1962). "The solution of sets of equations in groups". *Proc. Nat. Acad. Sci. U.S.A.* 48, pp. 1531–1533 (cit. on pp. 1812, 1813).

Lev Glebsky (2016). "Approximations of groups, characterizations of sofic groups, and equations over groups". *Journal of Algebra* (cit. on pp. 1798, 1799).

Lev Glebsky and Luis Manuel Rivera (2008). "Sofic groups and profinite topology on free groups". *Journal of Algebra* 320.9, pp. 3512–3518 (cit. on pp. 1800–1802).

Mikhael Gromov (1999). "Endomorphisms of symbolic algebraic varieties". *Journal of the European Mathematical Society* 1.2, pp. 109–197 (cit. on pp. 1798–1800).

Karsten Grove, Hermann Karcher, and Ernst A. Ruh (1974). "Jacobi fields and Finsler metrics on compact Lie groups with an application to differentiable pinching problems". *Math. Ann.* 211, pp. 7–21 (cit. on p. 1806).

Bernhard Herwig and Daniel Lascar (2000). "Extending partial automorphisms and the profinite topology on free groups". *Trans. Amer. Math. Soc.* 352.5, pp. 1985–2021 (cit. on p. 1802).

Derek F. Holt and Sarah Rees (2016). "Some closure results for $\mathcal{C}$-approximated groups". arXiv: 1601.01836 (cit. on p. 1798).

Heinz Hopf (1940). "Über den Rang geschlossener Liescher Gruppen". *Comment. Math. Helv.* 13, pp. 119–143 (cit. on p. 1812).

James Howie (1981). "On pairs of 2-complexes and systems of equations over groups". *J. Reine Angew. Math.* 324, pp. 165–174 (cit. on pp. 1812, 1813).

– (1984). "The p-adic topology on a free group: A counterexample". *Mathematische Zeitschrift* 187.1, pp. 25–27 (cit. on pp. 1800, 1803).

Ioan James and Emery Thomas (1959). "Which Lie groups are homotopy-abelian?" *Proc. Nat. Acad. Sci. U.S.A.* 45, pp. 737–740 (cit. on p. 1815).

Aditi Kar and Nikolay Nikolov (2014). "A non-LEA sofic group". *Proceedings – Mathematical Sciences*, pp. 1–5 (cit. on p. 1800).

David Kazhdan (1982). "On ε-representations". *Israel J. Math.* 43.4, pp. 315–323 (cit. on pp. 1806, 1807).

Daisuke Kishimoto and Akira Kono (2009). "On a conjecture of Ōshima". *Topology Appl.* 156.13, pp. 2189–2192 (cit. on p. 1814).

Anton Klyachko (1993). "A funny property of sphere and equations over groups". *Comm. Algebra* 21.7, pp. 2555–2575 (cit. on p. 1812).

Anton Klyachko and Andreas Thom (2017). "New topological methods to solve equations over groups". *Algebr. Geom. Topol.* 17.1, pp. 331–353 (cit. on pp. 1799, 1812–1815).

Frank Levin (1962). "Solutions of equations over groups". *Bull. Amer. Math. Soc.* 68, pp. 603–604 (cit. on p. 1812).

Martin W. Liebeck and Aner Shalev (2001). "Diameters of finite simple groups: sharp bounds and applications". *Annals of mathematics*, pp. 383–406 (cit. on pp. 1798, 1803).

Alex Lubotzky and Yehuda Shalom (2004). "Finite representations in the unitary dual and Ramanujan groups". *Discrete geometric analysis: proceedings of the first JAMS*

Symposium on Discrete Geometric Analysis, December 12-20, 2002, Sendai, Japan, Contemporary Mathematics 347, pp. 173 (cit. on p. 1808).

Dermot H. McLain (1954). "A characteristically simple group". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 50. Cambridge University Press, pp. 641–642 (cit. on p. 1803).

Bernhard H. Neumann (1943). "Adjunction of elements to groups". *J. London Math. Soc.* 18, pp. 4–11 (cit. on p. 1812).

Nikolay Nikolov, Jakob Schneider, and Andreas Thom (2017). "Some remarks on finitarily approximated groups". To appear in Journal de l'École polytechnique – Mathématiques. arXiv: 1703.06092 (cit. on pp. 1800–1805).

Nikolay Nikolov and Dan Segal (2012). "Generators and commutators in finite groups; abstract quotients of compact groups". *Invent. Math.* 190.3, pp. 513–602 (cit. on pp. 1801–1804).

Izhar Oppenheim (2015). "Vanishing of cohomology and property (T) for groups acting on weighted simplicial complexes". *Groups Geom. Dyn.* 9.1, pp. 67–101. MR: 3343347 (cit. on p. 1810).

Narutaka Ozawa (2004). "About the QWEP conjecture". *Internat. J. Math.* 15.5, pp. 501–530 (cit. on p. 1806).

Narutaka Ozawa, Mikael Rørdam, and Yasuhiko Sato (2015). "Elementary amenable groups are quasidiagonal". *Geom. Funct. Anal.* 25.1, pp. 307–316. MR: 3320894 (cit. on p. 1807).

Narutaka Ozawa, Andreas Thom, and Marcus De Chiffre (2017). "Operator algebraic approach to inverse and stability theorems for amenable groups". arXiv: 1706.04544 (cit. on p. 1806).

Vladimir G. Pestov (2008). "Hyperlinear and sofic groups: a brief guide". *Bull. Symbolic Logic* 14.4, pp. 449–480 (cit. on pp. 1805, 1806, 1813).

Anand Pillay (2017). "Remarks on compactifications of pseudofinite groups". *Fundamenta Mathematicae* 236.2, pp. 193–200 (cit. on p. 1805).

Vitaliĭ Roman'kov (2012). "Equations over groups". *Groups Complex. Cryptol.* 4.2, pp. 191–239 (cit. on p. 1812).

Jakob Schneider and Andreas Thom (2017). "A note on the normal subgroup lattice of ultraproducts of finite quasisimple groups". arXiv: 1709.06286 (cit. on p. 1801).

Dan Segal (2000). "Closed subgroups of profinite groups". *Proceedings of the London Mathematical Society* 81.1, pp. 29–54 (cit. on p. 1803).

– (2009). *Words: notes on verbal width in groups*. Vol. 361. Cambridge University Press (cit. on pp. 1801–1803).

Abel Stolz and Andreas Thom (2014). "On the lattice of normal subgroups in ultraproducts of compact simple groups". *Proc. Lond. Math. Soc. (3)* 108.1, pp. 73–102 (cit. on p. 1801).

Andreas Thom (2010). “Examples of hyperlinear groups without factorization property”. *Groups Geom. Dyn.* 4.1, pp. 195–208 (cit. on pp. 1800, 1807).
– (2012). “About the metric approximation of Higman’s group”. *J. Group Theory* 15.2, pp. 301–310 (cit. on pp. 1798, 1800).
– (2013). “Convergent sequences in discrete groups”. *Canad. Math. Bull.* 56.2, pp. 424–433 (cit. on p. 1814).
Andreas Thom and John S. Wilson (2014). “Metric ultraproducts of finite simple groups”. *C. R. Math. Acad. Sci. Paris* 352.6, pp. 463–466 (cit. on pp. 1800, 1801).
– (2016). “Some geometric properties of metric ultraproducts of finite simple groups”. arXiv: 1606.03863 (cit. on pp. 1800, 1801).
Aaron Tikuisis, Stuart White, and Wilhelm Winter (2017). “Quasidiagonality of nuclear C^*-algebras”. *Ann. of Math. (2)* 185.1, pp. 229–284 (cit. on p. 1806).
Alan M. Turing (1938). “Finite approximations to Lie groups”. *Annals of Mathematics*, pp. 105–111 (cit. on p. 1805).
Stanisław Ulam (1960). *A collection of mathematical problems*. Vol. 8. Interscience Tracts in Pure and Applied Mathematics. New York-London: Interscience Publishers (cit. on p. 1798).
Benjamin Weiss (2000). “Sofic groups and dynamical systems”. *Sankhyā Ser. A* 62.3. Ergodic theory and harmonic analysis (Mumbai, 1999), pp. 350–359 (cit. on p. 1798).
Nobuaki Yagita (1993). “Homotopy nilpotency for simply connected Lie groups”. *Bull. London Math. Soc.* 25.5, pp. 481–486 (cit. on p. 1815).
Boris Zilber (2014). “Perfect infinities and finite approximation”. *Infinity and Truth. IMS Lecture Notes Series* 25 (cit. on p. 1805).

Received 2017-12-05.

Andreas Thom
Andreas Thom
Institut für Geometrie
TU Dresden
Germany
andreas.thom@tu-dresden.de

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1819–1842)

STRUCTURE OF NUCLEAR C*-ALGEBRAS: FROM QUASIDIAGONALITY TO CLASSIFICATION AND BACK AGAIN

Wilhelm Winter

Abstract

I give an overview of recent developments in the structure and classification theory of separable, simple, nuclear C^*-algebras. I will in particular focus on the role of quasidiagonality and amenability for classification, and on the regularity conjecture and its interplay with internal and external approximation properties.

Introduction

A C^*-algebra is a (complex) Banach-* algebra such that $\|x^*x\| = \|x\|^2$ for all elements x. Equivalently, C^*-algebras may be thought of as norm-closed *-subalgebras of the bounded operators on Hilbert spaces. A von Neumann algebra is one that is even closed with respect to the weak operator topology. Examples of C^*-algebras include continuous functions on compact Hausdorff spaces, section algebras of vector bundles with matrix fibres, or suitable norm completions of group algebras. Group C^*-algebras come in different sizes; in particular there is a full one, which is universal with respect to all unitary representations of the group, and a reduced one, which is the norm completion of the left regular representation. Similar constructions can be associated with topological dynamical systems, via the crossed product construction.

From the 1970s on it became clear that the notion of amenability for groups, with its many equivalent formulations, can be rephrased, in almost as many ways, for operator algebras as well. Some of these notions are more or less directly carried over from groups to group C^*-algebras or group von Neumann algebras. But then it often turns out that they make perfect sense at a much more abstract level — and even there they remain closely related. Highlights are Choi–Effros' and Kirchberg's characterisation of nuclear

Supported by the Deutsche Forschungsgemeinschaft through SFB 878.
MSC2010: primary 46L05; secondary 46L35, 46L80, 43A07.
Keywords: C^*-algebra, amenability, nuclearity, quasidiagonality, classification.

C*-algebras by the completely positive approximation property, and of course Connes' classification of injective II_1 factors.

Connes' theorem kicked off an avalanche of further developments in von Neumann algebras, but it also remained an inspiration for C*-algebras. Elliott quite boldly conjectured that separable nuclear C*-algebras should be classifiable by K-theoretic data. After some refinements and adjustments, and many years of hard work, we now understand the conjecture and its scope fairly well, at least for simple and unital C*-algebras. Moreover, after a long detour, we now know that classification of simple nuclear C*-algebras is not only philosophically, but also technically, surprisingly analogous to the classification of injective factors. In particular, the von Neumann algebraic properties of being (o) injective, (i) hyperfinite, (ii) McDuff, and (iii) having tracial comparability of projections, do have C*-algebraic counterparts: (o) nuclearity, (i) finite (noncommutative) topological dimension, (ii) $\mathcal{Z}$-stability, and (iii) comparison of positive elements. Here, nuclearity is characterised via the completely positive approximation property, topological dimension is either decomposition rank or nuclear dimension, and $\mathcal{Z}$-stability is tensorial absorption of the Jiang–Su algebra $\mathcal{Z}$, the smallest possible C*-algebraic analogue of the hyperfinite II_1 factor $\mathcal{R}$. Comparison of positive elements can be described as a regularity property (lack of perforation) of the Cuntz semigroup.

A major difference between the C*- and the von Neumann algebra situation is that for von Neumann factors injectivity implies the three other properties, whereas a simple C*-algebra may be nuclear but fail to have finite nuclear dimension, be $\mathcal{Z}$-stable, or have strict comparison. However, at least conjecturally these three properties occur or fail simultaneously — and for C*-algebras with, say, not too complicated tracial state spaces, this is indeed a theorem. For now let us state a special case; we give a more comprehensive version later on.

THEOREM A: *For a separable, simple, unital, nuclear* C*-*algebra* $A \neq M_r(\mathbb{C})$ *with at most one tracial state, the following are equivalent:*

(i) *A has finite nuclear dimension.*

(ii) *A is $\mathcal{Z}$-stable, $A \cong A \otimes \mathcal{Z}$.*

(iii) *A has strict comparison of positive elements.*

After the classification of injective factors was complete, it took around fifteen years to finish the C*-analogue of the type III case. (I say 'finish', but this is only correct modulo the UCT problem. We will soon return to this little wrinkle.) So why did the type II analogue take almost forty years? For once, there is more information to keep track of: for type II_1 factors the invariant is simply a point, whereas for C*-algebras the invariant involves all possible ordered K-groups together with arbitrary Choquet simplices (arising as tracial state spaces). There is, however, also a deeper, and more mysterious reason.

This is related to both the universal coefficient theorem (UCT) and to quasidiagonality. The UCT problem asks whether all separable nuclear C*-algebras are—in a very weak sense—homotopy equivalent to commutative ones. Conceptually this has a topological flavour, so from this perspective it is reasonable that the UCT becomes an issue in the C*-algebraic (i.e., topological) setup, and not in the von Neumann algebraic (i.e., measure theoretic) situation. Nonetheless, I would like to understand this explanation at a more technical level — but maybe this is asking for too much as long as the UCT problem is not resolved. Quasidiagonality is an external approximation property; the quasidiagonality question (QDQ) asks whether all stably finite nuclear C*-algebras admit a separating set of finite dimensional approximate representations; cf. Blackadar and Kirchberg [1997]. For von Neumann algebras the situation is more clear, since here it is a 2-norm (i.e., *tracial*) version of quasidiagonality that matters.

The connection between amenability and quasidiagonality was perhaps first drawn in Hadwin [1987], where Rosenberg observed that discrete groups with quasidiagonal reduced group C*-algebras are amenable. The converse statement became known as Rosenberg's conjecture.

In Tikuisis, White, and Winter [2017], QDQ was answered for UCT C*-algebras with faithful tracial states.

THEOREM B: *Let A be a nuclear* C**-algebra with a faithful tracial state. Suppose A satisfies the UCT (one could also say that A is* KK*-equivalent to a commutative* C**-algebra). Then, A is quasidiagonal.*

By work of Tu (and Higson–Kasparov), amenable group C*-algebras do satisfy the UCT; since they also have a canonical faithful trace, this confirms Rosenberg's conjecture, and one arrives at a new characterisation of amenability.

COROLLARY C: *For a discrete group G, the reduced group* C**-algebra is quasidiagonal if and only if G is amenable.*

Upon combining Theorems A and B with the work of Elliott, Gong, Lin, and Niu, we are now in a position to state the most general classification result that can currently be expected in the simple and unital case.[1,2]

THEOREM D: *The class of separable, simple, unital, nuclear, and $\mathcal{Z}$-stable* C**-algebras satisfying the UCT is classified by* K*-theoretic invariants.*

Separability will always be necessary for a classification result of this type, and nuclearity and $\mathcal{Z}$-stability are known to be essential — and so, within its simple and unital scope, the theorem is complete modulo the UCT problem. For the time being this remains a sore point; however, one should note that in applications the C*-algebras of interest very

[1] To be precise, we need more general versions of Theorems A and B here; I'll state these in Sections 5 and 3.

[2] There are impressive results also in the non-unital and even in the non-simple situation, but I won't go into these here.

often come with sufficient additional geometric structure, so that the UCT can be verified directly.

With the benefit of hindsight, one might divide the classification programme into three major challenges. These are linked at many levels—not least by the final result—but *quasidiagonality* showcases these connections particularly beautifully. This is the point of view I will take in these notes. Let us have a quick look at each of these challenges and very briefly sketch how quasidiagonality enters the game; we will see some more details in the main part of the paper.

The first challenge: *Understand nuclearity and the interplay with finite dimensional approximation properties.*

The main step was the completely positive approximation property as established by Choi–Effros and Kirchberg. This has been refined in various ways since then; in particular it has been used to model finite covering dimension in a noncommutative setting. While these are *internal* approximations, quasidiagonality may be regarded as an *external* approximation property. Understanding when quasidiagonality holds is a major task of the theory.

The second challenge: *Understand the* C^**-algebraic regularity properties* (i), (ii), (iii) *above and their interplay.*

This is about the regularity conjecture for nuclear C^*-algebras, now often referred to as Toms–Winter conjecture. I will state the conjecture in its full form, and describe what we know and what we don't know. There are two C^*-algebraic counterparts of hyperfiniteness in this context: finite decomposition rank and finite nuclear dimension. The first occurs only for finite C^*-algebras, the second in greater generality. It was open for some time what the difference between the two notions is, and we now know that (at least for simple C^*-algebras) the dividing line is marked by quasidiagonal traces.

The third challenge: *Implement the actual classification procedure.*

This is technical, and not easy to describe in short. As an illustration I will at least state a stable uniqueness result, Theorem 2.4, which allows one to compare *-homomorphisms up to unitary equivalence. This is particularly important for Elliott's intertwining argument; cf. Elliott [2010]. I will also mention Kirchberg–Phillips classification, Lin's TAF classification, and the recent spectacular work of Gong–Lin–Niu, together with the, indeed quite final, classification theorem that is now in place.

I find the common history of quasidiagonality and classification most intriguing, since the two have met time and time again, and often very unexpectedly. First, Kirchberg used Voiculescu's result on quasidiagonality of suspensions to prove his famous $\mathcal{O}_2$-embedding theorem, which in turn led to Kirchberg–Phillips classification of purely infinite nuclear C^*-algebras. Next, Popa showed how to excise finite dimensional C^*-algebras in quasidiagonal ones with many projections. This was the starting point for Lin's TAF classification. (Later on, quasidiagonality also became crucial for the classification of specific types of examples, in particular for simple quotients of certain group C*-algebras and for crossed products associated to free and minimal $\mathbb{Z}^d$-actions on compact and finite dimensional Hausdorff spaces.) Quasidiagonality (of all traces, to be precise) was also a crucial hypothesis for the classification result of Elliott–Gong–Lin–Niu. Theorem B above marked a surprising turn of events, as it, conversely, invoked a classification result to arrive at quasidiagonality.

The title of this note refers to the chronological development of matters, as outlined above. The main body of the paper is arranged thematically, in order to give a better overview of the individual aspects of the theory.

In Section 1 we recall the notion of nuclearity, and various versions of the completely positive approximation property. Section 2 gives a very brief overview of K-theory, Elliott's invariant, and the role of the UCT, especially for stable uniqueness results. Section 3 is devoted to quasidiagonality, and a rough outline of the main theorem of Tikuisis, White, and Winter [2017]. In Section 4 we revisit Rosenberg's conjecture on the connection between amenability and quasidiagonality. Section 5 summarises what is known and what is not known about the Toms–Winter conjecture for simple, unital, nuclear C*-algebras. Finally, Section 6 highlights the state of the art of Elliott's classification programme.

I am indebted to Hannes Thiel, Aaron Tikuisis, and Stuart White for looking carefully at an earlier manuscript.

1 Internal approximation: nuclearity and exactness

1.1 C*-algebras form a category, with the most natural choice for morphisms being *-homomorphisms. (It follows from spectral theory that these are automatically norm-decreasing, hence continuous.)

Another important class of morphisms consists of completely positive maps (we write c.p., or c.p.c. if they are also contractive). These are linear and *-preserving, and send positive elements to positive elements, even after amplification with matrix algebras. By Stinespring's theorem, every c.p. map can be written as a compression of a *-homomorphism. More precisely, a map $\varphi : A \longrightarrow B$ is completely positive if and only if B embeds into another C*-algebra C and if there are a *-homomorphism $\pi : A \longrightarrow C$ and some $h \in C$ such that

$$\varphi(\,.\,) = h^*\pi(\,.\,)h.$$

1.2 A completely positive map $\varphi : A \longrightarrow B$ is order zero if it preserves orthogonality, i.e., whenever $a_1, a_2 \in A$ satisfy $a_1 a_2 = 0$, then $\varphi(a_1)\varphi(a_2) = 0$. By Winter and Zacharias [2009], c.p. order zero maps are precisely those for which there is a Stinespring dilation such that $h \in C$ is positive and commutes with $\pi(A)$.

As a consequence of this structure theorem, there is a bijective correspondence between c.p.c. order zero maps $A \longrightarrow B$ and *-homomorphisms $\mathrm{C}_0((0,1], A) \longrightarrow B$. Moreover, one can use functional calculus on the commutative C^*-algebra generated by h (which in fact is a quotient of $\mathrm{C}_0((0,1])$), to define a notion of functional calculus for c.p.c. order zero maps; cf. Winter and Zacharias [2009].

1.3 Definition: *A* C^**-algebra* A *has the completely positive approximation property, if the following holds:*

For any finite subset $\mathcal{F} \subset A$ *and any tolerance* $\epsilon > 0$*, there is a diagram*

$$A \xrightarrow{\psi} F \xrightarrow{\varphi} A \tag{1}$$

with F *a finite dimensional* C^**-algebra and completely positive contractive maps* ψ *and* φ*, such that* $\varphi\psi$ *agrees with the identity map up to* ϵ *on* $\mathcal{F}$*, in short*

$$\varphi\psi =_{\mathcal{F},\epsilon} \mathrm{id}_A, \text{ i.e., } \|a - \varphi\psi(a)\| < \epsilon \text{ for all } a \in \mathcal{F}.$$

We have asked the maps ψ and φ to be contractions. One could also ask them to be just bounded. As long as the norm bound is uniform, the resulting definitions will be equivalent.

1.4 A C^*-algebra A is nuclear if, for every other C^*-algebra B, there is only one C^*-norm on the algebraic tensor product $A \odot B$; equivalently, the maximal and minimal tensor products of A and B agree. A is exact if taking the minimal tensor product with another C^*-algebra B is an exact functor for any B. Since the maximal tensor product has this property, nuclear C^*-algebras are automatically exact.

Choi and Effros proved in Choi and Effros [1978] (and Kirchberg in Kirchberg [1977]) that a C^*-algebra is nuclear if and only if it has the completely positive approximation property. By the Choi–Effros lifting theorem, c.p.c. maps from nuclear C^*-algebras into quotient C^*-algebras always admit c.p.c. lifts.

Kirchberg's $\mathcal{O}_2$ embedding theorem says that any separable exact C^*-algebra can be embedded into the Cuntz algebra $\mathcal{O}_2$. Separable nuclear C^*-algebras are precisely those which in addition are the images of conditional expectations on $\mathcal{O}_2$.

1.5 In 2004, Eberhard Kirchberg and I defined a notion of covering dimension for C^*-algebras which is based on 1.3 and uses order zero maps to model disjointness of open sets in a noncommutative situation. This notion, called decomposition rank, was generalised in Winter and Zacharias [2010] by Joachim Zacharias and myself. Here are the precise definitions.

Definition: *A* C^**-algebra* A *has nuclear dimension at most* d*,* $\dim_{\mathrm{nuc}} A \leq d$*, if the following holds:*

For any finite subset $\mathfrak{F} \subset A$ and any tolerance $\epsilon > 0$, there is a diagram

$$A \xrightarrow{\psi} F \xrightarrow{\varphi} A$$

with F a finite dimensional C*-*algebra, ψ completely positive contractive and φ completely positive, such that*

(i) $\|a - \varphi\psi(a)\| < \epsilon$ *for all $a \in \mathfrak{F}$,*

(ii) there is a decomposition $F = F^{(0)} \oplus \ldots \oplus F^{(d)}$ such that each $\varphi^{(i)} := \varphi|_{F^{(i)}}$ is c.p.c. order zero.

If, moreover, the maps φ can be chosen to be contractive as well, then we say A has decomposition rank at most d, $\mathrm{dr}\, A \leq d$.

Both of these concepts generalise covering dimension for locally compact spaces. The values are zero precisely for AF algebras, and they can be computed (or at least bounded from above) for many concrete examples.

Note that for nuclear dimension, the maps φ are sums of $d + 1$ contractions, hence are uniformly bounded. Therefore, both finite nuclear dimension and finite decomposition rank imply nuclearity. The two concepts, as similar as they look, are genuinely different. In particular, unlike decomposition rank, nuclear dimension may be finite also for infinite C*-algebras such as the Toeplitz algebra or the Cuntz algebras. The problem of characterising the difference turned out to be close to the heart of the subject. It was shown in Kirchberg and Winter [2004] and Winter and Zacharias [2010], respectively, that the maps ψ can be arranged to be approximately multiplicative for decomposition rank, and approximately order zero for nuclear dimension. In the former case, this shows that finite decomposition rank implies quasidiagonality; cf. Section 3. We will see that this is close to nailing down the difference between decomposition rank and nuclear dimension precisely.

1.6 The approximations of 1.5 are more rigid than those of 1.3. This means for some purposes they are more useful, but we also know that not all nuclear C*-algebras have finite nuclear dimension. Building on Hirshberg, Kirchberg, and White [2012] (and extending Choi and Effros [1978]), Brown, Carrión, and White [2016] gave a refined version of the completely positive approximation property, which asks the involved maps to be somewhat more rigid. In particular, the 'downwards' maps ψ can be taken to be approximately order zero, and the 'upwards' maps φ to be sums of honest order zero maps. The precise statement is as follows.

Theorem: *A* C*-*algebra A is nuclear if and only if the following holds:*

For any finite subset $\mathfrak{F} \subset A$ and any tolerance $\epsilon > 0$, there is a diagram

$$A \xrightarrow{\psi} F \xrightarrow{\varphi} A$$

with F a finite dimensional C^**-algebra and c.p.c. maps ψ and φ, such that*

(i) $\|a - \varphi\psi(a)\| < \epsilon$ *for all* $a \in \mathfrak{F}$,

(ii) $\|\psi(a)\psi(b)\| < \epsilon$ *whenever* $a, b \in \mathfrak{F}$ *satisfy* $ab = 0$,

(iii) there is a decomposition $F = F^{(0)} \oplus \ldots \oplus F^{(k)}$, *such that the restrictions* $\varphi|_{F^{(i)}}$ *all have order zero, and such that* $\sum_{i=1}^{k} \|\varphi|_{F^{(i)}}\| \leq 1$.

Compared to 1.5, in this statement the number of summands (of which I think as colours) is not uniformly bounded, but unlike in the original completely positive approximation property of 1.3 one still has individual order zero maps. As an extra bonus, one can arrange the norms to add up to one, or, upon normalising, one can think of the maps φ as *convex* combinations of c.p.c. order zero maps. This kind of approximation is a little subtle to write down explicitly, even when A is a commutative C^*-algebra like $\mathrm{C}([0,1])$. (In particular, the number of colours in this setup will typically become very large; this is because in the proof at some point one has to pass from weak* to norm approximations via a convexity argument.)

1.7 One can think of the approximations of 1.5 and 1.6 as *internal* in the following sense:

By 1.2, each order zero map corresponds to a *-homomorphism from the cone over the domain C^*-algebra, which – for each matrix summand – essentially is given by an embedding of an algebra like $\mathrm{C}_0(X \setminus \{0\}) \otimes M_k$. Here, $X \subset [0,1]$ is a compact subset which is just the spectrum of the positive contraction $\varphi(1_{M_k}) \in A$. On the other hand, one may approximate $X \setminus \{0\}$ by the union of at most one half-open interval, finitely many closed intervals, and finitely many points. Now one can use order zero functional calculus to slightly modify $\varphi|_{M_k}$ in such a way that the image sits in an honest subalgebra of A which (after rescaling and relabeling the involved intervals) is isomorphic to

$$(\mathrm{C}_0((0,1]) \otimes M_k) \oplus (\mathrm{C}([0,1]) \otimes M_k) \oplus \ldots \oplus (\mathrm{C}([0,1]) \otimes M_k) \oplus M_k \oplus \ldots \oplus M_k.$$

The overall φ then maps into a (non-direct) sum of such subalgebras; with a little extra effort one can describe the map ψ in terms of associated conditional expectations, and of course one can also keep track of the convex coefficients in 1.6. It is usually more practical to write c.p. approximations like in (1), but I often do find it useful to think of them as genuinely internal.

2 K-theory, the UCT, and stable uniqueness

2.1 K-theory for C^*-algebras is a generalisation of topological K-theory; it is *the* homology theory which is at the same time homotopy invariant, half-exact, and compatible with stabilisation. For a (say unital) C*-algebra A, $\mathrm{K}_0(A)$ may be defined in terms of equivalence classes of projections in matrix algebras over A, or in terms of equivalence classes of finitely generated projective modules over A. (One first arrives at a semigroup, whose Grothendieck group is an ordered abelian group defined to be K_0.) $\mathrm{K}_1(A)$ may be defined in terms of equivalence classes of unitaries, or by taking K_0 of the suspension. Like for topological K-theory, one has Bott periodicity, so $\mathrm{K}_{*+2}(A) \cong \mathrm{K}_*(A)$.

Every tracial state on A naturally induces a state (i.e., a positive real valued character) on the ordered K_0-group.

2.2 For A simple and unital, the Elliott invariant consists of the ordered K_0-group (together with the position of the class of the unit), K_1, and of the trace space together with the pairing with K_0,

$$\mathrm{Ell}(A) := (\mathrm{K}_0(A), \mathrm{K}_0(A)_+, [1_A]_0, \mathrm{K}_1(A), \mathrm{T}(A), \mathrm{r}_A : \mathrm{T}(A) \longrightarrow \mathrm{S}(\mathrm{K}_0(A))).$$

$\mathrm{Ell}(\,.\,)$ is a functor in a natural manner.

We say a class $\mathcal{E}$ of simple unital C*-algebras is classified by the Elliott invariant, if the following holds:

Whenever A, B are in $\mathcal{E}$, and there is an isomorphism between $\mathrm{Ell}(A)$ and $\mathrm{Ell}(B)$, then there is an isomorphism between the algebras lifting the isomorphism of invariants. We will see in Section 6 that this actually happens, and in great generality.

2.3 Kasparov's KK-theory is a bivariant functor from C*-algebras to abelian groups which is contravariant in the first and covariant in the second variable. It has similar abstract properties as K-theory, and we have $\mathrm{K}_*(A) \cong \mathrm{KK}_*(\mathbb{C}, A)$.

Rosenberg and Schochet in Rosenberg and Schochet [1987] studied the sequence

$$(2) \quad 0 \longrightarrow \mathrm{Ext}^1(\mathrm{K}_*(A), \mathrm{K}_{*+1}(B)) \longrightarrow \mathrm{KK}(A, B) \longrightarrow \mathrm{Hom}(\mathrm{K}_*(A), \mathrm{K}_*(B)) \longrightarrow 0.$$

A separable C*-algebra A is said to *satisfy the universal coefficient theorem* (UCT for short), if the sequence (2) is exact for every σ-unital C*-algebra B. It follows from Rosenberg and Schochet [ibid.] (cf. Blackadar [1998, Theorem 23.10.5]), that A satisfies the UCT precisely if it is KK-equivalent to an abelian C*-algebra.

The UCT problem asks whether *all* separable *nuclear* C*-algebras satisfy the UCT. This is perhaps the most important open question about nuclear C*-algebras.

2.4 The sequence (2) allows one to lift homomorphisms between K-groups to KK-elements. The latter are already a little closer to *-homomorphisms between the C*-algebras, but to get there one needs fairly precise control over the extent to which *-homomorphisms are determined by their KK-classes. This is often done by so-called stable uniqueness theorems, as developed in particular by Lin, Dadarlat and Eilers, and others. Let us state here a slightly simplified version of Dadarlat and Eilers [2002, Theorem 4.5].

Theorem: *Let A, B be unital C*-algebras with A separable and nuclear. Let $\iota : A \longrightarrow B$ be a unital *-homomorphism which is totally full, i.e., for every nonzero positive element* *of A, its image under ι generates all of B as an ideal. Let $\phi, \psi : A \longrightarrow B$ be unital *-homomorphisms such that* $\mathrm{KK}(\phi) = \mathrm{KK}(\psi)$.

Then, for every finite subset $\mathcal{G} \subset A$ and every $\delta > 0$ there are $n \in \mathbb{N}$ and a unitary $u \in M_{n+1}(B)$ such that

$$\|u(\phi(a) \oplus \iota^{\oplus n}(a))u^* - (\psi(a) \oplus \iota^{\oplus n}(a))\| < \delta \textit{ for all } a \in \mathcal{G}.$$

2.5 In the theorem above, the number n depends on $\mathcal{G}$ and on δ, but also on the maps ϕ, ψ, and ι. However, in applications one often cannot specify these maps beforehand. Dadarlat and Eilers in Dadarlat and Eilers [2002] have found a way to deal with this issue (their original result only covers simple domains, but it can be pushed to the non-simple situation as well; cf. Lin [2005, Lemma 5.9] or Tikuisis, White, and Winter [2017, Theorem 3.5]). The idea is it to assume that n cannot be chosen independently of the maps, and then to construct sequences of maps which exhibit this behaviour. Now regard these sequences as product maps, and apply the original Theorem 2.4 to arrive at a contradiction. To this end, it is important to keep control over the KK-classes of the product maps — which is not easy, since KK-theory is not compatible with products in general. At this point the UCT saves the day, since (at least for the algebras involved) it guarantees that the map

$$\mathrm{KK}\big(A, \textstyle\prod_{\mathbb{N}} B_n\big) \longrightarrow \textstyle\prod_{\mathbb{N}} \mathrm{KK}(A, B_n)$$

is injective. Very roughly, if two sequences of KK-elements on the right hand side agree, they are connected by a sequence of homotopies. But since there is no uniform control over the length of these, it is not clear how to combine them to a single homotopy on the left hand side, at least not for general A. On the other hand, one can do this if the domain algebra A has some additional geometric structure — e.g., if it is commutative. But then of course it also suffices if A is KK-equivalent to a commutative C^*-algebra, i.e., if it satisfies the UCT.

3 External approximation: quasidiagonality

3.1 Halmos defined a set $S \subset \mathcal{B}(\mathcal{H})$ of operators on a Hilbert space to be quasidiagonal if there is an increasing net of finite rank projections converging strongly to the identity operator on the Hilbert space, such that the projections approximately commute with elements of S. One then calls a C^*-algebra quasidiagonal if it has a faithful representation on some Hilbert space, such that the image forms a quasidiagonal set of operators in Halmos' sense. Voiculescu in Voiculescu [1991, Theorem 1] rephrased this in a way highlighting quasidiagonality as an external approximation property.

Theorem: *A C^*-algebra A is quasidiagonal if, for every finite subset $\mathcal{F}$ of A and $\epsilon > 0$, there are a matrix algebra M_k and a c.p.c. map $\psi : A \longrightarrow M_k$ such that* *(i)* $\|\psi(ab) - \psi(a)\psi(b)\| < \epsilon$ *for all* $a, b \in \mathfrak{F}$*,*

(ii) $\|\psi(a)\| > \|a\| - \epsilon$ *for all* $a \in \mathfrak{F}$*.*

3.2 The maps of the theorem above may be thought of as approximate finite dimensional representations. This point of view has still not been fully exploited, partly because C*-algebras are not so accessible to representation theoretic methods. It has, on the other hand, turned out to be fruitful to think of quasidiagonality as an embeddability property. Let $\mathcal{Q}$ denote the universal UHF algebra, $\mathcal{Q} = M_2 \otimes M_3 \otimes \ldots$, so that each matrix algebra embeds unitally into $\mathcal{Q}$. Then, a separable C*-algebra A is quasidiagonal if and only if there is a commuting diagram of the form

$$\begin{tikzcd} & \prod_{\mathbb{N}} \mathcal{Q} \arrow[d] \\ A \arrow[ur, "\tilde{\psi}"] \arrow[r, "\bar{\psi}"] & \prod_{\mathbb{N}} \mathcal{Q} / \bigoplus_{\mathbb{N}} \mathcal{Q} \end{tikzcd}$$

with $\bar{\psi}$ an injective *-homomorphism and $\tilde{\psi}$ a completely positive contraction. If, in addition, A is nuclear, then the lift $\tilde{\psi}$ always exists by the Choi–Effros lifting theorem. Moreover, one may replace the sequence algebra $\prod_{\mathbb{N}} \mathcal{Q} / \bigoplus_{\mathbb{N}} \mathcal{Q}$ by an ultrapower $\mathcal{Q}_\omega$. As a result, a separable nuclear C*-algebra A is quasidiagonal if and only if there is an embedding

$$\kappa : A \longrightarrow \mathcal{Q}_\omega. \tag{3}$$

3.3 Every quasidiagonal C*-algebra is stably finite, i.e., neither the algebra nor any of its matrix amplifications contains a projections which is Murray–von Neumann equivalent to a proper subprojection (this is a finiteness condition, reminiscent of the absence of paradoxical decompositions). The quasidiagonality question (QDQ) asks whether this is the only obstruction, at least in the nuclear case.

QUESTION: (QDQ) Is every stably finite nuclear C*-algebra quasidiagonal?

After being around for some time this was first put in writing by Blackadar and Kirchberg in Blackadar and Kirchberg [1997]. There is a range of variations as discussed in Winter [2016].

3.4 By Ozawa [2013], like $\mathcal{Q}$ itself, the ultrapower $\mathcal{Q}_\omega$ also has a unique tracial state $\tau_{\mathcal{Q}_\omega}$. The composition $\tau_{\mathcal{Q}_\omega} \circ \kappa$ is a positive tracial functional on A. Whenever this is nonzero one may rearrange both κ and its lift so that $\tau_{\mathcal{Q}_\omega} \circ \kappa$ is a tracial state on A. A tracial state

τ which arises in this manner is called a quasidiagonal trace:

$$\begin{tikzcd} A \arrow[r, "\exists\kappa"] \arrow[dr, "\tau"'] & \mathcal{Q}_\omega \arrow[d, "\tau_{\mathcal{Q}_\omega}"] \\ & \mathbb{C} \end{tikzcd}$$

It is common to drop the injectivity requirement on κ in (3) in this context; this is largely for notational convenience since otherwise one would often have to factorise through the quotient by the trace kernel ideal.

A natural refinement of 3.3 is QDQ for traces; cf. Brown and Ozawa [2008] and Winter [2016]. Just as QDQ, this has been around for a while; to the best of my knowledge it appeared in Nate Brown's Brown [2006] for the first time explicitly. It became a quite crucial topic for Bosa, Brown, Sato, Tikuisis, White, and Winter [2015] and Tikuisis, White, and Winter [2017], and also for Elliott's classification programme, as we will see below.

QUESTION: Is every tracial state on a nuclear C^*-algebra quasidiagonal?

3.5 The fact that *unital* quasidiagonal C^*-algebras always have at least one quasidiagonal trace was first observed by Voiculescu in Voiculescu [1993]. On the other hand, an arbitrary embedding $\kappa : A \longrightarrow \mathcal{Q}_\omega$ may well end up in the trace kernel ideal of $\tau_{\mathcal{Q}_\omega}$, so that the composition $\tau_{\mathcal{Q}_\omega} \circ \kappa$ vanishes. For embeddings of cones as in Voiculescu [1991] this will always be the case. In that paper Voiculescu showed that quasidiagonality is homotopy invariant and concluded that cones and suspensions of arbitrary (say separable) C^*-algebras are always quasidiagonal. The method is completely general, but it does not allow to keep track of tracial states.

For cones over *nuclear* C^*-algebras, one can say more: Tikuisis, White, and Winter [2017, Lemma 2.6] introduced a way of mapping a cone over the nuclear C^*-algebra A to $\mathcal{Q}_\omega$ while at the same time controlling a prescribed trace on A. More precisely:

LEMMA: *Let A be a separable nuclear C^*-algebra with a tracial state τ_A. Then,*

(i) there is a c.p.c. order zero map

$$\varphi : A \longrightarrow \mathcal{Q}_\omega$$

such that $\tau_A = \tau_{\mathcal{Q}_\omega} \circ \varphi$, and

(ii) there is a *-homomorphism

$$\acute{\Lambda} : \mathrm{C}_0((0,1]) \otimes A \longrightarrow \mathcal{Q}_\omega$$

such that $\tau_\lambda \otimes \tau_A = \tau_{\mathcal{Q}_\omega} \circ \acute{\Lambda}$, where τ_λ denotes the Lebesgue integral on $\mathrm{C}_0((0,1])$.

Let us have a quick glance at the proof. Since $\mathcal{Q}$ is tensorially self-absorbing and since it is not very hard to find an embedding λ of the interval into $\mathcal{Q}$ in a Lebesgue trace preserving way, (ii) follows from (i) by extending the c.p.c. order zero map

$$A \xrightarrow{\lambda(\mathrm{id}_{(0,1]})\otimes\varphi} \mathcal{Q} \otimes \mathcal{Q}_\omega \longrightarrow (\mathcal{Q} \otimes \mathcal{Q})_\omega \cong \mathcal{Q}_\omega$$

to a *-homomorphism defined on the cone,

$$\acute{\Lambda} : \mathrm{C}_0((0,1]) \otimes A \longrightarrow \mathcal{Q}_\omega.$$

For (i), for the sake of simplicity let us assume that τ_A is extremal. Then since A is nuclear, by Connes' work Connes [1976], the weak closure of A in the GNS representation π_{τ_A} for τ_A is the hyperfinite II_1 factor $\mathcal{R}$. It follows from the Kaplansky density theorem that there is a surjection from $\mathcal{Q}_\omega$ onto $\mathcal{R}_\omega$. Now again by nuclearity, the Choi–Effros lifting theorem yields a c.p.c. lift $\widetilde{\varphi}$ of π_{τ_A}:

$$\begin{tikzcd} & & \mathcal{Q}_\omega \arrow[d, "q"] \\ A \arrow[urr, "\widetilde{\varphi}"] \arrow[r, "\pi_{\tau_A}"'] & \mathcal{R} \arrow[r, hook] & \mathcal{R}_\omega \end{tikzcd}$$

This lift $\widetilde{\varphi}$ has no reason to be order zero. However, for any approximate unit $(e_\lambda)_\Lambda$ of the kernel of the quotient map q, the maps $\widetilde{\varphi}_\lambda := (1-e_\lambda)^{1/2}\widetilde{\varphi}(\,.\,)(1-e_\lambda)^{1/2}$ will lift π_{τ_A} as well — and if one takes the approximate unit to be quasicentral with respect to Q_ω those maps are at least *approximately* order zero. Now use separability of A and a 'diagonal sequence argument' to turn the $\widetilde{\varphi}_\lambda$ into an honest order zero lift φ. This type of diagonal sequence argument appears inevitably when working with sequence algebras. In this case one can run it more or less by hand, but a better, and more versatile way to implement it in a C*-algebra context is Kirchberg's ϵ-test; cf. Kirchberg [2006, Lemma A1].

3.6 Let us take another look at the lemma above when A is unital. In this case, we have an embedding

$$\acute{\lambda} := \acute{\Lambda}|_{\mathrm{C}_0((0,1])\otimes 1_A} : \mathrm{C}_0((0,1]) \longrightarrow \mathcal{Q}_\omega$$

of the cone into $\mathcal{Q}_\omega$. One may unitise this map to arrive at an embedding

$$\bar{\lambda} : \mathrm{C}([0,1]) \longrightarrow \mathcal{Q}_\omega$$

which still induces the Lebesgue integral when composed with $\tau_{\mathcal{Q}_\omega}$. If only we could extend this map $\bar{\lambda}$ to $\mathrm{C}([0,1]) \otimes A$, then this would immediately prove quasidiagonality of the trace τ_A. Of course such an extension seems far too much to ask for, but it is not completely unreasonable either: The map $\bar{\lambda}$ restricts to the embeddings

$$\acute{\lambda} : \mathrm{C}_0((0,1]) \longrightarrow \mathcal{Q}_\omega \;\text{ and }\; \grave{\lambda} : \mathrm{C}_0([0,1)) \longrightarrow \mathcal{Q}_\omega.$$

Now since the Lebesgue integral is symmetric under flipping the interval, we see that it agrees with both maps $\tau_{\mathcal{Q}_\omega} \circ \acute{\lambda}$ and $\tau_{\mathcal{Q}_\omega} \circ \grave{\lambda} \circ \mathrm{flip}$ (where flip denotes the canonical isomorphism between $\mathrm{C}_0((0,1])$ and $\mathrm{C}_0([0,1))$. Moreover, by Ciuperca and Elliott [2008] this is enough to make the maps $\acute{\lambda} \circ \mathrm{flip}$ and $\grave{\lambda}$ approximately unitarily equivalent — and again by a diagonal sequence argument one can even make them honestly unitarily equivalent, i.e., one can find a unitary $u \in \mathcal{Q}_\omega$ such that

$$\grave{\lambda}(\,.\,) = u\,\acute{\lambda}(\mathrm{flip}(\,.\,))\,u^*. \tag{4}$$

Now this map can clearly be extended to all of $\mathrm{C}_0([0,1)) \otimes A$ by setting

$$\grave{\Lambda} := u\,\acute{\Lambda}((\mathrm{flip} \otimes \mathrm{id}_A)(\,.\,))\,u^*.$$

At this point we have two maps

$$\acute{\Lambda} : \mathrm{C}_0((0,1]) \otimes A \longrightarrow \mathcal{Q}_\omega \;\text{ and }\; \grave{\Lambda} : \mathrm{C}_0([0,1)) \otimes A \longrightarrow \mathcal{Q}_\omega,$$

which we would like to ‘superpose’ to a map defined on $\mathrm{C}([0,1]) \otimes A$. This can be done by means of a 2×2 matrix trick, involving the unitary u and rotation ‘along the interval’. The result will be a c.p.c. map

$$\bar{\Lambda} : \mathrm{C}([0,1]) \otimes A \longrightarrow M_2(\mathcal{Q}_\omega)$$

which will map $1_{[0,1]}$ to a projection of trace $1/2$. However, to arrive at quasidiagonality, $\bar{\Lambda}$ would also have to be multiplicative. This will indeed happen provided one can in addition choose the unitary u to implement the flip on all of the suspension $\mathrm{C}_0((0,1)) \otimes A$, or equivalently, to satisfy (4) as well as

$$\grave{\Lambda}(f \otimes a) = u\,\acute{\Lambda}(f \otimes a)\,u^* = \acute{\Lambda}(f \otimes a) \tag{5}$$

for all $a \in A$ and for all $f \in \mathrm{C}_0((0,1))$ which are symmetric in the sense that $\mathrm{flip}(f) = f$. In other words, we have to implement the flip on $\mathrm{C}_0((0,1))$ in the relative commutant of a certain suspension over A.

It is a lot to ask for such a unitary to begin with, and even approximate versions are just as hard to achieve, since in an ultrapower approximately implementing (4) and (5) will be as good as implementing them exactly. On the other hand, the domains of our maps are cones, or suspensions embedded in cones, hence zero-homotopic, and so there is no obstruction in K-theory to finding such a u. Luckily, there are powerful techniques from C*-algebra classification in place which do allow to compare maps when they sufficiently agree on K-theory; cf. Lin [2002] and Dadarlat and Eilers [2002], or 2.4 above. These require the target algebra to be 'admissible' (which is the case here), but there is also a price to pay: with these stable uniqueness theorems, one can only compare maps up to (approximate) unitary equivalence after adding a 'large' map to both sides. This largeness can be measured numerically using the trace of the target algebra (which recovers the trace τ_A on A via Λ; at this point it is important that τ_A is faithful). In general the largeness constant depends on the algebras involved, but also on the maps. In Tikuisis, White, and Winter [2017] we found a way to use large multiples of the original maps as correcting summands. It is then important for the largeness constant to not depend on the maps involved. As pointed out in 2.5, Dadarlat and Eilers have indeed developed such a stable uniqueness theorem which works when the domain (in our case the suspension $\mathrm{C}_0((0,1)) \otimes A$) in addition satisfies the UCT.

This is all made precise in Tikuisis, White, and Winter [ibid.], which also contains an extensive sketch of the proof (a slightly more informal sketch can be found in Winter [2016]). Here is the result.

3.7 Theorem: *Every faithful trace on a nuclear* C*-*algebra satisfying the UCT is quasidiagonal.*

In particular this answers the quasidiagonality question QDQ for UCT C*-algebras with faithful traces. We will see some more consequences in the subsequent sections.

3.8 In 2017, Gabe generalised the theorem above to the situation where A is only assumed to be exact (but still satisfying the UCT), and the trace is amenable.

In 2017, Schafhauser gave a different, and shorter, proof, which replaces the stable uniqueness theorem of Dadarlat and Eilers [2002] by a result from Elliott and Kucerovsky [2001].

4 Rosenberg's conjecture: amenability

4.1 In the appendix of Hadwin [1987], Rosenberg observed that for reduced group C*-algebras amenability and quasidiagonality are closely related.

Proposition: *Let G be a countable discrete group and suppose the reduced group* C*-*algebra $\mathrm{C}^*_\mathrm{r}(G) \subset \mathcal{B}(\ell^2(G))$ is quasidiagonal. Then, G is amenable.*

A proof is not hard, and worth looking at (the one below can be extracted from Brown and Ozawa [2008, Corollary 7.1.17, via Theorem 6.2.7]): For $g \in G$ let u_g denote the image of g in $\mathcal{B}(\ell^2(G))$ under the left regular representation. Let $(p_n)_\mathbb{N} \subset \mathcal{B}(\ell^2(G))$ be a sequence of finite rank projections strongly converging to the identity and approximately commuting with $\mathrm{C}^*_{\mathrm{r}}(G)$. Compression with the p_n yields unital c.p. maps

$$\varphi_n : \mathrm{C}^*_{\mathrm{r}}(G) \longrightarrow p_n \mathcal{B}(\ell^2(G)) p_n \cong M_{r_n}$$

(where r_n is just the rank of p_n). Since the p_n approximately commute with elements of $\mathrm{C}^*_{\mathrm{r}}(G)$, these maps are approximately multiplicative, so that the limit map

$$\varphi_\infty : \mathrm{C}^*_{\mathrm{r}}(G) \longrightarrow \prod_{\mathbb{N}} M_{r_n} / \bigoplus_{\mathbb{N}} M_{r_n}$$

is a *-homomorphism. Note that each φ_n extends to a unital c.p. map

$$\bar{\varphi}_n : \mathcal{B}(\ell^2(G)) \longrightarrow M_{r_n}$$

by Arveson's extension theorem. Now by Stinespring's theorem, $\mathrm{C}^*_{\mathrm{r}}(G)$ sits in the multiplicative domain of the limit map

$$\bar{\varphi}_\infty : \mathcal{B}(\ell^2(G)) \longrightarrow \prod_{\mathbb{N}} M_{r_n} / \bigoplus_{\mathbb{N}} M_{r_n},$$

which in particular means that for every $g \in G$ and every $x \in \mathcal{B}(\ell^2(G))$,

$$\bar{\varphi}_\infty(u_g x) = \bar{\varphi}_\infty(u_g)\bar{\varphi}_\infty(x).$$

Upon choosing a free ultrafilter $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$, the canonical tracial states on the M_{r_n}, evaluated along ω, yield a tracial state τ_ω on $\prod_{\mathbb{N}} M_{r_n} / \bigoplus_{\mathbb{N}} M_{r_n}$. Now for $x \in \mathcal{B}(\ell^2(G))$ we have

$$\begin{aligned}
\tau_\omega \circ \bar{\varphi}_\infty(u_g x u_g^*) &= \tau_\omega(\bar{\varphi}_\infty(u_g)\bar{\varphi}_\infty(x)\bar{\varphi}_\infty(u_g^*)) \\
&= \tau_\omega(\bar{\varphi}_\infty(u_g^*)\bar{\varphi}_\infty(u_g)\bar{\varphi}_\infty(x)) \\
&= \tau_\omega(\bar{\varphi}_\infty(u_g^* u_g)\bar{\varphi}_\infty(x)) \\
&= \tau_\omega \circ \bar{\varphi}_\infty(x).
\end{aligned}$$

This in particular holds for $x \in \ell^\infty(G)$ (regarded as multiplication operator), and we see that $\tau_\omega \circ \bar{\varphi}_\infty$ is a translation invariant state on $\ell^\infty(G)$. The existence of such an invariant mean is equivalent to G being amenable.

4.2 In the argument above, the p_n approximately commute with elements of $\mathrm{C}^*_{\mathrm{r}}(G)$ *in norm*. However, the construction almost forgets about this and really only requires the p_n to approximately commute with $\mathrm{C}^*_{\mathrm{r}}(G)$ *in trace*. The point is that $\tau_\omega \circ \bar{\varphi}_\infty$ is an *amenable trace*, which is enough to show amenability of G; cf. Brown and Ozawa [2008, Proposition 6.3.3]. If one conversely starts with an amenable group G with a sequence of Følner sets F_n, and chooses $p_n \in \mathcal{B}(\ell^2(G))$ to be the associated finite rank projections, then the same construction as above will again yield an invariant mean. Of course the p_n will approximately commute with $\mathrm{C}^*_{\mathrm{r}}(G)$ only in trace, and not in norm.

On the other hand, for $G = \mathbb{Z}$, one has $\mathrm{C}(S^1) \cong \mathrm{C}^*_{\mathrm{r}}(G) \subset \mathfrak{B}(\ell^2(G))$, and it is well known that commutative C*-algebras are quasidiagonal, not just in the abstract sense, but also when they are concretely represented on a Hilbert space. In this situation, one can even construct the quasi-diagonalising projections fairly explicity, from Følner sets, i.e., one can modify such Følner projections to make them approximately commute with $\mathrm{C}^*_{\mathrm{r}}(G)$ even in norm. (The idea is to 'connect' the left and right hand sides of Følner sets inside the matrix algebra hereditarily generated by the projection.)

4.3 The question of when (and how) one can find projections quasi-diagonalising $\mathrm{C}^*_{\mathrm{r}}(G)$ has turned out to be a hard one. Despite having ever so little evidence at hand at the time, Rosenberg did conjecture that amenable discrete groups are always quasidiagonal. He did not put the conjecture in writing in Hadwin [1987], but did promote the problem subsequently; see Brown [2006] and Brown and Ozawa [2008] for a more detailed discussion.

The conjecture received attention by a number of researchers, and was indeed confirmed for larger and larger classes of amenable groups. These arguments often start with the abelian case and use some kind of bootstrap argument to reach more general classes of groups. The problem usually is that quasidiagonality does not pass to extensions.

4.4 In 2015, Ozawa, Rørdam and Sato proved Rosenberg's conjecture for elementary amenable groups. The latter form a bootstrap class, containing many but not all amenable groups (Grigorchuk's examples with exponential growth are amenable but not elementary amenable — but their group C*-algebras were already known to be quasidiagonal for other reasons). The argument of Ozawa, Rørdam, and Sato [2015] relies on methods and results from the classification programme for simple nuclear C*-algebras. Therefore, already Ozawa, Rørdam, and Sato [ibid.] factorises through a stable uniqueness result like 2.4.

4.5 Eventually, Rosenberg's conjecture was confirmed in full generality as a consequence of the main result from Tikuisis, White, and Winter [2017].

Corollary: *If G is a discrete amenable group, then $\mathrm{C}^*_{\mathrm{r}}(G)$ is quasidiagonal.*

First, it is well known (and not too hard to show) that the canonical trace on $\mathrm{C}^*_{\mathrm{r}}(G)$ is faithful. Then one consults Tu's work Tu [1999] to conclude that amenable group C*-algebras always satisfy the UCT. Theorem 3.7 now says that $\mathrm{C}^*_{\mathrm{r}}(G)$ has a faithful quasidiagonal representation, i.e., it is quasidiagonal as an abstract C*-algebra. But even in its concrete representation on $\ell^2(G)$ it is quasidiagonal; cf. Brown and Ozawa [2008, Theorem 7.2.5].[3]

4.6 The Corollary above does indeed settle Rosenberg's conjecture, but of course only in a very abstract manner. In particular, at this point there seems no way to exhibit quasi-diagonalising projections explicitly, starting, say, with a Følner system for G.

[3] Note that countability / separability is not an issue since all properties involved can be tested locally.

5 Toms–Winter regularity

5.1 Kirchberg's $\mathcal{O}_\infty$-absorption theorem says that a separable, simple, nuclear C*-algebra is purely infinite precisely if it absorbs the Cuntz algebra $\mathcal{O}_\infty$, $A \cong A \otimes \mathcal{O}_\infty$; see Kirchberg [1995]. Next to his $\mathcal{O}_2$-embedding theorem, this is one of the cornerstones for Kirchberg–Phillips classification; cf. Rørdam [2002] for an overview.

At the time it was not at all clear whether one should expect a similar statement for stably finite C*-algebras. We now know that the Jiang–Su algebra $\mathcal{Z}$ of Jiang and Su [1999] really is the right analogue of $\mathcal{O}_\infty$ in this context. Moreover, we know that pure infiniteness can be interpreted as a regularity property of the Cuntz semigroup (almost unperforation, to be more specific) in the absence of traces; cf. Rørdam [2006]. On the other hand, the state of Elliott's classification programme in the early 2000s suggested that dimension type conditions should also play a crucial role.

5.2 In 2009, Andrew Toms and I exhibited a class of inductive limit C*-algebras for which finite decomposition rank, $\mathcal{Z}$-stability, and almost unperforation of the Cuntz semigroup occur or fail simultaneously. This class (Villadsen algebras of the first type) was somewhat artificial, and a bit thin, but still large enough to prompt our conjecture that these three conditions should be equivalent for separable, simple, unital, nuclear and stably finite C*-algebras. Once nuclear dimension was invented and tested, it became soon clear that the conjecture should be generalised to comprise both nuclear dimension and decomposition rank. The full version reads as follows.

Conjecture: *For a separable, simple, unital, nuclear* C*-*algebra $A \neq M_r$ the following are equivalent:*

(i) A has finite nuclear dimension.

(ii) A is $\mathcal{Z}$-stable.

(iii) A has strict comparison of positive elements.

Under the additional assumption that A is stably finite, one may replace (i) *by*

(i') A has finite decomposition rank.

5.3 I stated the above as a conjecture (perhaps in part for sentimental reasons), but after hard work by many people it is now almost a theorem (i.e., most of the implications have been proven in full generality). Let us recap what's known at this point.

When A has no trace, then (ii) $\Longleftrightarrow$ (iii) follows from Kirchberg's $\mathcal{O}_\infty$-absorption theorem as soon as one knows that an infinite exact C*-algebra is $\mathcal{Z}$-stable if and only if it is $\mathcal{O}_\infty$-stable; see Rørdam [2004].

In the finite case, (ii) $\Longrightarrow$ (iii) was shown by Rørdam in Rørdam [ibid.].

I showed first (i') $\Longrightarrow$ (ii) and then (i) $\Longrightarrow$ (ii) in Winter [2010] and Winter [2012], respectively.

In 2012, Matui and Sato showed (iii) $\Longrightarrow$ (ii) when A has only one tracial state. In each of Sato [2012], Kirchberg and Rørdam [2014], and Toms, White, and Winter [2015] this was generalised to the case where the tracial state space of A (always a Choquet simplex) has compact and finite dimensional extreme boundary.

In 2014, Matui and Sato showed (ii) $\Longrightarrow$ (i') when A has only finitely many extremal tracial states, and under the additional assumption that A is quasidiagonal. In Sato, White, and Winter [2015], (ii) $\Longrightarrow$ (i) was implemented in the case of a unique tracial state. In the six author paper Bosa, Brown, Sato, Tikuisis, White, and Winter [2015], Joan Bosa, Nate Brown, Yasuhiko Sato, Aaron Tikuisis, Stuart White and myself showed (ii) $\Longrightarrow$ (i) when the tracial state space of A has compact extreme boundary; (ii) $\Longrightarrow$ (i') was shown assuming in addition that every trace is quasidiagonal (by Theorem 3.7 this is automatic when A satisfies the UCT). For these last results, the ground was prepared by Ozawa's theory of von Neumann bundles from Ozawa [2013]. In upcoming work, Jorge Castillejos, Sam Evington, Aaron Tikuisis, Stuart White and I will show (ii) $\Longrightarrow$ (i) in full generality, and (ii) $\Longrightarrow$ (i') provided that every trace is quasidiagonal.

5.4 With all these results in place now, to sum up it is shorter to state what's not yet known: All we are missing is (iii) $\Longrightarrow$ (ii) for arbitrary trace spaces, and (i) $\Longrightarrow$ (i') without any quasidiagonality assumption.

I am still quite optimistic about the latter statement. For the former one, I also remain positive, but every once in a while I'm tempted to travel back in time to replace condition (iii) by

(iii') A has strict comparison and has almost divisible Cuntz semigroup.

(This condition is equivalent to saying that the Cuntz semigroups of A and $A \otimes \mathcal{Z}$ agree.) On the other hand, Thiel has recently shown that almost divisibility follows from strict comparison in the stable rank one case; see Thiel [2017]. Without this assumption, I wouldn't be too surprised if lack of divisibility was a new source of high dimensional examples in the spirit of Toms [2008].

5.5 Ever since its appearance, Connes' classification of injective II_1 factors was an inspiration for the classification and structure theory of simple nuclear C*-algebras. Once Conjecture 5.2 was formulated, it did not take that long to realise the surprising analogy with Connes' work, in particular Connes [1976, Theorem 5.1]. Roughly speaking, nuclearity on the C*-algebra side corresponds to injectivity for von Neumann algebras, finite nuclear dimension to hyperfiniteness, $\mathcal{Z}$-stability to $\mathfrak{R}$-absorption, i.e., being McDuff (cf. McDuff [1970]), and strict comparison corresponds to comparison of projections.

Matui and Sato in Matui and Sato [2012] and Matui and Sato [2014] have taken this analogy to another level, by turning it into actual theorems. This trend was further pursued in Sato, White, and Winter [2015] and Bosa, Brown, Sato, Tikuisis, White, and Winter [2015]. I find these extremely convincing; by now I am even optimistic that eventually we will be able to view Connes' result and Conjecture 5.2 as incarnations of the same abstract theorem.

6 Elliott's programme: classification

6.1 The first general classification result for nuclear C*-algebras was probably Glimm's classification of UHF algebras in terms of supernatural numbers. Bratteli observed that one can do essentially the same for AF algebras using Bratteli diagrams, but it was Elliott who classified AF algebras in terms of their ordered K_0-groups; Elliott [1976]. More classification results for larger classes were picked up in the 1980s and early 1990s; these prompted Elliott to conjecture that separable, simple, nuclear C*-algebra might be classifiable by K-theoretic invariants. (The precise form of the invariant underwent some adjustments as the theory and understanding of examples progressed.) Up to that point, all available results covered certain types of inductive limit C*-algebras. Then Kirchberg opened the door to classification in a much more abstract context; Kirchberg [1995].
6.2 I have already mentioned that Voiculescu showed quasidiagonality to be a homotopy invariant property. This in particular means that cones over separable C*-algebras are quasidiagonal, because the former are contractible; since quasidiagonality passes to subalgebras, suspensions are quasidiagonal as well. Kirchberg used this statement to prove his celebrated $\mathcal{O}_2$-embedding theorem (cf. Kirchberg [ibid.]; see also Rørdam [2002]), which was a cornerstone for Kirchberg–Phillips classification of separable, nuclear, simple, purely infinite C*-algebras.

This was perhaps the earliest indication that quasidiagonality should be relevant for the classification of nuclear C*-algebras, but of course in this situation we have quasidiagonal cones over C*-algebras which are themselves very far from being quasidiagonal.
6.3 Soon after, however, Popa in Popa [1997] carried over his local quantisation technique from von Neumann factors to simple C*-algebras with traces and sufficiently many projections. The theorem roughly says that such C*-algebras can be approximated locally by finite dimensional C*-subalgebras. It is safe to say that this result kicked off the systematic use of quasidiagonality in the classification of *stably finite* simple nuclear C*-algebras.
6.4 Next, Lin modified Popa's local approximation to the effect that the approximating subalgebras are moreover required to be large in a certain sense; this can often be measured tracially, hence the name TAF (tracially approximately finite dimensional) C*-algebras. Here is the precise definition.

DEFINITION: *A simple, unital* C**-algebra A is TAF, if the following holds: For every finite subset $\mathfrak{F} \subset A$, $\epsilon > 0$, and positive contraction $0 \neq e \in A$, there are a finite dimensional* C**-subalgebra $F \subset A$ and a partial isometry $s \in A$ such that*

(i) $\mathrm{dist}(1_F a 1_F, F) < \epsilon$ *for all $a \in \mathfrak{F}$,*

(ii) $\|1_F a - a 1_F\| < \epsilon$ *for all $a \in \mathfrak{F}$,*

(iii) $s^* s = 1_A - 1_F$ *and* $ss^* \in \overline{eAe}$.

The element e dominates the complement of 1_F, and therefore controls the size of F inside A. In many situations, this can be done in terms of tracial states on A, hence the name *tracially* AF.

6.5 In 2004, Lin managed to classify nuclear TAF algebras satisfying the UCT. The proof is inspired by Kirchberg–Phillips classification.

THEOREM: *The class of all separable, simple, unital, nuclear, infinite dimensional, TAF, UCT* C**-algebras is classified by the Elliott invariant.*

6.6 The theorem covers a fairly large class of stably finite C*-algebras which is characterised abstractly (as opposed to the various classes of inductive limit type algebras handled earlier). The UCT hypothesis remains mysterious, but in applications, for example to transformation group C*-algebras, this is often no issue since one can confirm the UCT directly. The scope of the theorem is nonetheless limited by the TAF assumption, which in particular requires the existence of many projections, and also the ordered K_0-group to be weakly unperforated. Weakly unperforated K-theory is implied by $\mathcal{Z}$-stability of the algebra; in Matui and Sato [2014] it was shown that nuclear TAF algebras are indeed $\mathcal{Z}$-stable. In Winter [2014] it was shown that classification of $\mathcal{Z}$-stable C*-algebras can be derived from classification of UHF-stable ones. This paved the road to applying Lin's TAF classification also in situations when the algebras contain no or only few projections.

In 2015, Gong, Lin and Niu generalised Definition 6.4 by admitting more general building blocks than just finite dimensional C*-algebras. At the same time, they managed to prove a classification result like Theorem 6.5 also in this context. This is a spectacular outcome, since the class covered by the result is no longer subject to K-theoretic restrictions other than those implied by $\mathcal{Z}$-stability anyway.

In Elliott, Gong, Lin, and Niu [2015], finally, it was shown that the UCT, together with finite nuclear dimension and quasidiagonality of all traces suffices to arrive at classification. In conjunction with 5.3 and Theorem 3.7, this confirms Elliott's conjecture in the $\mathcal{Z}$-stable, finite, UCT case. The $\mathcal{Z}$-stable, infinite, UCT case is precisely Kirchberg–Phillips classification. Moreover, Rørdam's and Toms' examples in Rørdam [2003] and in Toms [2008] (inspired by Villadsen [1999]) have shown that $\mathcal{Z}$-stability cannot be disposed of for classification via the Elliott invariant.

The upshot of this discussion is a classification result which—modulo the UCT problem—is as complete and final as can be. It is the culmination of decades of work, by many many hands. It reads as follows.

THEOREM: *The class of all separable, simple, unital, nuclear, $\mathcal{Z}$-stable UCT* C^**-algebras is classified by the Elliott invariant.*

References

Bruce Blackadar (1998). K*-theory for operator algebras*. Second. Vol. 5. Mathematical Sciences Research Institute Publications. Cambridge University Press, Cambridge, pp. xx+300 (cit. on p. 1827).

Bruce Blackadar and Eberhard Kirchberg (1997). "Generalized inductive limits of finite-dimensional C*-algebras". *Math. Ann.* 307.3, pp. 343–380 (cit. on pp. 1821, 1829).

Joan Bosa, Nathanial P. Brown, Yasuhiko Sato, Aaron Tikuisis, Stuart White, and Wilhelm Winter (2015). "Covering dimension of C*-algebras and 2-coloured classification". To appear in *Mem. Amer. Math. Soc.* arXiv: 1506.03974 (cit. on pp. 1830, 1837, 1838).

Nathanial P. Brown (2006). "Invariant means and finite representation theory of C*-algebras". *Mem. Amer. Math. Soc.* 184.865, pp. viii+105 (cit. on pp. 1830, 1835).

– (2011). "The symbiosis of C*- and W*-algebras". In: *Aspects of operator algebras and applications*. Vol. 534. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 121–155.

Nathanial P. Brown, José R. Carrión, and Stuart White (2016). "Decomposable approximations revisited". In: *Operator Algebras and Applications: The Abel Symposium 2015*. Abel Symposia 12. Springer, pp. 45–59 (cit. on p. 1825).

Nathanial P. Brown and Narutaka Ozawa (2008). C*-*algebras and finite-dimensional approximations*. Vol. 88. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (cit. on pp. 1830, 1833–1835).

Man Duen Choi and Edward G. Effros (1978). "Nuclear C*-algebras and the approximation property". *Amer. J. Math.* 100, pp. 61–79 (cit. on pp. 1824, 1825).

Alin Ciuperca and George A. Elliott (2008). "A remark on invariants for C*-algebras of stable rank one". *Int. Math. Res. Not. IMRN* 5, Art. ID rnm 158, 33 pp. (Cit. on p. 1832).

Alain Connes (1976). "Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$". *Ann. of Math. (2)* 104.1, pp. 73–115 (cit. on pp. 1831, 1837).

Joachim Cuntz (1977). "Simple C*-algebras generated by isometries". *Comm. Math. Phys.* 57.2, pp. 173–185.

Marius Dadarlat and Søren Eilers (2002). "On the classification of nuclear C*-algebras". *Proc. London Math. Soc. (3)* 85.1, pp. 168–210 (cit. on pp. 1827, 1828, 1833).

George A. Elliott (1976). “On the classification of inductive limits of sequences of semisimple finite-dimensional algebras”. *J. Algebra* 38.1, pp. 29–44 (cit. on p. 1838).
– (2010). “Towards a theory of classification”. *Adv. Math.* 223.1, pp. 30–48 (cit. on p. 1822).
George A. Elliott and David E. Evans (1993). “The structure of the irrational rotation C*-algebra”. *Ann. of Math. (2)* 138.3, pp. 477–501.
George A. Elliott, Guihua Gong, and Liangqing Li (2007). “On the classification of simple inductive limit C*-algebras. II. The isomorphism theorem”. *Invent. Math.* 168.2, pp. 249–320.
George A. Elliott, Guihua Gong, Huaxin Lin, and Zhuang Niu (2015). “On the classification of simple amenable C*-algebras with finite decomposition rank, II”. arXiv: `1507.03437v2` (cit. on p. 1839).
George A. Elliott and Dan Kucerovsky (2001). “An abstract Voiculescu–Brown–Douglas–Fillmore absorption theorem”. *Pacific J. Math.* 198, pp. 385–409 (cit. on p. 1833).
George A. Elliott and Andrew S. Toms (2008). “Regularity properties in the classification program for separable amenable C*-algebras”. *Bull. Amer. Math. Soc. (N.S.)* 45.2, pp. 229–245.
James Gabe (2017). “Quasidiagonal traces on exact C*-algebras”. *J. Funct. Anal.* 272, pp. 1104–1120 (cit. on p. 1833).
Guihua Gong, Huaxin Lin, and Zhuang Niu (2015). “Classification of finite simple amenable $\mathcal{Z}$-stable C*-algebras”. arXiv: `1501.00135v4` (cit. on p. 1839).
Uffe Haagerup (1985). “A new proof of the equivalence of injectivity and hyperfiniteness for factors on a separable Hilbert space”. *J. Funct. Anal.* 62.2, pp. 160–201.
– (2014). “Quasitraces on exact C*-algebras are traces”. *C. R. Math. Acad. Sci. Soc. R. Can.* 36.2-3, pp. 67–92.
Don Hadwin (1987). “Strongly quasidiagonal C*-algebras”. *J. Operator Theory* 18.1. with an appendix by Jonathan Rosenberg, pp. 3–18 (cit. on pp. 1821, 1833, 1835).
Nigel Higson and Gennadi Kasparov (2001). “E-theory and KK-theory for groups which act properly and isometrically on Hilbert space”. *Invent. Math.* 144.1, pp. 23–74.
Ilan Hirshberg, Eberhard Kirchberg, and Stuart White (2012). “Decomposable approximations of nuclear C*-algebras”. *Adv. Math.* 230, pp. 1029–1039 (cit. on p. 1825).
Xinhui Jiang and Hongbing Su (1999). “On a simple unital projectionless C*-algebra”. *Amer. J. Math.* 121, pp. 359–413 (cit. on p. 1836).
Eberhard Kirchberg (1977). “C*-nuclearity implies CPAP”. *Math. Nachr.* 76, pp. 203–212 (cit. on p. 1824).
– (1995). “Exact C*-algebras, tensor products, and the classification of purely infinite algebras”. In: *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*. Birkhäuser, Basel, pp. 943–954 (cit. on pp. 1836, 1838).

Eberhard Kirchberg (2006). “Central sequences in C*-algebras and strongly purely infinite algebras”. In: *Operator Algebras: The Abel Symposium 2004*. Vol. 1. Abel Symp. Berlin: Springer, pp. 175–231 (cit. on p. 1831).
Eberhard Kirchberg and Mikael Rørdam (2014). “Central sequence C*-algebras and tensorial absorption of the Jiang–Su algebra”. *J. Reine Angew. Math.* 695, pp. 175–214 (cit. on p. 1837).
Eberhard Kirchberg and Wilhelm Winter (2004). “Covering dimension and quasidiagonality”. *Internat. J. Math.* 15.1, pp. 63–85 (cit. on pp. 1824, 1825).
Huaxin Lin (2002). “Stable approximate unitary equivalence of homomorphisms”. *J. Operator Theory* 47.2, pp. 343–378 (cit. on p. 1833).
– (2004). “Classification of simple C*-algebras of tracial topological rank zero”. *Duke Math. J.* 125.1, pp. 91–119 (cit. on p. 1839).
– (2005). “An approximate universal coefficient theorem”. *Trans. Amer. Math. Soc.* 357, pp. 3375–3405 (cit. on p. 1828).
Hiroki Matui and Yasuhiko Sato (2012). “Strict comparison and $\mathcal{Z}$-absorption of nuclear C*-algebras”. *Acta Math.* 209.1, pp. 179–196 (cit. on p. 1837).
– (2014). “Decomposition rank of UHF-absorbing C*-algebras”. *Duke Math. J.* 163.14, pp. 2687–2708 (cit. on pp. 1837, 1839).
Dusa McDuff (1970). “Central sequence and the hyperfinite factor”. *Proc. London Math. Soc. (3)* 21, pp. 443–461 (cit. on p. 1837).
Narutaka Ozawa (2013). “Dixmier approximation and symmetric amenability for C*-algebras”. *J. Math. Sci. Univ. Tokyo* 20, pp. 349–374 (cit. on pp. 1829, 1837).
Narutaka Ozawa, Mikael Rørdam, and Yasuhiko Sato (2015). “Elementary amenable groups are quasidiagonal”. *Geom. Funct. Anal.* 25.1, pp. 307–316 (cit. on p. 1835).
Sorin Popa (1986). “A short proof of “Injectivity Implies Hyperfiniteness” for finite von Neumann algebras”. *J. Operator Theory* 16, pp. 261–272.
– (1997). “On local finite-dimensional approximation of C*-algebras”. *Pacific J. Math.* 181.1, pp. 141–158 (cit. on p. 1838).
Mikael Rørdam (2002). “Classification of nuclear, simple C*-algebras”. In: *Classification of nuclear* C*-*algebras. Entropy in operator algebras*. Vol. 126. Encyclopaedia Math. Sci. Berlin: Springer, pp. 1–145 (cit. on pp. 1836, 1838).
– (2003). “A simple C*-algebra with a finite and an infinite projection”. *Acta Math.* 191, pp. 109–142 (cit. on p. 1839).
– (2004). “The stable and the real rank of $\mathcal{Z}$-absorbing C*-algebras”. *Internat. J. Math.* 15.10, pp. 1065–1084 (cit. on p. 1836).
Mikael Rørdam (2006). “Structure and classification of C*-algebras”. In: *International Congress of Mathematicians. Vol. I*. Eur. Math. Soc., Zürich, pp. 1581–1598 (cit. on p. 1836).

Jonathan Rosenberg and Claude Schochet (1987). “The Künneth theorem and the universal coefficient theorem for Kasparov’s generalized K-functor”. *Duke Math. J.* 55.2, pp. 431–474 (cit. on p. 1827).

Yasuhiko Sato (2012). “Trace spaces of simple nuclear C*-algebras with finite-dimensional extreme boundary”. arXiv: 1209.3000v1 (cit. on p. 1837).

Yasuhiko Sato, Stuart White, and Wilhelm Winter (2015). “Nuclear dimension and $\mathcal{Z}$-stability”. *Invent. Math.* 202, pp. 893–921 (cit. on pp. 1837, 1838).

Christopher Schafhauser (2017). “A new proof of the Tikuisis–White–Winter Theorem”. arXiv: 1701.02180 (cit. on p. 1833).

Hannes Thiel (2017). “Ranks of operators in simple C*-algebras with stable rank one”. arXiv: 1711.04721 (cit. on p. 1837).

Aaron Tikuisis, Stuart White, and Wilhelm Winter (2017). “Quasidiagonality of nuclear C*-algebras”. *Ann. of Math. (2)* 185, pp. 229–284 (cit. on pp. 1821, 1823, 1828, 1830, 1833, 1835).

Andrew S. Toms (2008). “On the classification problem for nuclear C*-algebras”. *Ann. of Math. (2)* 167, pp. 1029–1044 (cit. on pp. 1837, 1839).

Andrew S. Toms, Stuart White, and Wilhelm Winter (2015). “$\mathcal{Z}$-stability and finite-dimensional tracial boundaries”. *Int. Math. Res. Not. IMRN* 10, pp. 2702–2727 (cit. on p. 1837).

Andrew S. Toms and Wilhelm Winter (2009). “The Elliott conjecture for Villadsen algebras of the first type”. *J. Funct. Anal.* 256.5, pp. 1311–1340 (cit. on p. 1836).

Jean-Louis Tu (1999). “La conjecture de Baum–Connes pour les feuilletages moyennables”. K-*Theory* 17.3, pp. 215–264 (cit. on p. 1835).

Jesper Villadsen (1999). “On the stable rank of simple C*-algebras”. *J. Amer. Math. Soc.* 12.4, pp. 1091–1102 (cit. on p. 1839).

Dan Voiculescu (1991). “A note on quasi-diagonal C*-algebras and homotopy”. *Duke Math. J.* 62.2, pp. 267–271 (cit. on pp. 1828, 1830).

– (1993). “Around quasidiagonal operators”. *Integral Equations Operator Theory* 17.1, pp. 137–149 (cit. on p. 1830).

Wilhelm Winter (2010). “Decomposition rank and $\mathcal{Z}$-stability”. *Invent. Math.* 179.2, pp. 229–301 (cit. on p. 1837).

– (2012). “Nuclear dimension and $\mathcal{Z}$-stability of pure C*-algebras”. *Invent. Math.* 187.2, pp. 259–342 (cit. on p. 1837).

– (2014). “Localizing the Elliott conjecture at strongly self-absorbing C*-algebras”. *J. Reine Angew. Math.* 692, pp. 193–231 (cit. on p. 1839).

– (2016). “QDQ vs. UCT”. In: *Operator Algebras and Applications: The Abel Symposium 2015*. Vol. 12. Abel Symposia. Springer, pp. 321–342 (cit. on pp. 1829, 1830, 1833).

Wilhelm Winter and Joachim Zacharias (2009). “Completely positive maps of order zero”. *Münster J. Math.* 2, pp. 311–324 (cit. on pp. 1823, 1824).

Wilhelm Winter and Joachim Zacharias (2010). "The nuclear dimension of C*-algebras". *Adv. Math.* 224.2, pp. 461–498 (cit. on pp. 1824, 1825).

Received 2017-12-02.

Wilhelm Winter
Mathematisches Institut
WWU Münster
Germany
wwinter@uni-muenster.de

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1843–1864)

ERGODIC OPTIMIZATION OF BIRKHOFF AVERAGES AND LYAPUNOV EXPONENTS

JAIRO BOCHI

Abstract

We discuss optimization of Birkhoff averages of real or vectorial functions and of Lyapunov exponents of linear cocycles, emphasizing whenever possible the similarities between the commutative and non-commutative settings.

Introduction

In this paper (X, T) denotes a topological dynamical system, that is, X is a compact metric space and $T: X \to X$ is a continuous map. Often we will impose additional conditions, but broadly speaking the dynamics that interest us the most are those that are sufficiently "chaotic", and in particular have many invariant probability measures.

Our subject is *ergodic optimization* in a broad sense, meaning the study of extremal values of asymptotic dynamical quantities, and of the orbits or invariant measures that attain them. More concretely, we will discuss the following topics:

1. maximization or minimization of the ergodic averages of a real-valued function;
2. optimization of the ergodic averages of a vectorial function, meaning that we are interested in the extrema of the ergodic averages of a function taking values in some euclidian space $\mathbb{R}^d$;
3. maximization or minimization of the top Lyapunov exponent of a linear cocycle over (X, T), or more generally, of the asymptotic average of a subadditive sequence of functions;
4. optimization of the whole vector of Lyapunov exponents of a linear cocycle.

Partially supported by projects Fondecyt 1180371 and Conicyt PIA ACT172001.
MSC2010: primary 37A99; secondary 37D30, 37E45, 37J50, 37H15, 90C05.

Unsurprisingly, many basic results and natural questions that arise in these topics are parallel. The aim of this paper is to provide an unified point of view, hoping that it will attract more attention to the many open problems and potential applications of the subject. The setting should also be convenient for the study of problems where one cares about classes of invariant measures that are not necessarily optimizing.

Disclaimer: This mandatorily short article is not a survey. We will not try to catalog the large corpus of papers that fit into the subject of ergodic optimization. We will neither provide a historical perspective of the development of these ideas, nor explore connections with fields such as Lagrangian Mechanics, Thermodynamical Formalism, Multifractal Analysis, and Control Theory.[1]

1 Optimization of Birkhoff averages

We denote by $\mathfrak{M}_T$ the set of T-invariant probability measures, which is a nonempty convex set and is compact with respect to the weak-star topology. Also let $\mathcal{E}_T \subseteq \mathfrak{M}_T$ be the subset formed by ergodic measures, which are exactly the extremal points of $\mathfrak{M}_T$.

Let $f : X \to \mathbb{R}$ be a continuous function. We use the following notation for Birkhoff sums:

$$f^{(n)} := f + f \circ T + \cdots + f \circ T^{n-1} .$$

By Birkhoff theorem, for every $\mu \in \mathfrak{M}_T$ and μ-almost every $x \in X$, the asymptotic average $\lim_{n\to\infty} \frac{1}{n} f^{(n)}(x)$ is well-defined. The infimum and the supremum of all those averages will be denoted by $\alpha(f)$ and $\beta(f)$, respectively; we call these numbers the *minimal* and *maximal ergodic averages* of f. Since $\alpha(f) = -\beta(-f)$, let us focus the discussion on the quantity β. It can also be characterized as:

$$\beta(f) = \sup_{\mu \in \mathfrak{M}_T} \int f \, \mathrm{d}\mu . \tag{1-1}$$

Compactness of $\mathfrak{M}_T$ implies the following *attainability property*: there exists at least one measure $\mu \in \mathfrak{M}_T$ for which $\int f \, \mathrm{d}\mu = \beta(f)$; such measures will be called *maximizing measures*.

Another characterization is given by the following *enveloping property*:

$$\beta(f) = \inf_{n \geq 1} \frac{1}{n} \sup_{x \in X} f^{(n)}(x) , \quad \text{and the inf is also a lim.} \tag{1-2}$$

Actually, upper semicontinuity of f suffices for these characterizations: see Jenkinson [2006a].

[1] I recommend Jenkinson's new survey Jenkinson [2017], which appeared shortly after the conclusion of this paper.

Recall that a function of the form $h \circ T - h$ with h continuous is called a *coboundary* (or C^0-*coboundary*). Two functions that differ by a coboundary are called *cohomologous*, and have the same maximal ergodic average β. Actually, $\beta(f)$ can be characterized as a minimax over the cohomology class of f:

$$\beta(f) = \inf_{h \in C^0(X)} \sup_{x \in X} (f + h \circ T - h)\,. \tag{1-3}$$

Following the terminology in linear programming, this is called the *dual characterization* of $\beta(f)$: see Radu [2004]. Formula (1-3) was discovered independently several times; it is Lemma 1.3 in the paper Furstenberg and Kifer [1983], where it is proved using Hahn–Banach theorem. Let us reproduce the more direct proof from Conze and Guivarc'h [1993]. The finite-time average $\frac{1}{n} f^{(n)}$ is cohomologous to f; indeed it equals $f + h \circ T - h$ for $h := \frac{1}{n} \sum_{i=1}^{n} f^{(i)}$. Using the enveloping property (1-2) we obtain the $\geq$ inequality in the dual formula (1-3). The reverse inequality is trivial.

Given f, a natural question arises: is the infimum in (1-3) attained? Well, if $h \in C^0(X)$ attains the infimum, then the inequality $f + h \circ T - h \leq \beta(f)$ holds everywhere on X. Consider the closed set K where equality holds. This set is nonempty; indeed by integrating the inequality we see that a measure $\mu \in \mathfrak{M}_T$ is maximizing if and only if $\mu(K) = 1$, that is, if $\operatorname{supp} \mu \subseteq K$. Any closed set with this property is called a *maximizing set*. So another question is whether the existence of such sets is guaranteed. Let us postpone answers a bit.

Every continuous function f can be seen as a linear functional on the vector space of signed Borel measures on X, and conversely. The quantity $\beta(f)$ is the maximum that this linear functional attains in the compact convex set $\mathfrak{M}_T$, and so its computation is a problem of infinite-dimensional linear programming.

Since every ergodic μ is an extremal point of $\mathfrak{M}_T$, it is maximizing for some f. Furthermore, since $\mathfrak{M}_T$ is a simplex (by uniqueness of ergodic decompositions), every ergodic μ is the unique maximizing measure for at least one function $f \in C^0(X)$ (see Jenkinson [2006b] for the precise arguments). Conversely, if f has a unique maximizing measure μ then μ must be ergodic (because the ergodic decomposition of a maximizing measure is formed by maximizing measures).

Uniqueness of the maximizing measure is a *(topologically) generic* property, i.e., it holds for every function in a dense G_δ subset of $C^0(X)$. Furthermore, the same is true if $C^0(X)$ is replaced by any Baire vector space $\mathfrak{F}$ of functions that embeds continuously and densely in $C^0(X)$: see Jenkinson [2006a, Thrm. 3.2].

The properties of maximizing measures of generic functions in $\mathfrak{F}$ may be very different according to the space under consideration. Consider $\mathfrak{F} = C^0(X)$ first. Suppose that T is sufficiently hyperbolic (more precisely, T satisfies Bowen's specification property); to avoid trivialities also assume that X is a perfect set. Then, as shown by Morris [2010], the unique maximizing measure of a generic function $f \in C^0(X)$ satisfies any chosen generic property in the space of measures $\mathfrak{M}_T(X)$; in particular maximizing measures generically have zero entropy and full support. Note that if a function f admits a maximizing set K and has a maximizing measure of full support, then necessarily $K = X$ and therefore all probability measures have the same integral. Since the latter property is obviously non-generic (our T's are not uniquely ergodic), we conclude that generic continuous functions f admit *no* maximizing set, and in particular the infimum in the dual formula (1-3) is *not* attained.

The situation is radically different for more regular functions. A central result in ergodic optimization, found in different forms via various methods by many authors, roughly states that if the dynamics T is sufficiently hyperbolic and the function f is sufficiently regular, then the infimum in the dual formula (1-3) is attained. Such results are called *non-positive Livsic theorems*, *Mañé–Conze–Guivarch lemmas*, or *Mañé lemmas* for short. One of the simplest versions is this: if $T : X \to X$ is a one-sided subshift of finite type, and f is a θ-Hölder function (assuming X metrized in the usual way), then there exists a θ-Hölder function h such that

$$f + h \circ T - h \le \beta(f), \tag{1-4}$$

Similar statements also hold for uniformly expanding maps, Anosov diffeomorphisms and flows, etc. Some references are Conze and Guivarc'h [1993], Savchenko [1999], Contreras, Lopes, and Thieullen [2001], Bousch [2001], Lopes and Thieullen [2003], Pollicott and Sharp [2004], Bousch [2011], and Garibaldi [2017]. The methods of proof are also diverse: some proofs use Thermodynamical Formalism, some use fixed point theorems, and some use bare hands. A function h solving the cohomological inequality (1-4) is called a *subaction*. As negative result, it is shown in Bousch and Jenkinson [2002] that the regularity of the subaction h is not always as good as f: it may be impossible to find a C^1 subaction h even if T and f are C^ω. The study of subactions forms a subject by itself: see Garibaldi [2017].

So it is natural to focus the study on regular functions f and hyperbolic dynamics T, for which the theory is richer. Yuan and Hunt [1999] showed that only measures μ supported on periodic orbits can have the *locking property*, which means that μ is the unique maximizing measure for some f and also for functions sufficiently close to f. Much more recently, Contreras [2016] settled a main open problem and proved that maximizing measures are generically supported on periodic orbits. More precisely, he proved that if T is a expanding map then a generic Lipschitz function has a unique maximizing measure, which is supported on a periodic orbit and has the locking property.

Contreras' theorem provides some confirmation of the experimental findings of Hunt and Ott published two decades before Hunt and Ott [1996]. They basically conjectured that for typical chaotic systems T and typical smooth functions f, the maximizing measure is supported on a periodic orbit. However, their concept of typicality was a probabilistic one: Hunt and Ott actually conjectured that for typical parameterized families of functions, the Lebesgue measure of the parameters corresponding to maximizing orbits of period p or greater is exponentially small in terms of p. This type of conjecture remains open.

A conceptually clean probabilistic notion of typicality in function spaces was introduced in Hunt, Sauer, and Yorke [1992] (basically rediscovering Christensen [1972]); it is called *prevalence*. See Broer, Hasselblatt, and Takens [2010] for several examples of prevalent properties (not all of them topologically generic) in Dynamical Systems. Bochi and Zhang [2016] have obtained the following result in the direction of Hunt–Ott conjectures: if T is the one-sided shift on two symbols, and $\mathfrak{F}$ is a space of functions with a very strong modulus of regularity, then every f in a prevalent subset of $\mathfrak{F}$ has a unique maximizing measure, which is supported on a periodic orbit and has the locking property. Furthermore, we have obtained a sufficient condition for periodicity in terms of the wavelet coefficients of f. There is experimental evidence that this condition is prevalent not only in the space $\mathfrak{F}$ but also on bigger spaces of Hölder functions, but a proof is still missing.

For full shifts (and for other sufficiently hyperbolic dynamics as well), the set of measures supported on periodic orbits is dense in $\mathfrak{M}_T$. In particular, $\mathfrak{M}_T$ is a *Poulsen simplex*: its set $\mathcal{E}_T$ of extremal points is dense. It seems fanciful to try to form a mental image of such an object, but let us try anyway. There are natural ways (see Bochi and Zhang [ibid.]) to approximate the Poulsen simplex by a nested sequence $R_1 \subset R_2 \subset \cdots \subset \mathfrak{M}_T$ of (finite-dimensional) polyhedra whose vertices are measures supported on periodic orbits. Moreover, each polyhedron R_n is a projection of the next one R_{n+1}, and the whole simplex $\mathfrak{M}_T$ is the inverse limit of the sequence. These polyhedra are not simplices: on the contrary, they have a huge number of vertices and their faces are small. Moreover, the polyhedra are increasingly non-round: the height of R_{n+1} with respect to R_n is small. In particular, each interior point of R_n can be well-approximated by a vertex of some R_m with $m > n$; this resembles the Poulsen property. Furthermore, among the vertices of R_n, only a few of them are "pointy", and the others are "blunt"; these pointy vertices are the measures supported on orbits of low period. If we take at random a linear functional on the finite-dimensional span of R_n, then the vertex of R_n that attains the maximum is probably a pointy vertex. This is a speculative justification for the Hunt–Ott conjectures.

As mentioned before, not every $\mu \in \mathcal{E}_T$ can appear as a unique maximizing measure of a *regular* function. So it is natural to ask exactly which measures can appear once the regularity class is prescribed. Motivated by a class of examples that we will explain in the next section, Jenkinson formulated the following fascinating question Jenkinson [2006a, Probl. 3.12]: if T is the doubling map on the circle and f is an analytic function with a unique maximizing measure μ, can μ have positive entropy?

2 Optimization of vectorial Birkhoff averages

Now consider a continuous vectorial function $f : X \to \mathbb{R}^d$. The *rotation set* of f is the set $R(f)$ of all averages $\int f \, d\mu$ where $\mu \in \mathfrak{M}_T$. This is a compact convex subset of $\mathbb{R}^d$. Furthermore, by ergodic decomposition, it equals the convex hull of the averages of f with respect to ergodic measures; in symbols:

$$R(f) = \operatorname{co}\left\{\int f \, d\mu \; ; \; \mu \in \mathcal{E}_T\right\} . \tag{2-1}$$

If $d = 1$ then $R(f) = [\alpha(f), \beta(f)]$, using the notation of the previous section. The prime type of examples of rotation set, which justifies the terminology, are those when f equals the displacement vector of a map of the d-torus homotopic to the identity. Other examples of rotation sets, where the dynamics is actually a (geodesic) flow, are Schwartzman balls: see Paternain [1999] and Pollicott and Sharp [2004].

The measures $\mu \in \mathfrak{M}_T$ for which $\int f \, d\mu$ is an extremal point of $R(f)$ are called *extremal measures*. Of course, each of these measures is also a maximizing measure for a real-valued function $\langle c, f(\cdot)\rangle$, for some nonzero vector $c \in \mathbb{R}^d$, so we can use tools from one-dimensional ergodic optimization.

Let us describe a very important example that appeared in many of the early results in ergodic optimization Conze and Guivarc'h [1993], Hunt and Ott [1996], Jenkinson [1996, 2000], and Bousch [2000]. Let $T(z) := z^2$ be the doubling map on the unit circle, and let $f : S^1 \to \mathbb{C}$ be the inclusion function. The associated rotation set $R(f) \subset \mathbb{C} = \mathbb{R}^2$ is called the *fish*. Confirming previous observations from other researchers, Bousch [2000] proved that the extremal measures are exactly the Sturmian measures. These measures form a family μ_ρ parametrized by rotation number $\rho \in \mathbb{R}/\mathbb{Z}$. If $\rho = p/q$ is rational then μ_ρ is supported on a periodic orbit of period q, while if ρ is irrational then μ_ρ is supported on an extremely thin Cantor set (of zero Hausdorff dimension, in particular) where T is semiconjugated to an irrational rotation. In particular, Jenkinson's problem stated at the end of Section 2 has a positive answer if f is a trigonometric polynomial of degree 1: Sturmian measures not only have zero entropy (i.e., subexponential complexity), but in fact they have linear complexity.

Let us come back to arbitrary T and f. Analogously to (1-2), we have the following *enveloping property*:

$$R(f) = \bigcap_{n \geq 1} \frac{1}{n} \operatorname{co}\left(f^{(n)}(X)\right) , \tag{2-2}$$

and the intersection is also a limit (in both the set-theoretic and the Hausdorff senses). Using the same trick as in the proof of (1-3), it follows that for every neighborhood U of $R(f)$ there exists g cohomologous to f whose image is contained in U.

The following question arises: when can we find g cohomologous to f taking values in $R(f)$? As we have already learned in the previous section, to hope for this to be true we need at least some hyperbolicity and regularity assumptions. If moreover $d = 1$ then the answer of the question becomes positive: Bousch [2002] showed that whenever T and f satisfy the assumptions of a Mañé Lemma, there exists g cohomologous to f taking values in $R(f)$, i.e. such that $\alpha(f) \leq g \leq \beta(f)$. What about $d \geq 2$? Unfortunately the answer is negative. The following observation was found by Vincent Delecroix and the author:

Proposition 2.1. *Let (X, T) and f be as in the definition of the fish. There exists no g C^0-cohomologous to f taking values in the fish.*

Proof. For a contradiction, suppose that there exists a continuous function $h\colon S^1 \to \mathbb{C}$ such that $g := f + h \circ T - h$ takes values in the fish. For each integer $n \geq 0$, let $z_n := e^{2\pi i/2^n}$. These points form a homoclinic orbit

$$\cdots \mapsto z_2 \mapsto z_1 \mapsto z_0 \hookleftarrow \quad \text{with} \quad \lim z_n = z_0 \,.$$

We claim that $g(z_n) = 1$ for every n. This leads to a contradiction, because on one hand, the series $\sum_{n=1}^{\infty}(f(z_n)-1)$ is absolutely convergent to a non-zero sum (as the imaginary part is obviously positive), and on the other hand, by telescopic summation and continuity of h, the sum should be zero.

In order to prove the claim, note that g must constant equal to $\int f \, d\mu_\rho$ on the support of each Sturmian measure μ_ρ. In particular, the compact set $K := (g \circ T - g)^{-1}(0)$ contains all those supports. Consider the obvious semiconjugacy φ between the one-sided two-shift and the doubling map T, namely the map which associates to an infinite word $w = b_0 b_1 \ldots$ in zeros and ones the complex number $\varphi(w) = e^{2\pi i t}$ where $t = 0.b_1 b_2 \ldots$ in binary. Then $\varphi(0^n 1 0^\infty) = z_n$. On the other hand, for each $k \geq 0$, the periodic infinite word $(0^n 1 0^k)^\infty$ is Sturmian, and it tends to $0^n 1 0^\infty$ as $k \to \infty$. This shows that each z_n is the limit of points that belong to supports of Sturmian measures. In particular, each z_n belongs to K. This means that the value $g(z_n)$ is independent of n. Since $z_0 = 1$ is a fixed point, we conclude that this value is 1, as claimed. □

Following Bousch and Jenkinson [2002], a function f is called a *weak coboundary* if $\int f \, d\mu = 0$ for every $\mu \in \mathfrak{M}_T$, or equivalently if f is a uniform limit of coboundaries. There exist weak coboundaries f that are not coboundaries; indeed this happens whenever T is a non-periodic homeomorphism: see Kocsard [2013]. The paper Bousch and Jenkinson [2002] contains an explicit example of such f in the case T is the doubling map. The reason why their function f is not a coboundary is that the sum over a homoclinic orbit is nonzero. This is exactly the obstruction we used in the proof of Proposition 2.1.

Therefore we pose the problem: Does it exist a function g weakly cohomologous to f taking values in the fish?

Naturally, there are many other questions about rotation sets. We can ask about their shape, either for typical or for all functions with some prescribed regularity. It is shown in Kucherenko and Wolf [2014] that any compact convex subset of $\mathbb{R}^d$ is the rotation set of some continuous function. In the case of the fish, the boundary is not differentiable, and has a dense subset of corners, one at each extremal point corresponding to a measure supported on a periodic orbit; furthermore, and all the curvature of the boundary is concentrated at these corners. This seems to be the typical situation of rotation sets of regular functions. Let us note that the boundary of Schartzmann balls is never differentiable: see Pollicott and Sharp [2004] and references therein.

Another property of the fish is the following: the closure of the union of the supports of the extremal measures has zero topological entropy (actually it has cubic complexity; see Mignosi [1991, Corol. 18]). Is this phenomenon typical?

3 Optimization of the top Lyapunov exponent

We now replace Birkhoff sums by matrix products. That is, given a continuous map $F\colon X \to \mathrm{Mat}(d,\mathbb{R})$ taking values into the space of $d \times d$ real matrices, we consider the products

$$F^{(n)}(x) := F(T^{n-1}x) \cdots F(Tx)F(x) .$$

The triple (X, T, F) is called a *linear cocycle* of dimension d. It induces a skew-product dynamics on $X \times \mathbb{R}^d$ by $(x, v) \mapsto (Tx, F(x)v)$, whose n-th iterate is therefore $(x, v) \mapsto (T^n x, F^{(n)}(x)v)$. More generally, we could replace $X \times \mathbb{R}^d$ by any vector bundle over X and then consider bundle endomorphisms that fiber over $T\colon X \to X$, but for simplicity will refrain from doing so.

As an immediate consequence of the Kingman's subadditive ergodic theorem, for any $\mu \in \mathfrak{M}_T$ and μ-almost every $x \in X$, the following limit, called the *top Lyapunov exponent* at x, exists:

$$\lambda_1(F, x) := \lim_{n\to\infty} \frac{1}{n} \log \|F^{(n)}(x)\| \in [-\infty, +\infty),$$

which is clearly independent of the choice of norm on the space of matrices. Similarly to what we did in Section 1, we can either minimize or maximize this number; the corresponding quantities will be denoted by $\alpha(F)$ and $\beta(F)$. However, this time the maximization and the minimization problems are fundamentally different. While $\beta(F)$ is always attained by at least one measure (which will be called *Lyapunov maximizing*), that is not necessarily the case for $\alpha(F)$. Indeed, absence of Lyapunov minimizing measures may be locally generic Bochi and Morris [2015, Rem. 1.13], and may occur even for "derivative cocycles": see Cao, Luzzatto, and Rios [2006].

More generally, one can replace $\log \|F^{(n)}\|$ by any subadditive sequence of continuous functions, or even upper semicontinuous ones, and optimize the corresponding asymptotic average. One show check that a maximizing measure always exists, and that a enveloping property similar to (1-2) holds. See the appendix of the paper Morris [2013] for the proofs of these and other basic results on subadditive ergodic maximization. From these general results one can derive immediately those of Cao [2003] and Dragičević [2017], for example.

The maximization of the linear escape rate of a cocycle of isometries also fits in the context of subadditive ergodic optmization. Under a nonpositive curvature assumption, this maximal escape rate satisfies a duality formula resembling (1-3): see Bochi and Navas [2015]. For some information on the maximal escape rate in the case of isometries of Gromov-hyperbolic spaces, see Oregón-Reyes [2016].

Returning to linear cocycles, note that if the matrices $F(x)$ are invertible then we can define a skew-product transformation T_F on the compact space $X \times \mathbb{RP}^{d-1}$ by $(x, [v]) \mapsto (Tx, [F(x)v])$. Then $\beta(F)$ can be seen as the maximal ergodic average of the function $f(x, [v]) := \log(\|F(x)v\|/\|v\|)$. In this way, maximization of the top Lyapunov exponent can be reduced to commutative ergodic optimization. Note, however, that the space of T_F-invariant probability measures depends on F in a complicated way.

There is a specific setting where optimization of the top Lyapunov exponent has been studied extensively. An *one-step cocycle* is a linear cocycle (X, T, F) where (X, T) is the full shift (either one- or two-sided) on a finite alphabet, say $\{1, \dots, k\}$, and the matrix $F(x)$ only depends on the zeroth symbol of the sequence x. Therefore an one-step cocycle is completely specified by a k-tuple of matrices $(A_1, \dots, A_k)$. It is possible to consider also compact alphabets, but for simplicity let us stick with finite ones.

The *joint spectral radius* and the *joint spectral subradius* of a tuple of matrices are respectively defined as the numbers $e^{\beta(F)}$ and $e^{\alpha(F)}$, where (X, T, F) is the corresponding one-step cocycle. The joint spectral radius was introduced in 1960 by Rota and Strang [1960], and it became a popular subject in the 1990's as applications to several areas (wavelets, control theory, combinatorics, etc) were found. The joint spectral subradius was introduced later by Gurvits [1995], and has also been the subject of some pure and applied research. See the book Jungers [2009] for more information.

The first examples of one-step cocycles with finite alphabets without a Lyapunov-maximizing measure supported on a periodic orbit were first constructed in dimension $d = 2$ by Bousch and Mairesse [2002], refuting the finiteness conjecture from Lagarias and Wang [1995]. Other constructions appeared later: see Blondel, Theys, and Vladimirov [2003]. Counterexamples to the finiteness conjecture seem to be very non-typical: Maesumi [2008] conjectures that they have zero Lebesgue measure in the space of tuples of matrices, and Morris and Sidorov [2013] exhibit one-parameter families of pairs of matrices where counterexamples form a Cantor set of zero Hausdorff dimension.

The Lyapunov maximizing measures for the one-step cocycles in the examples from Bousch and Mairesse [2002] and Morris and Sidorov [2013] (among others) are Sturmian and so have linear complexity. There are higher-dimensional examples with arbitrary polynomial complexity: see Hare, Morris, and Sidorov [2013]. In all known examples where the Lyapunov maximizing measure is unique, it has subexponential complexity, i.e., zero entropy. So the following question becomes inevitable: is this always the case?

A partial result in that direction was obtained by Bochi and Rams [2016]. We exhibit a large class of 2-dimensional one-step cocycles for which the Lyapunov-maximizing and Lyapunov-minimizing measures have zero entropy; furthermore the class includes counterexamples to the finiteness conjecture. Our sufficient conditions for zero entropy are simple: existence of strictly invariant families of cones satisfying a non-overlapping condition. (In particular, our cocycles admit a dominated splitting; see Section 4 below for the definition.) To prove the result, we identify a certain order structure on Lyapunov-optimal orbits (or more precisely on the Oseledets directions associated to those orbits) that leaves no room for positive entropy.

In the setting considered in Bochi and Rams [ibid.] (or more generally for cocycles that admit a dominated splitting of index 1), Lyapunov-minimizing measures do exist, and moreover the minimal top Lyapunov exponent α is continuous among such cocycles. For one-step cocycles that admit no such splitting, β is still continuous, but α is not: see Bochi and Morris [2015]. In fact, even though discontinuities of α are topologically non-generic, we believe that they form a set of positive Lebesgue measure: Bochi and Morris [ibid., Conj. 7.7].

Let us now come back to general linear cocycles, but let us focus the discussion on the maximal top Lyapunov exponent β. As we have seen in Section 1, Mañé Lemma is a basic tool in 1-dimensional commutative ergodic optimization. Let us describe related a notion in the setting of Lyapunov exponents.

Let (X, T, F) be a linear cocycle of dimension d. A *Finsler norm* is a family $\{\|\cdot\|_x\}$ of norms in $\mathbb{R}^d$ depending continuously on $x \in X$. An *extremal norm* is a Finsler norm such that

$$\|F(x)v\|_{Tx} \le e^{\beta(F)}\|v\|_x \quad \text{for all } x \in X \text{ and } v \in \mathbb{R}^d.$$

In the case $d = 1$, the linear maps $F(x) \colon \mathbb{R} \to \mathbb{R}$ consist of multiplication by scalars $\pm e^{f(x)}$, and the maximal Lyapunov exponent $\beta(F)$ of the cocycle equals the maximal ergodic average $\beta(f)$ of the function f. Moreover, an arbitrary Finsler norm can be written as $\|v\|_x = e^{h(x)}|v|$, and it will be an extremal norm if and only if $f + h \circ T - h \le \beta(f)$, which is the cohomological inequality (1-4). So the relation with Mañé Lemma becomes apparent and we see that existence of an extremal norm is far from automatic even in dimension $d = 1$.

Extremal norms were first constructed by Barabanov [1988a] in the case of one-step cocycles: he showed that under an irreducibility assumption (no common invariant subspace, except for the trivial ones), there exists an extremal norm that is constant (i.e. $\|v\|_x$ is independent of the basepoint x).[2] These extremal norms provide a fundamental tool in the study of the joint spectral radius: see also Wirth [2002] and Jungers [2009].

Beyond one-step cocycles, when can we guarantee the existence of an extremal norm? Bochi and Garibaldi [n.d.] consider the situation where T is a hyperbolic homeomorphism and $F : X \to \mathrm{GL}(d, \mathbb{R})$ is a θ-Hölder continuous map. As it happens often in this context, it is useful to assume *fiber bunching*, which roughly means that the largest rate under which the matrices $F(x)$ distort angles is bounded by τ^θ, where $\tau > 1$ is a constant related to the hyperbolicity of T. (Note that locally constant cocycles, being locally constant, are θ-Hölder for arbitrarily large θ, and so always satisfy fiber bunching; in this sense, our setting generalizes the classical one.) We say that the cocycle is *irreducible* if has no θ-Hölder invariant subbundles, except for the trivial ones. The main result of Bochi and Garibaldi [ibid.] is that strong fiber bunching together with irreducibility implies the existence of an extremal norm. Let us also mention a curious fact: there are examples where the extremal norm cannot be Riemannian.

The existence of an extremal norm is a first step towards more refined study of maximizing measures: for example, it implies the existence of a *Lyapunov maximizing set*, similarly to the maximizing sets discussed in Section 1. Such sets were studied in Morris [2013] for one-step cocycles.

We can recast in the present context the same type of questions discussed above: How complex are Lyapunov-maximizing measures, either for typical cocycles, or (assuming uniqueness) among all cocycles within a prescribed regularity class?

4 Optimization of all Lyapunov exponents

In this final section, we consider all Lyapunov exponents and not only the top one. Our aim is modest: to introduce an appropriate setting for ergodic optimization of Lyapunov exponents, and to check that the most basic properties seen in the previous sections are still valid.

Let $\mathbf{s}_1(g) \geq \cdots \geq \mathbf{s}_d(g)$ denote the *singular values* of a matrix $g \in \mathrm{GL}(d, \mathbb{R})$. These are the semi-axes of the ellipsoid $g(S^{d-1})$, where S^{d-1} is the unit sphere in $\mathbb{R}^d$. The

[2] Barabanov's norms also have an extra property that does not concern us here. Previously, Rota and Strang [1960] have already considered the weaker notion of extremal *operator* norms.

Cartan projection is the map

$$\vec{\sigma}(g) := \big(\log \mathbf{s}_1(g), \dots, \log \mathbf{s}_d(g)\big)\,, \tag{4-1}$$

which takes values in the *positive chamber*

$$\mathfrak{a}^+ := \left\{(\xi_1, \dots, \xi_d) \in \mathbb{R}^d \; ; \; \xi_1 \geq \cdots \geq \xi_d\right\}.$$

The Cartan projection has the subadditive property

$$\vec{\sigma}(gh) \preccurlyeq \vec{\sigma}(g) + \vec{\sigma}(h)\,, \tag{4-2}$$

where $\preccurlyeq$ denotes the *majorization partial order* in $\mathbb{R}^d$ defined in as follows: $\xi \preccurlyeq \eta$ (which reads as ξ is *majorized* by η) if ξ is a convex combination of vectors obtained by permutation of the entries of η. The group of automorphisms of $\mathbb{R}^d$ consisting of permutation of coordinates is called the *Weyl group* and is denoted $\mathfrak{W}$; so $\xi \preccurlyeq \eta$ if and only if ξ belongs to the polyhedron $\mathrm{co}(\mathfrak{W}\eta)$ (called a *permuthohedron*). For vectors $\xi = (\xi_1, \dots, \xi_d)$ and $\eta = (\eta_1, \dots, \eta_d)$ in the positive chamber $\mathfrak{a}^+$, majorization can be characterized by the following system of inequalities:

$$\xi \preccurlyeq \eta \quad \Leftrightarrow \quad \forall i \in \{1, \dots, d\},\ \xi_1 + \cdots + \xi_i \leq \eta_1 + \cdots + \eta_i \text{ , with equality if } i = d.$$

For a proof, see the book Marshall, Olkin, and Arnold [2011], which contains plenty of information on majorization, including applications.

Now let us consider a linear cocycle (X, T, F). For simplicity, let us assume that the matrices $F(x)$ are invertible. Using Kingman's theorem, one shows that for every $\mu \in \mathfrak{M}_T$ and μ-almost every $x \in X$, the limit

$$\vec{\lambda}(F, x) := \lim_{n \to \infty} \frac{1}{n} \vec{\sigma}(F^{(n)}(x))$$

exists; it is called the *Lyapunov vector* of the point x. Its entries are called the *Lyapunov exponents*. If μ is ergodic then the Lyapunov vector is μ-almost surely equal to a constant $\vec{\lambda}(F, \mu)$. The *Lyapunov spectrum* of the cocycle is defined as:

$$L^+(F) := \left\{\vec{\lambda}(F, \mu) \; ; \; \mu \in \mathcal{E}_T\right\} \subset \mathfrak{a}^+ .$$

By analogy with the rotation set (2-1), we introduce the *inner envelope* of the cocycle as the closed-convex hull of the Lyapunov spectrum, that is, $I^+(F) := \overline{\mathrm{co}}(L^+(F))$. [3]

[3] In recent work, Sert [2017] considers one-step cocycles taking values on more general Lie groups and satisfying a Zariski denseness assumption, introduces and studies a subset of the positive chamber called *joint spectrum*, and applies it to obtain results on large deviations. It turns out that the joint spectrum coincides with our inner envelope $I^+(F)$, in the $\mathrm{SL}(d, \mathbb{R})$ case at least.

Differently from the commutative situation, however, extremal points of this convex set are not necessarily attained. Therefore we introduce other sets (see Figure 1):

$$\begin{aligned} L(F) &:= \mathfrak{W} \cdot L^+(F) && \textit{(symmetric Lyapunov spectrum)}; \\ I(F) &:= \mathfrak{W} \cdot I^+(F) && \textit{(symmetric inner envelope)}; \\ O(F) &:= \mathrm{co}(I(F)) && \textit{(symmetric outer envelope)}; \\ O^+(F) &:= O(F) \cap \mathfrak{a}^+ && \textit{(outer envelope)}. \end{aligned}$$

Then the extremal points of the symmetric outer envelope $O(F)$ are attained as (perhaps reordered) Lyapunov vectors of ergodic measures:

$$\operatorname{ext}(O(F)) \subseteq L(f)\,. \tag{4-3}$$

Sometimes this is the best we can say about attainability, but sometimes we can do better. There is one situation where all extremal points of the symmetric inner envelope $I(f)$ are attained, namely if the cocycle admits a dominated splitting into one-dimensional bundles, because then we are essentially reduced to rotation sets in $\mathbb{R}^d$.

Let us explain the concept of domination. For simplicity, let us assume that $T\colon X \to X$ is a homeomorphism. Let us also assume that there is a fully supported T-invariant probability measure (otherwise we simply restrict T to the minimal center of attraction; see Akin [1993, Prop. 8.8(c)]).

Suppose that V and W are two F-invariant subbundles of constant dimensions. We say that V *dominates* W if there are constants $\kappa_0 > 1$ and $n_0 \geq 1$ such that for every $x \in X$ and every $n \geq n_0$, the smallest singular value of $F^{(n)}(x)|_{V(x)}$ is bigger than κ_0^n times the biggest singular value of $F^{(n)}(x)|_{W(x)}$. We say that V is the *dominating* bundle, and W is the *dominated* bundle. The terminology *exponentially separated splitting* is more common in ODE and Control Theory, and other terms also appear especially in the earlier literature, but we will stick to the terminology *dominated splitting*, though grammatically inferior. The bundles V and W are in fact continuous, and they are robust with respect to perturbations of F: see Bonatti, Díaz, and Viana [2005] for this and other basic properties of domination.

The cocycle has a unique *finest dominated splitting*; this is a finite collection of invariant subbundles $V_1, V_2, \ldots, V_k$, each one dominating the next one, and with maximal k. It is indeed a splitting in the sense that $V_1(x) \oplus \cdots \oplus V_k(x) = \mathbb{R}^d$ for every x. If $k = 1$ then the splitting is called *trivial*.

We say that $i \in \{1, \ldots, d-1\}$ is a *index of domination* of the cocycle if there exists a dominated splitting with a dominating bundle of dimension i; otherwise we say that that i is a *index of non-domination*. So, if $V_1 \oplus \cdots \oplus V_k$ is the finest dominated splitting of the cocycle, and $d_i := \dim V_i$, then the indices of domination are d_1, $d_1 + d_2$, …, $d_1 + d_2 + \cdots + d_{k-1}$.

There is a way of detecting domination without referring to invariant subbundles or cones. As shown in Bochi and Gourmelon [2009], $i \in \{1, \dots, d-1\}$ is a index of domination if and only if there is an exponential gap between i-th and $(i+1)$-th singular values; more precisely: there are constants $\kappa_1 > 1$ and $n_1 \geq 1$ such that $\mathbf{s}_i(F^{(n)}(x))/\mathbf{s}_{i+1}(F^{(n)}(x)) \geq \kappa_1^n$ for all $x \in X$ and all $n \geq n_1$. In terms of the sets

$$\Sigma_n(F) := \left\{\vec{\sigma}(F^{(n)}(x)) \; ; \; x \in X\right\} \subset \mathfrak{a}^+ , \tag{4-4}$$

we have the following geometric characterization: i is an index of domination if and only if for all sufficiently large n, the sets $\frac{1}{n}\Sigma_n(F)$ are uniformly away from the *wall* $\xi_i = \xi_{i+1}$ (a hyperplane that contains part of the boundary of the positive chamber $\mathfrak{a}^+$).

If Θ is a subset of $\{1, 2, \dots, d-1\}$, define the Θ-*superchamber* as the following closed convex subset of $\mathbb{R}^d$:

$$\mathfrak{a}^\Theta := \left\{(\xi_1, \dots, \xi_d) \in \mathbb{R}^d \; ; \; 1 \leq i \leq j < k \leq d \text{ integers}, \; j \notin \Theta \; \Rightarrow \; \xi_i \geq \xi_k\right\} .$$

For example, $\mathfrak{a}^\Theta = \mathfrak{a}^+$ if Θ is empty, and $\mathfrak{a}^\Theta = \mathbb{R}^d$ if $\Theta = \{1, 2, \dots, d-1\}$. (Moreover, $\mathfrak{a}^\Theta$ equals the orbit of $\mathfrak{a}^+$ under an appropriate subgroup of $\mathfrak{W}$, but we will not need this fact.) If C is any subset of $\mathbb{R}^d$, the *closed-Θ-convex* hull of C, denoted by $\overline{\mathrm{co}}_\Theta(C)$, is defined as the smallest closed subset of $\mathbb{R}^d$ that contains C, is invariant under the Weyl group $\mathfrak{W}$, and whose intersection with the superchamber $\mathfrak{a}^\Theta$ is convex.

Let Θ be the set of indices of non-domination of the cocycle (X, T, F). We define the following two sets:

$$\begin{aligned} M(F) &:= \overline{\mathrm{co}}_\Theta\left(L^+(F)\right) \quad && \textit{(symmetric Morse spectrum)}; \\ M^+(F) &:= \mathfrak{a}^+ \cap M(F) \quad && \textit{(Morse spectrum)}. \end{aligned}$$

The Morse spectra (symmetric or not) are sandwiched between the inner and outer envelopes: see Figure 1). If $\Theta = \varnothing$ then $M(F) = I(F)$, while if $\Theta = \{1, \dots, d-1\}$ then $M(F) = O(F)$.

The Morse spectrum allows us to state the following *attainability property*, which is stronger than (4-3):

$$\operatorname{ext}(M(F)) \subseteq L(F) . \tag{4-5}$$

The name "Morse" comes from Morse decompositions in Conley theory Conley [1978]. Morse spectra were originally defined by Colonius and Kliemann [1996]; see also Colonius, Fabbri, and Johnson [2007]. The Morse spectra defined here are more closely related to the ones considered by San Martin and Seco [2010] (who in fact dealt with more general Lie groups). In concrete terms, we have the following characterization: $\xi \in M^+(F)$ if

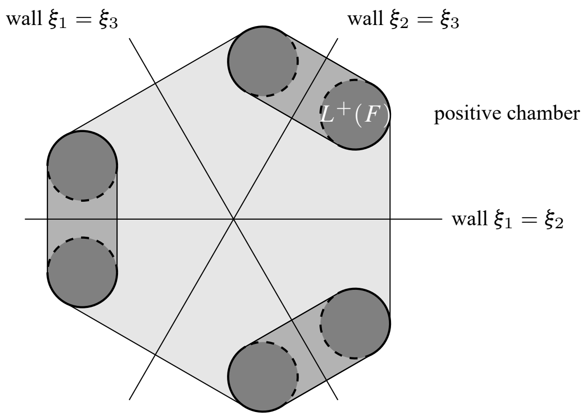


Figure 1: Suppose F takes values in $\mathrm{SL}(3,\mathbb{R})$; then all spectra are contained in the plane $\{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 \ ; \ \xi_1 + \xi_2 + \xi_3 = 0\}$. The figure shows a possibility for the three sets $I(F) \subseteq M(F) \subseteq O(F)$, which are are pictured in decreasing shades of gray, assuming that the unique index of domination is 1, i.e., $\Theta = \{2\}$.

and only if there exist sequences $n_i \to \infty$ and $\varepsilon_i \to 0$, ε_i-pseudo orbits $(x_{i,0}, x_{i,1}, \dots)$, and matrices $g_{i,j} \in \mathrm{GL}(d,\mathbb{R})$ with $\|g_{i,j} - F(x_{i,j})\| < \varepsilon_i$ such that:

$$\frac{1}{n_i} \vec{\sigma}\left(g_{i,n_i-1} \cdots g_{i,1} g_{i,0}\right) \to \xi \,.$$

We will not provide a proof. For background on Morse spectra defined in terms of pseudo orbits and relations with dominated splittings and Lyapunov exponents (and without Lie algebra terminology), see the book Colonius and Kliemann [2014]. Let us remark that all those type of Morse spectra contain more information that the Mather [1968] and Sacker–Sell spectra Dragičević [2017]. Indeed, the finest dominated splitting refines the Sacker–Sell decomposition, which may be seen as the finest *absolute* (as opposed to *pointwise*) dominated splitting.

Let us inspect other basic properties of the Morse spectra which are similar to properties of the rotation sets seen in Section 2. The following *enveloping property* is analogous to (2-2):

$$M(F) = \bigcap_{n \in \mathbb{N}} \frac{1}{n} \overline{\mathrm{co}}_\Theta(\Sigma_n(F)) \,, \tag{4-6}$$

and again the intersection above is a limit. In particular, the Morse spectrum $M(\cdot)$ is an upper semicontinuous function of F (with respect to the uniform topology). If the cocycle has a dominated splitting into $k = d$ bundles then $M(\cdot)$ is continuous at F.

Two cocycles F and G over the same base (X, T) are called *conjugate* if there is continuous map $H: X \to \mathrm{GL}(d, \mathbb{R})$ such that $G(x) = H(Tx)^{-1} F(x) H(x)$. The Morse spectrum is invariant under cocycle conjugation.

Let us state a result similar to the duality property (1-3) and its vectorial counterpart explained in Section 2:

Proposition 4.1. *Given a neighborhood U of $M^+(F)$, there exists a cocycle G conjugate to F such that $\Sigma_1(G) \subset U$.*

The proof requires some preliminaries. Let $\mathcal{S}$ denote the space of inner products in $\mathbb{R}^d$. The group $\mathrm{GL}(d, \mathbb{R})$ acts transitively on $\mathcal{S}$ as follows:

$$\langle\!\langle \cdot, \cdot \rangle\!\rangle_2 = g * \langle\!\langle \cdot, \cdot \rangle\!\rangle_1 \quad \Leftrightarrow \quad \langle\!\langle u, v \rangle\!\rangle_2 = \langle\!\langle g^{-1}u, g^{-1}v \rangle\!\rangle_1 \,.$$

Using the standard inner product $\langle \cdot, \cdot \rangle =: o$ as a reference, every element $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ of $\mathcal{S}$ can be uniquely represented by a positive (i.e. positive-definite symmetric) matrix p such that $\langle\!\langle u, v \rangle\!\rangle = \langle p^{-1}u, v \rangle$. In this way we may identify $\mathcal{S}$ with the set of positive matrices, and o is identified with the identity matrix. In these terms, the group action becomes:

$$g * p = gpg^\dagger \,.$$

The *vectorial distance* is defined as the following map:

$$\vec{\delta} \colon \mathcal{S} \times \mathcal{S} \to \mathfrak{a}^+ \,, \qquad \vec{\delta}(p, q) := 2\vec{\sigma}(p^{-1/2} q^{1/2}) \,,$$

where $\vec{\sigma}$ is the Cartan projection (4-1). It has the following properties:

1. $\vec{\delta}(o, q) = \vec{\sigma}(q)$;
2. $\vec{\delta}$ is a complete invariant for the group action on pairs of points, that is, $\vec{\delta}(p_1, q_1) = \vec{\delta}(p_2, q_2)$ if and only if there exists $g \in \mathrm{GL}(d, \mathbb{R})$ such that $g * p_1 = p_2$ and $g * q_1 = q_2$;
3. $\vec{\delta}(p, p) = 0$;
4. $\vec{\delta}(q, p) = \mathbf{i}(\vec{\delta}(p, q))$, where $\mathbf{i}(\xi_1, \dots, \xi_d) := (-\xi_d, \dots, -\xi_1)$ is the *opposition involution*;
5. triangle inequality: $\vec{\delta}(p, r) \preccurlyeq \vec{\delta}(p, q) + \vec{\delta}(q, r)$; this follows from (4-2).

In particular, the euclidian norm of $\vec{\delta}$ is a true distance function, and it invariant under the action; indeed that is the usual way to metrize $\mathcal{S}$.

A parameterized curve $\gamma\colon [0,1] \to \mathcal{S}$ is called a *geodesic segment* if there is a vector $\xi \in \mathfrak{a}^+$ such that $\vec{\delta}(f(t), f(s)) = (s-t)\xi$, provided $t \le s$. A geodesic segment is determined uniquely by its endpoints $p = f(0)$ and $q = f(1)$; it is given by the formula $f(s) = q^s$ if $p = o$. The image of f is denoted $[p,q]$ and by abuse of terminology is also called a geodesic segment. The *midpoint* of the geodesic segment is $\operatorname{mid}[p,q] := f(1/2)$.

We shall prove the following vectorial version of the Busemann nonpositive curvature inequality:

$$\vec{\delta}\big(\operatorname{mid}[r,p], \operatorname{mid}[r,q]\big) \preccurlyeq \tfrac{1}{2}\vec{\delta}(p,q)\,, \quad \text{for all } r, p, q \in \mathcal{S}\,. \tag{4-7}$$

Parreau [2017] has announced a general version of this inequality that holds in other symmetric spaces and affine buildings, using the appropriate partial order.

In order to prove (4-7), consider the *Jordan projection* $\vec{\chi}\colon \mathrm{GL}(d,\mathbb{R}) \to \mathfrak{a}^+$ defined by $\vec{\chi}(g) := \big(\log|z_1|, \dots, \log|z_d|\big)$, where $z_1, \dots, z_d$ are the eigenvalues of g, ordered so that $|z_1| \ge \cdots \ge |z_d|$. The Jordan projection is cyclically invariant, that is, $\vec{\chi}(gh) = \vec{\chi}(hg)$. The Cartan and Jordan projections are related by $\vec{\sigma}(g) = \frac{1}{2}\vec{\chi}(gg^\dagger)$. Another property is that Cartan majorizes Jordan: $\vec{\sigma}(g) \succcurlyeq \vec{\chi}(g)$. This follows from the fact that the spectral radius of a matrix is less than or equal to its top singular value, applied to g and its exterior powers.

Proof of the vectorial Busemann NPC inequality (4-7). Take arbitraries $r, p, q \in \mathcal{S}$. Since the vectorial distance is invariant under the group action, it is sufficient to consider the case where $r = o$. Then the midpoints under consideration are $p^{1/2}$ and $q^{1/2}$. Using the definition of $\vec{\delta}$ and the properties of the projections $\vec{\sigma}$ and $\vec{\chi}$, we have:

$$\begin{aligned}\vec{\delta}(p^{1/2}, q^{1/2}) = 2\vec{\sigma}\big(p^{-1/4}q^{1/4}\big) &= \vec{\chi}\big(p^{-1/4}q^{1/2}p^{-1/4}\big)\\ &= \vec{\chi}\big(p^{-1/2}q^{1/2}\big) \preccurlyeq \vec{\sigma}\big(p^{-1/2}q^{1/2}\big) = \tfrac{1}{2}\vec{\delta}(p,q)\,. \qquad \square\end{aligned}$$

Proof of Proposition 4.1. We will adapt the argument from Bochi and Navas [2015, p. 383–384], using (4-7) instead of the ordinary Busemann NPC inequality. Consider the case $\Theta = \{1, \dots, d-1\}$, i.e., the cocycle no nontrivial dominated splitting. Then $M^+(F)$ is closed under majorization, in the sense that:

$$\xi \in \mathfrak{a}^+,\ \eta \in M^+(F),\ \xi \preccurlyeq \eta \quad \Rightarrow \quad \xi \in M^+(F)$$

Fix a neighborhood $U \supset M^+(F)$. Without loss of generality, we may assume that $U \cap \mathfrak{a}^+$ is closed under majorization. We want to find G conjugate to F such that $\Sigma_1(G) \subset U$. By (4-6), for sufficiently large N we have $\frac{1}{N}\Sigma_N(F) \subset U$. Fix such N of the form $N = 2^k$.

Let us define recursively continuous maps $\psi_0, \psi_1, \dots, \psi_k\colon X \to \mathcal{S}$ as follows: ψ_0 is constant equal to o, and

$$\psi_{j+1}(x) := \operatorname{mid}\big[(F^{(2^{k-j-1})}(x))^{-1} * \psi_j(T^{2^{k-j-1}}x), \psi_j(x)\big]\,,$$

Then:

$$\underbrace{\psi_{j+1}(T^{2^{k-j-1}}x)}_{①} = \operatorname{mid}\big[\underbrace{(F^{(2^{k-j-1})}(T^{2^{k-j-1}}x))^{-1} * \psi_j(T^{2^{k-j}}x)}_{②}, \psi_j(T^{2^{k-j-1}}x)\big]$$

and, using equivariance of midpoints,

$$\underbrace{F^{(2^{k-j-1})}(x) * \psi_{j+1}(x)}_{③} = \operatorname{mid}\big[\psi_j(T^{2^{k-j-1}}x), \underbrace{F^{(2^{k-j-1})}(x) * \psi_j(x)}_{④}\big].$$

By (4-7), we have $\vec{\delta}(①,③) \preccurlyeq \frac{1}{2}\vec{\delta}(②,④)$, which, by the invariance of the vectorial distance, amounts to

$$\vec{\delta}\big(\psi_{j+1}(T^{2^{k-j-1}}x), F^{(2^{k-j-1})}(x)*\psi_{j+1}(x)\big) \preccurlyeq \tfrac{1}{2}\vec{\delta}\big(\psi_j(T^{2^{k-j}}x), F^{(2^{k-j})}(x)*\psi_j(x)\big).$$

Combining the whole chain of these inequalities we obtain:

$$\vec{\delta}\big(\psi_k(Tx), F(x) * \psi_k(x)\big) \preccurlyeq \tfrac{1}{2^k}\vec{\delta}\big(\psi_0(T^{2^k}x), F^{(2^k)}(x) * \psi_0(x)\big).$$

Equivalently, denoting $\varphi := \psi_k$,

$$\vec{\delta}\big(\varphi(Tx), F(x) * \varphi(x)\big) \preccurlyeq \tfrac{1}{N}\vec{\delta}\big(o, F^{(N)}(x) * o\big). \tag{4-8}$$

Take a continuous map $H : X \to \mathrm{GL}(d,\mathbb{R})$ such that $H(x) * \varphi(x) = o$ for every x (e.g., $H := \varphi^{-1/2}$), and let $G(x) := H(Tx)F(x)H(x)^{-1}$. Then

$$\begin{aligned}\vec{\delta}\big(\varphi(Tx), F(x) * \varphi(x)\big) &= \vec{\delta}\big(o, G(x) * o\big)\\ &= \vec{\sigma}\big(G(x) * o\big) \preccurlyeq \tfrac{1}{N}\vec{\sigma}\big(F^{(N)}(x) * o\big) \in \tfrac{1}{N}\Sigma_N(F) \subseteq U \cap \mathfrak{a}^+.\end{aligned}$$

Since $U \cap \mathfrak{a}^+$ is closed under majorization, we conclude that $\vec{\sigma}\big(G(x) * o\big) \in U$. That is, $\Sigma_1(G) \subseteq U$, as we wanted to show.

In the case the cocycle admits a nontrivial dominated splitting, we take a preliminary conjugation to make the bundles of the finest dominated splitting orthogonal. Then the exact same procedure above leads to the desired conjugation, but we omit the verifications. □

As a corollary of Proposition 4.1, we reobtain a result of Gourmelon [2007], which says that it is always possible to find an adapted Riemannian norm for which dominations are seen in the first iterate (i.e., $n_0 = 1$ in our definition). Indeed, the corresponding inner product at the point x is $\varphi(x)$, where φ is the map constructed in the proof of Proposition 4.1.

Furthermore, the construction gives as extra property which is essential to certain applications in Bochi, Katok, and Rodriguez Hertz [n.d.], namely: fixed a favored ergodic measure $\mu_0 \in \mathcal{E}_T$, we can choose the adapted metric φ with respect to which the expansion rates in the first iterate are close to the Lyapunov exponents with respect to μ_0, except on a set of small μ_0 measure.[4] More precisely, we can take N large enough so that the RHS in (4-8) is $L^1(\mu_0)$-close to the Lyapunov vector $\vec{\lambda}(F, \mu_0)$. On the other hand, the integral of the LHS majorizes the Lyapunov vector. It follows that the RHS is also $L^1(\mu_0)$-close to the Lyapunov vector.

The measures $\mu \in \mathfrak{M}_T$ for which the Lyapunov vector $\vec{\lambda}(F, \mu)$ is an extremal point of the symmetric Morse set $M(F)$ are called *extremal measures* for the linear cocycle (X, T, F).

As an example, consider the one-step cocycle generated by the pair of matrices $A_1 := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $A_2 := \left(\begin{smallmatrix} 2 & 0 \\ 2 & 2 \end{smallmatrix}\right)$. The extremal measures of this cocycle are Sturmian: this can be deduced from a result of Hare, Morris, Sidorov, and Theys [2011][5]. Moreover, results of Morris and Sidorov [2013] imply that the boundary of the symmetric Morse set is not differentiable, with a dense subset of corners, just like the fish seen in Section 2. Again, we ask: are these phenomena typical?

Acknowledgments. I thank Ian D. Morris and Cagri Sert for corrections and suggestions.

References

Ethan Akin (1993). *The general topology of dynamical systems*. Vol. 1. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. x+261. MR: 1219737 (cit. on p. 1855).

N. E. Barabanov (1988a). "On the Lyapunov exponent of discrete inclusions. I". *Avtomat. i Telemekh.* 2, pp. 40–46. MR: 940263 (cit. on p. 1853).

– (1988b). "On the Lyapunov exponent of discrete inclusions. I". *Automat. Remote Control* 49.2, pp. 152–157.

Vincent D. Blondel, Jacques Theys, and Alexander A. Vladimirov (2003). "An elementary counterexample to the finiteness conjecture". *SIAM J. Matrix Anal. Appl.* 24.4, pp. 963–970. MR: 2003315 (cit. on p. 1851).

J. Bochi and E. Garibaldi (n.d.). *Extremal norms for fiber-bunched cocycles*. In preparation (cit. on p. 1853).

[4]Lyapunov metrics used in Pesin theory (see e.g. Katok and Hasselblatt [1995, p. 668]) satisfy such a property in a set of full measure, but they are only measurable, and are not necessarily adapted to the finest dominated splitting.

[5]Namely: for every $\alpha > 0$, the one-step cocycle generated by $(A_0, \alpha A_1)$ has a unique Lyapunov-maximizing measure, which is Sturmian.

J. Bochi, A. Katok, and F. Rodriguez Hertz (n.d.). *Flexibility of Lyapunov exponents among conservative diffeormophisms*. In preparation (cit. on p. 1861).

Jairo Bochi and Nicolas Gourmelon (2009). "Some characterizations of domination". *Math. Z.* 263.1, pp. 221–231. MR: 2529495 (cit. on p. 1856).

Jairo Bochi and Ian D. Morris (2015). "Continuity properties of the lower spectral radius". *Proc. Lond. Math. Soc. (3)* 110.2, pp. 477–509. MR: 3335285 (cit. on pp. 1851, 1852).

Jairo Bochi and Andrés Navas (2015). "A geometric path from zero Lyapunov exponents to rotation cocycles". *Ergodic Theory Dynam. Systems* 35.2, pp. 374–402. MR: 3316917 (cit. on pp. 1851, 1859).

Jairo Bochi and Michał Rams (2016). "The entropy of Lyapunov-optimizing measures of some matrix cocycles". *J. Mod. Dyn.* 10, pp. 255–286. MR: 3538864 (cit. on p. 1852).

Jairo Bochi and Yiwei Zhang (2016). "Ergodic optimization of prevalent super-continuous functions". *Int. Math. Res. Not. IMRN* 19, pp. 5988–6017. MR: 3567264 (cit. on p. 1847).

Christian Bonatti, Lorenzo J. Díaz, and Marcelo Viana (2005). *Dynamics beyond uniform hyperbolicity*. Vol. 102. Encyclopaedia of Mathematical Sciences. A global geometric and probabilistic perspective, Mathematical Physics, III. Springer-Verlag, Berlin, pp. xviii+384. MR: 2105774 (cit. on p. 1855).

Thierry Bousch (2000). "Le poisson n'a pas d'arêtes". *Ann. Inst. H. Poincaré Probab. Statist.* 36.4, pp. 489–508. MR: 1785392 (cit. on p. 1848).

– (2001). "La condition de Walters". *Ann. Sci. École Norm. Sup. (4)* 34.2, pp. 287–311. MR: 1841880 (cit. on p. 1846).

– (2002). "Un lemme de Mañé bilatéral". *C. R. Math. Acad. Sci. Paris* 335.6, pp. 533–536. MR: 1936826 (cit. on p. 1849).

– (2011). "Le lemme de Mañé-Conze-Guivarc'h pour les systèmes amphidynamiques rectifiables". *Ann. Fac. Sci. Toulouse Math. (6)* 20.1, pp. 1–14. MR: 2829831 (cit. on p. 1846).

Thierry Bousch and Oliver Jenkinson (2002). "Cohomology classes of dynamically non-negative C^k functions". *Invent. Math.* 148.1, pp. 207–217. MR: 1892849 (cit. on pp. 1846, 1849).

Thierry Bousch and Jean Mairesse (2002). "Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture". *J. Amer. Math. Soc.* 15.1, pp. 77–111. MR: 1862798 (cit. on pp. 1851, 1852).

H. W. Broer, B. Hasselblatt, and F. Takens, eds. (2010). *Handbook of dynamical systems. Vol. 3*. Elsevier/North-Holland, Amsterdam, pp. xi+543. MR: 3292649 (cit. on p. 1847).

Yongluo Cao (2003). "Non-zero Lyapunov exponents and uniform hyperbolicity". *Nonlinearity* 16.4, pp. 1473–1479. MR: 1986306 (cit. on p. 1851).

Yongluo Cao, Stefano Luzzatto, and Isabel Rios (2006). "Some non-hyperbolic systems with strictly non-zero Lyapunov exponents for all invariant measures: horseshoes with internal tangencies". *Discrete Contin. Dyn. Syst.* 15.1, pp. 61–71. MR: 2191385 (cit. on p. 1851).

Jens Peter Reus Christensen (1972). "On sets of Haar measure zero in abelian Polish groups". *Israel J. Math.* 13, 255–260 (1973). MR: 0326293 (cit. on p. 1847).

Fritz Colonius, Roberta Fabbri, and Russell Johnson (2007). "Chain recurrence, growth rates and ergodic limits". *Ergodic Theory Dynam. Systems* 27.5, pp. 1509–1524. MR: 2358975 (cit. on p. 1856).

Fritz Colonius and Wolfgang Kliemann (1996). "The Morse spectrum of linear flows on vector bundles". *Trans. Amer. Math. Soc.* 348.11, pp. 4355–4388. MR: 1329532 (cit. on p. 1856).

– (2014). *Dynamical systems and linear algebra*. Vol. 158. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xvi+284. MR: 3242107 (cit. on p. 1857).

Charles Conley (1978). *Isolated invariant sets and the Morse index*. Vol. 38. CBMS Regional Conference Series in Mathematics. American Mathematical Society, Providence, R.I., pp. iii+89. MR: 511133 (cit. on p. 1856).

G. Contreras, A. O. Lopes, and Ph. Thieullen (2001). "Lyapunov minimizing measures for expanding maps of the circle". *Ergodic Theory Dynam. Systems* 21.5, pp. 1379–1409. MR: 1855838 (cit. on p. 1846).

Gonzalo Contreras (2016). "Ground states are generically a periodic orbit". *Invent. Math.* 205.2, pp. 383–412. MR: 3529118 (cit. on p. 1846).

Jean-Pierre Conze and Yves Guivarc'h (1993). "Croissance des sommes ergodiques et principe variationnel". Manuscript (cit. on pp. 1845, 1846, 1848).

Davor Dragičević (2017). "A note on the theorem of Johnson, Palmer and Sell". *Period. Math. Hungar.* 75.2, pp. 167–171. MR: 3718510 (cit. on pp. 1851, 1857).

H. Furstenberg and Y. Kifer (1983). "Random matrix products and measures on projective spaces". *Israel J. Math.* 46.1-2, pp. 12–32. MR: 727020 (cit. on p. 1845).

Eduardo Garibaldi (2017). *Ergodic optimization in the expanding case*. SpringerBriefs in Mathematics. Concepts, tools and applications. Springer, Cham, pp. viii+73. MR: 3701349 (cit. on p. 1846).

Nikolaz Gourmelon (2007). "Adapted metrics for dominated splittings". *Ergodic Theory Dynam. Systems* 27.6, pp. 1839–1849. MR: 2371598 (cit. on p. 1860).

Leonid Gurvits (1995). "Stability of discrete linear inclusion". *Linear Algebra Appl.* 231, pp. 47–85. MR: 1361100 (cit. on p. 1851).

Kevin G. Hare, Ian D. Morris, and Nikita Sidorov (2013). "Extremal sequences of polynomial complexity". *Math. Proc. Cambridge Philos. Soc.* 155.2, pp. 191–205. MR: 3091514 (cit. on p. 1852).

Kevin G. Hare, Ian D. Morris, Nikita Sidorov, and Jacques Theys (2011). "An explicit counterexample to the Lagarias-Wang finiteness conjecture". *Adv. Math.* 226.6, pp. 4667–4701. MR: 2775881 (cit. on p. 1861).

Brian R. Hunt, Tim Sauer, and James A. Yorke (1992). "Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces". *Bull. Amer. Math. Soc. (N.S.)* 27.2, pp. 217–238. MR: 1161274 (cit. on p. 1847).

Brian R Hunt and Edward Ott (1996). "Optimal periodic orbits of chaotic systems occur at low period". *Physical Review E* 54.1, pp. 328–337 (cit. on pp. 1846, 1848).

Oliver Jenkinson (1996). "Conjugacy rigidity, cohomological triviality and barycentres of invariant measures". PhD thesis. University of Warwick (cit. on p. 1848).

– (2000). "Frequency locking on the boundary of the barycentre set". *Experiment. Math.* 9.2, pp. 309–317. MR: 1780215 (cit. on p. 1848).

– (2006a). "Ergodic optimization". *Discrete Contin. Dyn. Syst.* 15.1, pp. 197–224. MR: 2191393 (cit. on pp. 1844, 1845, 1847).

– (2006b). "Every ergodic measure is uniquely maximizing". *Discrete Contin. Dyn. Syst.* 16.2, pp. 383–392. MR: 2226487 (cit. on p. 1845).

– (Dec. 2017). "Ergodic optimization in dynamical systems". arXiv: 1712.02307 (cit. on p. 1844).

Raphaël Jungers (2009). *The joint spectral radius*. Vol. 385. Lecture Notes in Control and Information Sciences. Theory and applications. Springer-Verlag, Berlin, pp. xiv+144. MR: 2507938 (cit. on pp. 1851, 1853).

Anatole Katok and Boris Hasselblatt (1995). *Introduction to the modern theory of dynamical systems*. Vol. 54. Encyclopedia of Mathematics and its Applications. With a supplementary chapter by Katok and Leonardo Mendoza. Cambridge University Press, Cambridge, pp. xviii+802. MR: 1326374 (cit. on p. 1861).

Alejandro Kocsard (2013). "On cohomological C^0-(in)stability". *Bull. Braz. Math. Soc. (N.S.)* 44.3, pp. 489–495. MR: 3124747 (cit. on p. 1849).

Tamara Kucherenko and Christian Wolf (2014). "Geometry and entropy of generalized rotation sets". *Israel J. Math.* 199.2, pp. 791–829. MR: 3219558 (cit. on p. 1850).

Jeffrey C. Lagarias and Yang Wang (1995). "The finiteness conjecture for the generalized spectral radius of a set of matrices". *Linear Algebra Appl.* 214, pp. 17–42. MR: 1311628 (cit. on p. 1851).

Artur O. Lopes and Philippe Thieullen (2003). "Sub-actions for Anosov diffeomorphisms". *Astérisque* 287. Geometric methods in dynamics. II, pp. xix, 135–146. MR: 2040005 (cit. on p. 1846).

Mohsen Maesumi (2008). "Optimal norms and the computation of joint spectral radius of matrices". *Linear Algebra Appl.* 428.10, pp. 2324–2338. MR: 2408030 (cit. on p. 1851).

Albert W. Marshall, Ingram Olkin, and Barry C. Arnold (2011). *Inequalities: theory of majorization and its applications*. Second. Springer Series in Statistics. Springer, New York, pp. xxviii+909. MR: 2759813 (cit. on p. 1854).

John N. Mather (1968). "Characterization of Anosov diffeomorphisms". *Nederl. Akad. Wetensch. Proc. Ser. A 71 = Indag. Math.* 30, pp. 479–483. MR: 0248879 (cit. on p. 1857).

Filippo Mignosi (1991). "On the number of factors of Sturmian words". *Theoret. Comput. Sci.* 82.1, Algorithms Automat. Complexity Games, pp. 71–84. MR: 1112109 (cit. on p. 1850).

Ian D. Morris (2010). "Ergodic optimization for generic continuous functions". *Discrete Contin. Dyn. Syst.* 27.1, pp. 383–388. MR: 2600778 (cit. on p. 1845).

– (2013). "Mather sets for sequences of matrices and applications to the study of joint spectral radii". *Proc. Lond. Math. Soc. (3)* 107.1, pp. 121–150. MR: 3083190 (cit. on pp. 1851, 1853).

Ian D. Morris and Nikita Sidorov (2013). "On a devil's staircase associated to the joint spectral radii of a family of pairs of matrices". *J. Eur. Math. Soc. (JEMS)* 15.5, pp. 1747–1782. MR: 3082242 (cit. on pp. 1852, 1861).

Eduardo Oregón-Reyes (June 2016). "Properties of sets of isometries of Gromov hyperbolic spaces". To appear in *Groups Geom. Dyn.* arXiv: 1606.01575 (cit. on p. 1851).

Anne. Parreau (2017). *Vectorial metric compactification of symmetric spaces and affine buildings* (cit. on p. 1859).

Gabriel P. Paternain (1999). *Geodesic flows*. Vol. 180. Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA, pp. xiv+149. MR: 1712465 (cit. on p. 1848).

Mark Pollicott and Richard Sharp (2004). "Livsic theorems, maximizing measures and the stable norm". *Dyn. Syst.* 19.1, pp. 75–88. MR: 2038273 (cit. on pp. 1846, 1848, 1850).

Lucian Radu (2004). "Duality in thermodynamic formalism". Manuscript (cit. on p. 1845).

Gian-Carlo Rota and Gilbert Strang (1960). "A note on the joint spectral radius". *Nederl. Akad. Wetensch. Proc. Ser. A 63 = Indag. Math.* 22, pp. 379–381. MR: 0147922 (cit. on pp. 1851, 1853).

Luiz A. B. San Martin and Lucas Seco (2010). "Morse and Lyapunov spectra and dynamics on flag bundles". *Ergodic Theory Dynam. Systems* 30.3, pp. 893–922. MR: 2643716 (cit. on p. 1856).

S. V. Savchenko (1999). "Homological inequalities for finite topological Markov chains". *Funktsional. Anal. i Prilozhen.* 33.3, pp. 91–93. MR: 1724277 (cit. on p. 1846).

Cagri Sert (2017). "Joint spectrum and large deviation principle for random matrix products". *C. R. Math. Acad. Sci. Paris* 355.6, pp. 718–722. MR: 3661555 (cit. on p. 1854).

Fabian Wirth (2002). "The generalized spectral radius and extremal norms". *Linear Algebra Appl.* 342, pp. 17–40. MR: 1873424 (cit. on p. 1853).

Guocheng Yuan and Brian R. Hunt (1999). "Optimal orbits of hyperbolic systems". *Nonlinearity* 12.4, pp. 1207–1224. MR: 1709845 (cit. on p. 1846).

Received 2018-02-02.

JAIRO BOCHI
FACULTAD DE MATEMÁTICAS
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
jairo.bochi@mat.uc.cl

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1865–1884)

A BRIEF INTRODUCTION TO SOFIC ENTROPY THEORY

LEWIS P. BOWEN

Abstract

Sofic entropy theory is a generalization of the classical Kolmogorov-Sinai entropy theory to actions of a large class of non-amenable groups called sofic groups. This is a short introduction with a guide to the literature.

1 Introduction

Classical entropy theory is concerned with single transformations of a topological or measure space. This can be generalized straightforwardly to actions of the lattice $\mathbb{Z}^d$. However, one encounters real difficulty in any attempt to generalize to actions of non-amenable groups. This short survey will begin with the free group of rank 2, $\mathbb{F}_2 := \langle a, b \rangle$. The Cayley graph $G = (V, E)$ of this group has vertex set $V = \mathbb{F}_2$ and directed edges $(g, ga), (g, gb)$ for $g \in \mathbb{F}_2$. It is a 4-regular tree. It is non-amenable because any finite subset $F \subset V$ has the property that if ∂F is the set of edges $e \in E$ with one end in F and one end outside of F then $|\partial F| \geq 2|F|$. After understanding the special case of the free group (from a dynamicist's view), we will generalize to residually finite groups and sofic groups and then briefly survey recent developments; namely the classification of Bernoulli shifts, Bernoulli factors, Rokhlin entropy theory, algebraic dynamics and the geometry of model spaces.

We will not define amenability here (see Kerr and Li [2016] for example). We will also not cover classical entropy theory. The interested reader is encouraged to consult one of the standard texts (e.g. Petersen [1989]). Other introductions and surveys on sofic entropy theory include Weiss [2015], Kerr and Li [2016], Gaboriau [2016], and L. Bowen [2017a].

Supported in part by NSF grant DMS-1500389 and a Simons Fellowship.
MSC2010: 37A35.
Keywords: entropy, sofic groups.

1.1 The Ornstein-Weiss factor. In 1987, Ornstein and Weiss exhibited a curious example D. S. Ornstein and Weiss [1987]. To explain it, let $X := (\mathbb{Z}/2\mathbb{Z})^{\mathbb{F}_2}$ be the set of all maps $x : \mathbb{F}_2 \to \mathbb{Z}/2\mathbb{Z}$. This is a compact abelian group under pointwise addition. We can identify $X \times X$ with the group of maps from $\mathbb{F}_2 \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Define $\Phi : X \to X \times X$ by

$$\Phi(x)(g) = (x(g) - x(ga), x(g) - x(gb)).$$

This a surjective homomorphism. It is also $\mathbb{F}_2$-equivariant where $\mathbb{F}_2$ acts on X by

$$(fx)(g) := x(f^{-1}g).$$

And its kernel consists of the two constant maps. So it is a continuous, algebraic, 2-1 factor map from $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{F}_2}$ onto $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})^{\mathbb{F}_2}$.

This appears to give a contradiction to entropy theory because if K is a finite set then the entropy of the action of $\mathbb{F}_2$ on $K^{\mathbb{F}_2}$ should be $\log|K|$. So the Ornstein-Weiss factor map increases entropy! At the time of D. S. Ornstein and Weiss [ibid.] it was unknown whether or not the two actions $\mathbb{F}_2 \curvearrowright X$ and $\mathbb{F}_2 \curvearrowright X \times X$ could be measurably conjugate (with respect to Haar measure on X and $X \times X$). We will show by Theorem 7.1 below that they are not measurably conjugate.

2 Topological entropy for $\mathbb{Z}$-actions

Here we will develop entropy theory for $\mathbb{Z}$-actions in a slightly non-traditional way which generalizes to actions of free groups. To begin, consider a homeomorphism $T : X \to X$ of a compact metric space (X, d). A *partial orbit* of length n is a tuple of the form $(x, Tx, \dots, T^{n-1}x) \in X^n$. Define a metric $d_\infty^{(n)}$ on X^n by

$$d_\infty^{(n)}(\underline{x}, \underline{y}) = \max_i d(x_i, y_i)$$

where, for example $\underline{x} = (x_1, \dots, x_n)$. Then Rufus Bowen's definition of the topological entropy of T is

$$h(T) := \sup_{\epsilon>0} \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\{\text{length } n \text{ partial orbits}\}, d_\infty^{(n)})$$

where if $S \subset X^n$ then $\mathrm{Sep}_\epsilon(S, d_\infty)$ denotes the maximum cardinality of an ϵ-separated subset $Y \subset S$ R. Bowen [1971]. (Recall that a subset $Y \subset S$ is *ϵ-separated* if $d(y,z) > \epsilon$ for any $y, z \in Y$ with $y \neq z$).

Instead of counting partial orbits to compute entropy, we can count pseudo-orbits. To be precise, an n-tuple $\underline{x} \in X^n$ is an *(n, δ)-pseudo orbit* if

$$\frac{1}{n} \sum_{i=1}^{n-1} d(Tx_i, x_{i+1}) < \delta.$$

By Markov's inequality and continuity, any pseudo orbit contains a long subword that is close to a partial orbit. Hence

$$h(T) = \sup_{\epsilon>0} \inf_{\delta>0} \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\{(n,\delta) \text{ pseudo orbits}\}, d_\infty^{(n)}).$$

This is similar to Katok's treatment of entropy in Katok [1980] and follows from Kerr-Li's approach to entropy in Kerr and Li [2011b].

It is of classical interest to count periodic orbits too. We will say that $\underline{x} \in X^n$ is *periodic with period* $\leq n$ if $Tx_i = x_{i+1}$ for all $1 \leq i \leq n-1$ and $Tx_n = x_1$. The growth rate of periodic orbits is a lower bound for the entropy rate but in general they are not equal. To remedy this, let us consider pseudo-periodic orbits. To be precise, a n-tuple $\underline{x} \in X^n$ is an (n,δ)-*pseudo-periodic orbit* if

$$\frac{1}{n}\left(\sum_{i=1}^{n-1} d(Tx_i, x_{i+1}) + d(Tx_n, x_1)\right) < \delta.$$

Since (n,δ)-pseudo-periodic orbits are (n,δ)-pseudo orbits and (n,δ)-pseudo orbits are $(n, \delta + o_n(1))$-pseudo-periodic orbits, it follows that

$$h(T) = \sup_{\epsilon>0} \inf_{\delta>0} \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\{(n,\delta) \text{ pseudo-periodic orbits}\}, d_\infty^{(n)}).$$

Now it might look like we have not gained much by these observations since pseudo-periodic orbits are almost the same as pseudo orbits and the latter shadow partial orbits of only slightly less length. However, there is a conceptual advantage. This is because pseudo-periodic orbits can be thought of as maps from an external model, namely $\mathbb{Z}/n\mathbb{Z}$, to X that approximate the dynamics. It is a very useful observation that we are not required to count only partial orbits or periodic orbits, both of which are too restrictive to generalize to actions of free groups.

3 Topological sofic entropy for actions of free groups

Suppose the free group $\mathbb{F}_2$ acts on a compact space X by homeomorphisms. A *periodic orbit* of this action consists of a finite set V_0, an action of $\mathbb{F}_2$ on V_0 and an $\mathbb{F}_2$-equivariant map $\phi : V_0 \to X$.

It can be helpful to visualize the action of $\mathbb{F}_2$ on V_0 by making the *action graph* $G_0 = (V_0, E_0)$ whose edges consist of all pairs of the form $(v, a\cdot v)$ and $(v, b\cdot v)$ for $v \in V_0$. It is a directed graph in which every vertex has in-degree and out-degree 2.

Given a finite subset $F \subset \mathbb{F}_2$ and $\delta > 0$, a (V_0, δ, F)-*pseudo-periodic orbit* is a map $\phi : V_0 \to X$ that is approximately equivariant in the sense that

$$|V_0|^{-1} \sum_{v \in V_0} d(\phi(g \cdot v), g \cdot \phi(v)) < \delta$$

for every $g \in F$.

Now suppose that we fix a sequence $\Sigma := \{\mathbb{F}_2 \curvearrowright V_i\}_{i=1}^{\infty}$ of actions of $\mathbb{F}_2$ on finite sets V_i. Tentatively, we will call the *topological sofic entropy of* $\mathbb{F}_2 \curvearrowright X$ *with respect to* Σ the quantity

$$h_\Sigma(\mathbb{F}_2 \curvearrowright X) := \sup_{\epsilon > 0} \inf_{\delta > 0} \inf_{F \Subset \mathbb{F}_2} \limsup_{n \to \infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\{(V_n, \delta, F) \text{ pseudo-periodic orbits}\}, d_\infty^{V_n}).$$

where $d_\infty^{V_n}$ is the metric on X^{V_n} defined by

$$d_\infty^{V_n}(\phi, \psi) := \max_{v \in V_n} d(\phi(v), \psi(v)).$$

To state the next result, we need some terminology. Given a countable group Γ and continuous actions $\Gamma \curvearrowright X$, $\Gamma \curvearrowright Y$, an *embedding* of $\Gamma \curvearrowright X$ into $\Gamma \curvearrowright Y$ is a continuous Γ-equivariant injective map $\Phi : X \to Y$. If it is also surjective then it is a *topological conjugacy*.

Theorem 3.1. *If* $\mathbb{F}_2 \curvearrowright X$ *embeds into* $\mathbb{F}_2 \curvearrowright Y$ *then for any* Σ*,*

$$h_\Sigma(\mathbb{F}_2 \curvearrowright X) \le h_\Sigma(\mathbb{F}_2 \curvearrowright Y).$$

In particular, topological sofic entropy is a topological conjugacy invariant.

The proof of this is straightforward, see Kerr and Li [2011b, 2016] for details. (The definition of topological sofic entropy is due to Kerr and Li [2011b] which was inspired by my earlier work L. Bowen [2010a]).

3.1 Examples.

3.1.1 A boring example and asymptotic freeness. Suppose that V_n is a single point for all n. In this case, $h_\Sigma(\mathbb{F}_2 \curvearrowright X)$ is simply the logarithm of the number of fixed points of the action. While this is a topological conjugacy invariant, it is not what one usually means by entropy. To avoid this kind of example, we require that the actions $\{\mathbb{F}_2 \curvearrowright V_n\}_n$ are *asymptotically free*. This means that for every nonidentity element $g \in \mathbb{F}_2$

$$\lim_{n \to \infty} |V_n|^{-1} \#\{v \in V_n : g \cdot v = v\} = 0. \tag{1}$$

A countable group Γ admits a sequence $\{\Gamma \curvearrowright V_n\}_{n=1}^{\infty}$ of actions on finite sets satisfying asymptotic freeness if and only if Γ is *residually finite*. We will come back to this point later. From now on, we assume the actions $\{\mathbb{F}_2 \curvearrowright V_n\}_n$ are asymptotically free.

3.1.2 A curious example. Let $X = \mathbb{Z}/2\mathbb{Z}$ and consider the action $\mathbb{F}_2 \curvearrowright X$ defined by $s \cdot x = x + 1$ for $s \in \{a, b\}$. Let $\Sigma = \{\Gamma \curvearrowright V_n\}_{n=1}^{\infty}$ be a sequence of actions on finite sets with the property that the corresponding action graphs $G_n := (V_n, E_n)$ are bipartite. Let $V_n = P_n \sqcup Q_n$ be the bi-partition. Then define $\phi : V_n \to X$ by $\phi(P_n) = \{0\}$ and $\phi(Q_n) = \{1\}$. This map is a (V_n, δ, F)-pseudo-periodic orbit for all δ, F. So $h_\Sigma(\mathbb{F}_2 \curvearrowright X) \geq 0$. It can be shown that in fact any pseudo-periodic orbit must be close to either ϕ or $\phi + 1$ which implies $h_\Sigma(\mathbb{F}_2 \curvearrowright X) = 0$.

Next let $\Sigma' = \{\Gamma \curvearrowright V'_n\}_{n=1}^{\infty}$ be a sequence of actions such that the corresponding action graphs $G'_n := (V'_n, E'_n)$ are far from bi-partite. For example, it is known (and will be explained in Section 7.3) that if the action $\Gamma \curvearrowright V'_n$ is chosen uniformly at random and $|V'_n| \to \infty$ as $n \to \infty$ then with high probability the action graphs will be far from bi-partite in the following sense. For small enough $\delta > 0$ and $F = \{a, b, a^{-1}, b^{-1}\}$ there are no (V'_n, δ, F)-pseudo-periodic orbits. Since $\log(0) = -\infty$ this implies $h_{\Sigma'}(\Gamma \curvearrowright X) = -\infty$.

This example shows (1) entropy depends on the choice of sequence Σ and (2) it is possible for the entropy to be $-\infty$, even for very simple systems.

3.1.3 The Ornstein-Weiss example revisited. In Section 7 below we will sketch a proof that the full shift action $\mathbb{F}_2 \curvearrowright K^{\mathbb{F}_2}$ has topological sofic entropy $\log |K|$ for any finite set K. So the Ornstein-Weiss factor map does indeed increase entropy. How can this happen? The answer is that if $\mathbb{F}_2 \curvearrowright X$ *factors* onto $\mathbb{F}_2 \curvearrowright Y$ (meaning there is a continuous $\mathbb{F}_2$-equivariant surjective map $\Phi : X \to Y$) then, generally speaking, there is no way to "lift" pseudo-periodic orbits of the downstairs action $\mathbb{F}_2 \curvearrowright Y$ up to the source action $\mathbb{F}_2 \curvearrowright X$, even approximately.

Let us see this in detail for the Ornstein-Weiss map. Suppose $\Sigma := \{\mathbb{F}_2 \curvearrowright V_n\}_n$ is a sequence of actions on finite sets and form the action graphs $G_n = (V_n, E_n)$. Given any map $\psi : V_n \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, we can define the *pullback* or *pullback name* of ψ by

$$\widetilde{\psi} : V_n \to (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})^{\mathbb{F}_2}, \quad \widetilde{\psi}(v)(g) = \psi(g^{-1} \cdot v).$$

This map is a (V_n, δ, F)-pseudo-periodic orbit for the shift-action $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})^{\mathbb{F}_2}$ for any $\delta > 0$ and finite $F \subset \mathbb{F}_2$. In particular, this can be used to show that the entropy of $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})^{\mathbb{F}_2}$ is at least $\log 4$.

However most of these maps do not "lift" via the Ornstein-Weiss map. To be precise, define

$$\Phi_n : (\mathbb{Z}/2\mathbb{Z})^{V_n} \to (\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})^{V_n}$$

$$\Phi_n(\psi)(v) = (\psi(v) - \psi(a^{-1} \cdot v), \psi(v) - \psi(b^{-1} \cdot v)).$$

This map is induced from the Ornstein-Weiss map. The point is that while the Ornstein-Weiss map is surjective, its finite approximations Φ_n are from surjective. This is obvious since the domain of Φ_n is exponentially smaller than $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})^{V_n}$. Another argument is homological.

Given $\psi : V_n \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, let $\psi' : E_n \to \mathbb{Z}/2\mathbb{Z}$ be the map defined by

$$\psi(v) = \left(\psi'(v, a^{-1} \cdot v), \psi'(v, b^{-1} \cdot v)\right).$$

Then ψ is in the image of Φ_n if and only if ψ' is a coboundary. So the "reason" the Ornstein-Weiss factor map increases entropy is that the $\mathbb{Z}/2\mathbb{Z}$-homology of the approximating graphs G_n grows exponentially. This observation generalizes: Gaboriau and Seward show in Gaboriau and Seward [2015] that if Γ is any sofic group and k is a finite field then the sofic entropy of $\Gamma \curvearrowright k^\Gamma/k$ is at least $(1 + \beta^1_{(2)}(\Gamma)) \log |k|$ where $\beta^1_{(2)}(\Gamma)$ is the first ℓ^2-Betti number of Γ and k^Γ/k is the quotient of k^Γ by the constant functions.

By contrast, any pseudo-periodic orbit of a $\mathbb{Z}$-action is close to a partial orbit of slightly less length. Partial orbits always lift. This explains why entropy is monotone decreasing for actions of $\mathbb{Z}$.

4 Sofic groups

You might have noticed that we have not used any special properties of free groups. In fact, the definition of topological sofic entropy stated above works for any residually finite group Γ in place of $\mathbb{F}_2$. Recall that Γ is *residually finite* if there exists a decreasing sequence $\Gamma \geq \Gamma_1 \geq \Gamma_2 \cdots$ such that each Γ_n is normal and finite-index in Γ and $\cap_n \Gamma_n = \{e\}$ (this is equivalent to the previous definition in Section 3.1.1). In this case, the sequence of actions $\Gamma \curvearrowright \Gamma/\Gamma_n$ is asymptotically free and so the above definition of topological sofic entropy makes sense.

However, the actions $\Gamma \curvearrowright V_n$ do not really have to be actions! To explain, let $\sigma_n : \Gamma \to \mathrm{sym}(V_n)$ be a sequence of maps from Γ to the symmetric groups $\mathrm{sym}(V_n)$. We do not require these maps to be homomorphisms but we do require that they are *asymptotically multiplicative* in the following sense: for every $g, h \in \Gamma$ we require:

(2) $$\lim_{n\to\infty} |V_n|^{-1} \#\{v \in V_n : \sigma_n(gh)v = \sigma_n(g)\sigma_n(h)v\} = 1.$$

We still require that they are asymptotically free, which means for every nonidentity $g \in \Gamma$

$$\lim_{n\to\infty} |V_n|^{-1}\#\{v \in V_n : \sigma_n(g)\cdot v = v\} = 0. \tag{3}$$

Any sequence $\Sigma = \{\sigma_n\}_{n=1}^{\infty}$ satisfying equations (2, 3) is called a *sofic approximation to* Γ and Γ is called *sofic* if it has a sofic approximation. The definition of topological sofic entropy given above makes sense for arbitrary sofic groups with respect to a sofic approximation Σ once we replace $g \cdot v$ in the definition of a pseudo-periodic orbit with $\sigma_n(g)\cdot v$.

Sofic groups were defined implicitly by Gromov in Gromov [1999]. They were given their name by Benjy Weiss in Weiss [2000]. It is known that amenable groups and residually finite groups are sofic. The class of sofic groups is closed under a large number of group operations including passing to subgroups, direct limits, inverse limits, extensions by amenable groups, direct products, free products with amalgamation over amenable subgroups, graph products and wreathe products Elek and Szabó [2006], Dykema, Kerr, and Pichot [2014], Păunescu [2011], Elek and Szabó [2011], Ciobanu, Holt, and Rees [2014], and Hayes and Sale [2016]. By Malcev's Theorem, finitely generated linear groups are residually finite Malcev [1940]. Since soficity is closed under direct limits, all countable linear groups are sofic. Sofic groups solve special cases of a number of general conjectures including Connes Embedding Conjecture Elek and Szabó [2005], the Determinant Conjecture Elek and Szabó [ibid.], the Algebraic Eigenvalue Conjecture Thom [2008] and Gottschalk's Surjunctivity Conjecture (more on that later on). It is a major open problem whether all countable groups are sofic. Surveys on sofic groups include Pestov [2008], Pestov and Kwiatkowska [2012], and Capraro and Lupini [2015].

5 An application to Gottschalk's Surjunctivity Conjecture

Conjecture 1. *Gottschalk [1973] Let k be a finite set, Γ a countable group and $\Phi : k^\Gamma \to k^\Gamma$ a continuous Γ-equivariant map (where k^Γ is given the product topology). If Φ is injective then it is also surjective.*

This conjecture was proven to be true whenever Γ is sofic by Gromov [1999]. Another proof was given by Weiss [2000] and then another by Kerr and Li [2011b]. Here is a sketch of Kerr-Li's proof: the sofic entropy of $\Gamma \curvearrowright k^\Gamma$ is $\log|k|$. However, the sofic entropy of any proper closed Γ-invariant subset $X \subset k^\Gamma$ is strictly less than $\log|k|$. Since entropy is a topological invariant, this proves the conjecture.

By the way, Gottschalk's conjecture implies Kaplansky's Direct Finiteness Conjecture which states: if k is a finite field, x, y are elements of the group ring $k\Gamma$ and $xy = 1$ then $yx = 1$. To see the connection, observe that x and y induce linear Γ-equivariant maps

Φ_x, Φ_y from k^Γ to itself. If $xy = 1$ then $\Phi_x \Phi_y$ is the identity and so Φ_y is injective. If Gottschalk's conjecture holds then Φ_y must also be surjective and therefore Φ_x is its inverse. So $\Phi_y \Phi_x = 1$, which implies $yx = 1$. In fact this result holds for all fields k, even infinite fields because every field is embeddable into an ultraproduct of finite fields Capraro and Lupini [2015]. When Γ is sofic, direct finiteness of $k\Gamma$ also holds whenever k is a division ring Elek and Szabó [2004] or a unital left Noetherian ring Li and Liang [2016].

6 Measure sofic entropy

Before defining measure sofic entropy, let us revisit topological sofic entropy. The new notation will be useful in treating the measure case. Again, let $\Gamma \curvearrowright X$ be an action by homeomorphisms on a compact metric space (X, d) and fix a sofic approximation $\Sigma = \{\sigma_n\}_{n=1}^{\infty}$ where $\sigma_n : \Gamma \to \mathrm{sym}(V_n)$. Let $\Omega(\sigma_n, \delta, F) \subset X^{V_n}$ be the set of all (σ_n, δ, F)-pseudo-periodic orbits. To be precise, $\phi \in X^{V_n}$ is in $\Omega(\sigma_n, \delta, F)$ if and only if

$$|V_n|^{-1} \sum_{v \in V_n} d(\phi(\sigma_n(g) \cdot v), g \cdot \phi(v)) < \delta$$

for every $g \in F$. So $\Omega(\sigma_n, \delta, F)$ depends implicitly on the action $\Gamma \curvearrowright X$. The topological sofic entropy of $\Gamma \curvearrowright X$ is defined by

$$h_\Sigma(\Gamma \curvearrowright X) := \sup_{\epsilon > 0} \inf_{\delta > 0} \inf_{F \Subset \Gamma} \limsup_{n \to \infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\Omega(\sigma_n, \delta, F), d_\infty^{V_n}).$$

To define measure sofic entropy, we need a few more preliminaries. Let $\mathrm{Prob}(X)$ denote the space of Borel probability measures on X. It is a compact metrizable space with respect to the weak* topology which is defined by: a sequence of measures $\mu_n \in \mathrm{Prob}(X)$ converges to a measure μ if and only if: for every continuous function f on X,

$$\int f \, d\mu_n \to \int f \, d\mu$$

as $n \to \infty$.

Given a map $\phi : V_0 \to X$ (where V_0 is a finite set), the *empirical distribution* of ϕ is the measure

$$P_\phi := |V_0|^{-1} \sum_{v \in V_0} \delta_v \in \mathrm{Prob}(X).$$

Given an open subset $\mathcal{O} \subset \mathrm{Prob}(X)$, let $\Omega(\sigma_n, \delta, F, \mathcal{O})$ be the set of all pseudo-periodic orbits $\phi \in \Omega(\sigma_n, \delta, F)$ such that $P_\phi \in \mathcal{O}$. Then the *sofic entropy of a measure-preserving* *action* $\Gamma \curvearrowright (X, \mu)$ *with respect to* Σ is

$$h_\Sigma(\Gamma \curvearrowright (X, \mu)) := \sup_{\epsilon>0} \inf_{\delta>0} \inf_{F \Subset \Gamma} \inf_{\mathcal{O} \ni \mu} \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\Omega(\sigma_n, \delta, F, \mathcal{O}), d_\infty^{V_n}).$$

This is a measure-conjugacy invariant! For the proof see Kerr and Li [2011b, 2016]. This definition is due to Kerr and Li [2011b] which was inspired by my earlier efforts L. Bowen [2010a]. The proof is straightforward if the measure-conjugacy is a topological conjugacy. The general case is obtained by approximating a measure-conjugacy and its inverse by continuous maps.

There is also a variational principle (due to Kerr and Li [2011b]):

Theorem 6.1 (Variational Principle). $h(\Gamma \curvearrowright X) = \sup_\mu h(\Gamma \curvearrowright (X, \mu))$ *where the sup is over all* Γ*-invariant measures* $\mu \in \mathrm{Prob}(X)$*. If no such measures exist then* $h(\Gamma \curvearrowright X) = -\infty$.

There is also a notion of sofic pressure and a corresponding variational principle Chung [2013]. See also Zhang [2012] for local versions. Moreover, sofic entropy (both topological and measure) agrees with classical entropy whenever Γ is amenable Kerr and Li [2013] and L. Bowen [2012b].

7 Symbolic actions: the topological case

Let $(\mathfrak{A}, d_\mathfrak{A})$ be a compact metric space. In most applications, $\mathfrak{A}$ is either finite or a torus (thought of as a compact abelian group). Let $\mathfrak{A}^\Gamma$ be the space of all functions $x : \Gamma \to \mathfrak{A}$ with the topology of pointwise convergence on finite sets. The group Γ acts on this space by $(g \cdot x)(f) = x(g^{-1} f)$.

Now suppose $X \subset \mathfrak{A}^\Gamma$ is a closed Γ-invariant subspace. There is a more convenient definition of the sofic entropy of $\Gamma \curvearrowright X$ based on maps $\phi : V_n \to \mathfrak{A}$ (instead of maps $\phi : V_n \to X$). Given $\phi : V_n \to \mathfrak{A}$ and $v \in V_n$, define the *pullback of* ϕ by

$$\Pi_v^{\sigma_n}(\phi) \in \mathfrak{A}^\Gamma, \quad \Pi_v^{\sigma_n}(\phi)(g) := \phi(\sigma_n(g)^{-1} v).$$

Given an open neighborhood $\mathcal{U}$ of X in $\mathfrak{A}^\Gamma$, let $\Omega'(\sigma_n, \delta, \mathcal{U})$ be the set of all maps $\phi : V_n \to \mathfrak{A}$ such that

$$|V_n|^{-1} \#\{v \in V_n : \Pi_v^{\sigma_n}(\phi) \in \mathcal{U}\} \geq 1 - \delta.$$

We call such a map a $(\sigma_n, \delta, \mathcal{U})$*-microstate* (this terminology is inspired by Voiculescu's free entropy Voiculescu [1995]). We also call such a map a *microstate* if the parameters are understood or intentionally left ambiguous. Then

$$h_\Sigma(\Gamma \curvearrowright X) = \sup_{\epsilon>0} \inf_{\delta>0} \inf_{\mathcal{U} \supset X} \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\Omega'(\sigma_n, \delta, \mathcal{U}), d_\infty^{V_n})$$

where (by abuse of notation) the metric $d_\infty^{V_n}$ on $\mathcal{A}^{V_n}$ is defined by

$$d_\infty^{V_n}(\phi, \psi) = \max_{v \in V_n} d_{\mathcal{A}}(\phi(v), \psi(v))$$

The reason this works is that given any $\phi' \in \Omega'(\sigma_n, \delta, \mathcal{U})$ we can find a pseudo-periodic orbit $\phi : V_n \to X$ such that $\phi(v)$ is close to $\Pi_v^{\sigma_n}(\phi')$ for most v. Conversely, given a pseudo-periodic orbit ϕ we can define $\phi' : V_n \to \mathcal{A}$ by $\phi'(v) =$ the projection of ϕ to the identity-coordinate. Then $\Pi_v^{\sigma_n}(\phi')$ will be close to $\phi(v)$ for most v. The measure case of this statement is proven in Austin [2016a]. The topological case follows from the variational principle.

If $\mathcal{A}$ is finite then we can simplify further by setting $\epsilon = 0$. To be precise,

$$h_\Sigma(\Gamma \curvearrowright X) = \inf_{\delta > 0} \inf_{\mathcal{U} \supset X} \limsup_{n \to \infty} \frac{1}{n} \log \#\Omega'(\sigma_n, \delta, \mathcal{U}).$$

The reason this works is that if $\phi, \psi \in \mathcal{A}^{V_n}$ are any distinct elements then $d_\infty^{V_n}(\phi, \psi) \geq c > 0$ where c is the minimum distance between distinct elements of $\mathcal{A}$.

Using this definition of sofic entropy, it is easy to check that $h_\Sigma(\Gamma \curvearrowright \mathcal{A}^\Gamma) = \log |\mathcal{A}|$. Indeed, this is true because $\Omega'(\sigma_n, \delta, \mathcal{U}) = \mathcal{A}^{V_n}$.

7.1 Symbolic actions: the measure case. Suppose $\mu \in \mathrm{Prob}(\mathcal{A}^\Gamma)$ is a probability measure preserved under the action. For any open neighborhood $\mathcal{O} \ni \mu$, a $(\sigma_n, \mathcal{O})$-*microstate* for μ is a map $\phi : V_n \to \mathcal{A}$ such that its *empirical measure*, defined by

$$P_\phi := |V_n|^{-1} \sum_{v \in V_n} \delta_{\Pi_v^{\sigma_n}(\phi)}$$

is contained in $\mathcal{O}$. Let $\Omega'(\sigma_n, \mathcal{O})$ be the set of all $(\sigma_n, \mathcal{O})$-microstates for μ. Then

$$h_\Sigma(\Gamma \curvearrowright (\mathcal{A}^\Gamma, \mu)) = \sup_{\epsilon > 0} \inf_{\mathcal{O} \ni \mu} \limsup_{n \to \infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\Omega'(\sigma_n, \mathcal{O}), d_\infty^{V_n}).$$

This approach is proven in Austin [ibid.].

As in the topological case, we can simplify further if $\mathcal{A}$ is finite by setting $\epsilon = 0$ to obtain

$$h_\Sigma(\Gamma \curvearrowright (\mathcal{A}^\Gamma, \mu)) = \inf_{\mathcal{O} \ni \mu} \limsup_{n \to \infty} \frac{1}{n} \log \#\Omega'(\sigma_n, \mathcal{O}).$$

The above is essentially the same as my original definition of sofic entropy in L. Bowen [2010a].

7.2 Bernoulli shifts. Let κ be a Borel probability measure on $\mathcal{A}$. The *Shannon entropy* of κ is

$$H(\kappa) := -\sum_{a \in \mathcal{A}} \kappa(\{a\}) \log \kappa(\{a\})$$

if κ is supported on a countable set (and $0 \log 0 := 0$). If κ is not supported on a countable set then $H(\kappa) := +\infty$.

Let κ^Γ be the product measure on $\mathcal{A}^\Gamma$. The action $\Gamma \curvearrowright (\mathcal{A}^\Gamma, \kappa^\Gamma)$ is called the *Bernoulli shift over* Γ *with base space* $(\mathcal{A}, \kappa)$.

Theorem 7.1. *For any sofic approximation* Σ,

$$h_\Sigma(\Gamma \curvearrowright (\mathcal{A}^\Gamma, \kappa^\Gamma)) = H(\kappa).$$

The finite entropy case of this was obtained in L. Bowen [ibid.] and the infinite entropy case is in Kerr and Li [2011a]. Here is a sketch in the special case in which $\mathcal{A}$ is finite.
The lower bound. Let ϕ be a random map $V_n \to \mathcal{A}$ with distribution κ^{V_n}. A second moment argument shows that, for any open neighborhood $\mathcal{O}$ of κ, if n is sufficiently large then with high probability ϕ is a $(\sigma_n, \mathcal{O})$-microstate for κ^Γ. By the law of large numbers or the Shannon-McMillan Theorem, any subset $S \subset \mathcal{A}^{V_n}$ with measure close to 1 has cardinality at least $e^{|V_n|H(\kappa)-o(|V_n|)}$. This proves the lower bound.
The upper bound. Suppose that $\mathcal{O}$ consists of all measures $\mu \in \mathrm{Prob}(\mathcal{A}^\Gamma)$ such that if $P : \mathcal{A}^\Gamma \to \mathcal{A}$ denotes projection onto the identity coordinate then $\|P_*\mu - \kappa\|_{TV} < \epsilon$. Then the number of $(\sigma_n, \mathcal{O})$-microstates is approximately the multinomial

$$|V_n|! \left(\prod_{a \in \mathcal{A}} \lfloor \kappa(\{a\}) V_n \rfloor ! \right)^{-1}$$

which, by Stirling's formula, is approximately $e^{|V_n|H(\kappa)+o(|V_n|)}$. This proves the upper bound.

7.3 The f-invariant and RS-entropy. For this section, consider the special case in which $\Gamma = \mathbb{F}_r = \langle s_1, \ldots, s_r \rangle$ is the rank r free group. Instead of fixing a sofic approximation, set $V_n := \{1, \ldots, n\}$ and let $\sigma_n : \mathbb{F}_r \to \mathrm{sym}(V_n)$ be a uniformly random homomorphism. The *f-invariant or RS-entropy* of a measure-preserving action $\mathbb{F}_r \curvearrowright (X, \mu)$ is defined in the same way as sofic entropy except that ones takes an expected value before the logarithm:

$$f(\mu) := h^{RS}(\mu) := \sup_{\epsilon>0} \inf_{\delta>0} \inf_{F \Subset \Gamma} \inf_{\mathcal{O} \ni \mu} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\sigma_n}\big[\mathrm{Sep}_\epsilon(\Omega(\sigma_n, \delta, F, \mathcal{O}), d_\infty^{V_n})\big].$$

The RS stands for "replica-symmetric" to emphasize the analogy with the corresponding notions in the literature in statistical physics and theoretical computer science Dembo and Montanari [2010]. In the special case in which μ is a shift invariant measure on $\mathcal{A}^{\mathbb{F}_r}$ and $\mathcal{A}$ is finite, the definition reduces to

$$f(\mu) = h^{RS}(\mu) = \inf_{\mathcal{O} \ni \mu} \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}_{\sigma_n}[\#\Omega'(\sigma_n, \mathcal{O})].$$

Alternatively, let $\mathcal{P}$ be a countable measurable partition of X. Its Shannon entropy is defined by

$$H_\mu(\mathcal{P}) := -\sum_{P \in \mathcal{P}} \mu(P) \log \mu(P). \tag{4}$$

Define

$$F_\mu(\mathcal{P}) := -(2r-1)H_\mu(\mathcal{P}) + \sum_{i=1}^{r} H_\mu(\mathcal{P} \vee s_i \mathcal{P})$$

where $\mathcal{P} \vee s_i \mathcal{P}$ is the smallest partition refining both $\mathcal{P}$ and $s_i \mathcal{P}$. For any finite $W \subset \mathbb{F}_r$, let $\mathcal{P}^W$ be the smallest partition containing $w\mathcal{P}$ for all $w \in W$.

Theorem 7.2. *L. Bowen [2010c] $f(\mu) = \inf_{R>0} F_\mu(\mathcal{P}^{B(R)})$ where $B(R) \subset \mathbb{F}_r$ is the ball of radius $R > 0$ with respect to the word metric and $\mathcal{P}$ is any generating partition for the action such that $H_\mu(\mathcal{P}) < \infty$. (A partition $\mathcal{P}$ is* generating *if the smallest complete Γ-invariant σ-sub-algebra containing $\mathcal{P}$ consists of all measurable sets).*

The theorem above was taken as the definition of f in L. P. Bowen [2010] where the f-invariant was first proven to be a measure-conjugacy invariant without using model spaces. Conditional on the existence of a finite generating partition, the f-invariant is additive under direct products, it satisfies an ergodic decomposition formula Seward [2014b], a subgroup formula Seward [2014a], an Abraham-Rokhlin formula L. Bowen [2010b] and a (restricted) Yuzvinskii addition formula L. Bowen and Gutman [2014]. Sofic entropy does not, in general, satisfy such formulas L. Bowen [2017a].

Example 1. The f-invariant of any action Γ on a finite set X is $-(r-1)\log|X|$. So $\mathbb{E}_{\sigma_n}[\#\Omega'(\sigma_n, \mathcal{O})] \approx |X|^{-(r-1)|V_n|}$ (for $\mathcal{O}$ small and n large). Since $\#\Omega'(\sigma_n, \mathcal{O}) \in \{0, 1, 2, \ldots\}$, with high probability $\Omega'(\sigma_n, \mathcal{O})$ is empty (if $\mathcal{O}$ is small enough, $r \geq 2$ and $|X| \geq 2$). This explains why, as claimed in Section 3.1.2, if $\Gamma \curvearrowright V_n'$ is uniformly random then with high probability the graphs $G_n' = (V_n', E_n')$ are far from bipartite in the sense that there are no pseudo-periodic orbits (or microstates) for the action $\Gamma \curvearrowright \mathbb{Z}/2\mathbb{Z}$ where each generator acts nontrivially.

The paper L. Bowen [2010b] defines Markov processes over the free group and shows that, for such processes, $F_\mu(\mathcal{P}) = f(\mu)$. In particular, it is very easy to compute. It can be shown that some mixing Markov processes (e.g., the Ising process with small transition probability and free boundary conditions) have negative f-invariant. These cannot be isomorphic to Bernoulli shifts. By contrast, all mixing Markov processes over the integers are isomorphic to Bernoulli shifts. It is a major open problem to classify (mixing) Markov processes over a free group up to measure-conjugacy.

8 Classification of Bernoulli shifts

Theorem 7.1 shows that if Γ is sofic then Bernoulli shifts with different base space entropies are not measurably conjugate. Surprisingly, the converse is true even without soficity:

Theorem 8.1. *If Γ is any countably infinite group and $(\mathcal{A}, \kappa)$, $(\mathcal{B}, \lambda)$ are two probability spaces with equal Shannon entropies $H(\kappa) = H(\lambda)$ then $\Gamma \curvearrowright (\mathcal{A}, \kappa)^\Gamma$ is measurably conjugate to $\Gamma \curvearrowright (\mathcal{B}, \lambda)^\Gamma$.*

Remark 1. The special case in which $\Gamma = \mathbb{Z}$ is Ornstein's famous theorem of 1970 D. Ornstein [1970a]. Stepin observed that if Γ is any countable group containing a copy of $\mathbb{Z}$ then the theorem above holds for Γ because one can build an isomorphism for Γ-actions from an isomorphism for $\mathbb{Z}$-actions, coset-by-coset Stepin [1975]. Ornstein-Weiss extended Ornstein's Theorem to all countably infinite amenable groups through the technology of quasi-tilings D. S. Ornstein and Weiss [1980]. I showed in L. Bowen [2012a] that the above theorem is true whenever the supports of κ and λ each contain more than 2 elements. The proof is essentially a "measurable version" of Stepin's trick. The last remaining case (when say $|\mathcal{A}| = 2$) has been handled recently in soon-to-be-published work of Brandon Seward.

9 Bernoulli factors

Theorem 9.1. *Let Γ be any non-amenable group. Then every nontrivial Bernoulli shift over Γ factors onto every other nontrivial Bernoulli shift.*

This theorem is obtained in L. Bowen [2017b]. The special case of the free group $\mathbb{F}_2$ was handled in L. Bowen [2011b] using the Ornstein-Weiss map and Sinai's Factor Theorem (for actions of $\mathbb{Z}$). It immediately follows for any group containing a copy of $\mathbb{F}_2$ since we can build the factor map coset-by-coset. In Ball [2005], it was shown that if Γ is any non-amenable group then there is some Bernoulli shift over a finite base space that factors onto all other Bernoulli shifts. The argument used a rudimentary form of the

Gaboriau-Lyons Theorem (before that theorem existed) which states that if Γ is any non-amenable group then there is some Bernoulli shift (over a finite base space) $\Gamma \curvearrowright (\mathfrak{K}, \kappa)^\Gamma$ and an ergodic essentially free action of the free group $\mathbb{F}_2 \curvearrowright \mathfrak{K}^\Gamma$ such that the orbits of the free group action are contained in the Γ-orbits Gaboriau and Lyons [2009]. We can then view this free group action as being essentially like having a free subgroup of Γ and build the factor map coset-by-coset as before. The main new result of L. Bowen [2017b] is that the Gaboriau-Lyons Theorem holds for arbitrary Bernoulli shifts. Theorem 9.1 then follows from an argument similar to Ball [2005].

10 Rokhlin entropy

A measurable partition $\mathcal{P}$ of a measure space (X, μ) is *generating* for an action $\Gamma \curvearrowright (X, \mu)$ if the smallest Γ-invariant sigma-algebra containing $\mathcal{P}$ consists of all measurable sets (modulo sets of measure zero). The Rokhlin entropy of an ergodic action $\Gamma \curvearrowright (X, \mu)$ is the infimum of the Shannon entropies of generating partitions (the Shannon entropy is defined by (4)). In the special case that $\Gamma = \mathbb{Z}$, Rokhlin proved that this agrees with Kolmogorov-Sinai entropy Rohlin [1967]. A modern proof of this, that holds for all amenable Γ, is in Seward and Tucker-Drob [2016].

Rokhlin entropy is an upper bound to sofic entropy. It is unknown whether they are equal, conditioned on the sofic entropy not being minus infinity. For example, this is unknown even for principal algebraic actions (see Section 11.0.2). Unfortunately, sofic entropy is the only known lower bound for Rokhlin entropy; hence we do not even know how to compute the Rokhlin entropy of Bernoulli shifts, except when the group is sofic, in which case the Rokhlin entropy equals the Shannon entropy of the base. Indeed, it is shown in Seward [2015a] that if the Rokhlin entropy of Bernoulli shifts is positive (for all groups) then Gottschalk's conjecture holds for all groups. It is also known that Rokhlin entropy equals sofic entropy for Gibbs measures satisfying a strong spatial mixing condition Alpeev [2017] and Austin and Podder [2017].

Rokhlin entropy has mainly been used as a hypothesis rather than a conclusion. There are two main theorems of this form; they generalize Krieger's Generator Theorem and Sinai's Factor Theorem:

Theorem 10.1. *Seward [2014c] If $\Gamma \curvearrowright (X, \mu)$ is ergodic and has Rokhlin entropy $< \log(n)$ for some integer $n > 1$ then there exists a generating partition for the action with n parts.*

This Theorem is the simplest version of a large variety of far more refined results contained in Seward [2014c, 2015a] and Alpeev and Seward [2016].

Theorem 10.2. *Seward [2015b] If $\Gamma \curvearrowright (X, \mu)$ is ergodic and has positive Rokhlin entropy then it factors onto a Bernoulli shift.*

11 Algebraic actions

An *algebraic action* is an action of a countable group Γ on a compact group X by group-automorphisms. The main problem is to relate algebraic or analytic properties of the image of Γ in $\mathrm{Aut}(X)$ with purely dynamical properties. The study of single automorphisms (that is, $\Gamma = \mathbb{Z}$) goes back at least to Juzvinskiĭ [1965a,b] and R. Bowen [1971]. The special case $\Gamma = \mathbb{Z}^d$ was studied intensively in the 80's and 90's Schmidt [1995]. Here we will highlight a few recent achievements and open problems extending classical results to the realm of sofic group actions.

11.0.1 Topological versus measure entropy. Yuzvinskii and R. Bowen showed that, when $\Gamma = \mathbb{Z}$, the topological and measure-entropy of an algebraic action $\Gamma \curvearrowright X$ agree (where the measure on X is Haar measure) Juzvinskiĭ [1965b] and R. Bowen [1971]. This was extended to amenable Γ by Deninger [2006]. It was an open problem since 2011 whether this result could be extended to sofic groups. There were computations of entropy showing that it was true in a number of special cases Kerr and Li [2011b], L. Bowen [2011a], L. Bowen and Li [2012], and Hayes [2016b] but these all proceeded by computing the topological entropy and the measure entropy separately in terms of analytic data and then showing their equality. So it is astonishing that just recently Ben Hayes proved under a mild hypothesis on the actions that the topological entropy agrees with the measure entropy Hayes [2016a]. The proof uses Austin's lde-sofic-entropy Austin [2016a].

11.0.2 Principal algebraic actions. Let f be an element of the integer group ring $\mathbb{Z}\Gamma$ and consider the principal ideal $\mathbb{Z}\Gamma f \subset \mathbb{Z}\Gamma$ generated by f. Then $\mathbb{Z}\Gamma/\mathbb{Z}\Gamma f$ is a countable abelian group that Γ acts on by automorphisms (namely the action is $g(x+\mathbb{Z}\Gamma f) := gx + \mathbb{Z}\Gamma f$). Let $X_f := \mathrm{Hom}(\mathbb{Z}\Gamma/\mathbb{Z}\Gamma f, \mathbb{R}/\mathbb{Z})$ be the Pontryagin dual. This is a compact abelian group under pointwise addition. Moreover, the action of Γ on $\mathbb{Z}\Gamma/\mathbb{Z}\Gamma f$ induces an action of Γ on X_f by automorphisms.

In the special case in which $\mathbb{Z} = \Gamma$, we can write f as $f = \sum_{i=-m}^{n} c_i x^i$ for some coefficients $c_i \in \mathbb{Z}$ and some $0 \le m, n$. After multiplying f by some x^k we may assume that $m = 0$ and $c_0 \neq 0$. This change does not affect the dynamics. So we can think of f as a polynomial. Yuzvinskii and R. Bowen showed that the entropy of $\mathbb{Z}$ acting on X_f is the log-Mahler measure of f. This result was extended to $\mathbb{Z}^d$ by Lind, Schmidt, and Ward [1990]. Chris Deninger observed that the Fuglede-Kadison determinant generalizes Mahler measure to non-abelian Γ and conjectured that, when Γ is amenable, $h(\Gamma \curvearrowright X_f)$ is the log of the Fuglede-Kadison determinant of f Deninger [2006]. Special cases were handled in Deninger [2006] and Deninger and Schmidt [2007] before the general case was completed by Hanfeng Li in Li [2012].

The first result in the setting of a non-amenable acting group Γ was the case of expansive principal algebraic actions of residually finite groups L. Bowen [2011a]. This was extended in L. Bowen and Li [2012] to some non-expansive actions (harmonic mod 1 points). Then in a stunning breakthrough Hayes proved that for an arbitrary sofic Γ and arbitrary $f \in \mathbb{Z}\Gamma$ either f is not injective as a convolution operator on $\ell^2(\Gamma)$ (in which case the sofic entropy is infinite) or it is injective and the sofic entropy equals the log of the Fuglede-Kadison determinant of f Hayes [2016b].

11.0.3 Yuzvinskii's addition formula. Suppose that $N \lhd X$ is a closed Γ-invariant normal subgroup. We say the *addition formula* holds for $(\Gamma \curvearrowright X, N)$ if the entropy of $\Gamma \curvearrowright X$ equals the sum of the entropy of $\Gamma \curvearrowright N$ with the entropy of $\Gamma \curvearrowright X/N$. In the special case in which $\Gamma = \mathbb{Z}$, this result is due to Yuzvinskii Juzvinskiĭ [1965a,b]. It was extended to $\mathbb{Z}^d$ in Lind, Schmidt, and Ward [1990] and to arbitrary amenable groups in Li [2012] (and independently in unpublished work of Lind-Schmidt). It is an important structural result which when combined with the principal algebraic case yields a general procedure for computing entropy of algebraic actions satisfying mild hypotheses Schmidt [1995]. Using it, Li and Thom relate entropy to L^2-torsion Li and Thom [2014] thereby obtaining new results about L^2-torsion and algebraic dynamics.

The Ornstein-Weiss example shows that addition formulas fail for sofic entropy in general. Indeed, it has been shown in Bartholdi and Kielak [2017] that for any non-amenable Γ and any field k, there exists an embedding $(k\Gamma)^n \to (k\Gamma)^{n-1}$ for some $n \in \mathbb{N}$ (as $k\Gamma$-modules). Taking the dual and setting k equal to a finite field gives a contradiction to the addition formula.

However, the f-invariant satisfies the addition formula when X is totally disconnected and satisfies some technical hypothesis L. Bowen and Gutman [2014]. The reason this does not contradict the previous paragraph is that the f-invariant, unlike sofic entropy, can take finite negative values. It is an open problem whether the addition formula holds for the f-invariant in general.

11.0.4 Pinsker algebra. In recent work, Hayes shows that under mild hypotheses, the outer Pinsker algebra of an algebraic action of a sofic group is algebraic; that is, it comes from an invariant closed normal subgroup Hayes [2016c]. To explain a little more, the outer sofic entropy of a factor is the growth rate of the number of microstates that lift to the source action. The outer Pinsker algebra is the maximal σ-sub-algebra such that the corresponding factor has zero outer sofic entropy. It follows that, in order, to prove an algebraic action has completely positive entropy (CPE), it is sufficiently to check all of its algebraic factors.

In Schmidt [1995] it is shown that if $\Gamma = \mathbb{Z}^d$ then all CPE algebraic actions are Bernoulli. Could this be true more generally? Even for amenable groups, this problem is open.

12 Geometry of model spaces

Recall from Section 7.1 that

$$h_\Sigma(\Gamma \curvearrowright (\mathcal{A}^\Gamma, \mu)) = \sup_{\epsilon>0} \inf_{\mathcal{O} \ni \mu} \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{Sep}_\epsilon(\Omega'(\sigma_n, \mathcal{O}), d_\infty^{V_n}).$$

The spaces $\Omega'(\sigma_n, \mathcal{O})$ are called *model spaces*. We consider them with the metrics $d_1^{V_n}$ defined by

$$d_1^{V_n}(\phi, \psi) := \frac{1}{|V_n|} \sum_{v \in V_n} d(\phi(v), \psi(v)).$$

In the special case in which $\mathcal{A}$ is finite and $d(x, y) \in \{0, 1\}$ for all $x, y \in \mathcal{A}$, $d_1^{V_n}$ is the normalized Hamming metric. The asymptotic *geometry* of these model spaces can be used to define new invariants.

In Austin [2016b] Tim Austin introduced a notion of asymptotic coarse connectedness for model spaces that depends on a choice of sofic approximation Σ. He calls this notion *connected model spaces rel* Σ. For fixed Σ, it is a measure-conjugacy invariant. He shows that Bernoulli shifts have connected model spaces rel Σ (for any Σ). On the other hand, if Γ is residually finite, has property (T) and $\mathbb{T}$ is the 1-torus then there is a sofic approximation Σ such that the model spaces for the action $\Gamma \curvearrowright \mathbb{T}^\Gamma/\mathbb{T}$ are not asymptotically coarsely connected. In particular, this action is a factor of a Bernoulli shift that is not isomorphic to a Bernoulli shift (by contrast, all factors of Bernoulli shifts of $\mathbb{Z}$-actions are Bernoulli D. Ornstein [1970b]). This was known earlier from work of Popa-Sasyk Popa and Sasyk [2007] and the example is the same, but the proof is different since it goes through this new measure-conjugacy invariant.

In forthcoming work, I will generalize model-connectedness to asymptotic coarse homological invariants. These new invariants will be applied to show that there are Markov processes over the free group that do not have the Weak Pinsker Property. By contrast, Austin recently proved that all processes over $\mathbb{Z}$ have the Weak Pinsker Property Austin [2017].

In Austin [2016a] Austin defines a notion of convergence for sequences of probability measures μ_n on $\mathcal{A}^{V_n}$ (with respect to Σ) and uses this to define new notions of sofic entropy. One of these notions was called "doubly-quenched sofic entropy" in Austin [ibid.] but has now been renamed to "locally doubly empirical entropy" to avoid conflict with physics notions. This particular version of sofic entropy is additive under direct products.

Under very mild hypotheses, it agrees with the power-stabilized entropy which is defined by

$$h^{PS}_{\Sigma}(\Gamma \curvearrowright (X,\mu)) = \lim_{n\to\infty} \frac{1}{n} h^{PS}_{\Sigma}(\Gamma \curvearrowright (X,\mu)^n)$$

where $\Gamma \curvearrowright X^n$ acts diagonally $g(x_1,\ldots,x_n) := (gx_1,\ldots,gx_n)$. Thus, it seems to be the 'right' version of sofic entropy if one requires additivity under direct products.

Acknowledgments. Thanks to Tim Austin and Andrei Alpeev for helpful comments.

References

Andrei Alpeev (2017). "Random ordering formula for sofic and Rokhlin entropy of Gibbs measures". *preprint* (cit. on p. 1878).

Andrei Alpeev and Brandon Seward (2016). "Krieger's finite generator theorem for ergodic actions of countable groups III". *preprint* (cit. on p. 1878).

Tim Austin (2016a). "Additivity properties of sofic entropy and measures on model spaces". *Forum Math. Sigma* 4, e25, 79. MR: 3542515 (cit. on pp. 1874, 1879, 1881).

– (2016b). "The Geometry of Model Spaces for Probability-Preserving Actions of Sofic Groups". *Anal. Geom. Metr. Spaces* 4, Art. 6. MR: 3543677 (cit. on p. 1881).

– (2017). "Measure concentration and the weak Pinsker property". *preprint* (cit. on p. 1881).

Tim Austin and Moumanti Podder (2017). "Gibbs measures over locally tree-like graphs and percolative entropy over infinite regular trees". *preprint* (cit. on p. 1878).

Karen Ball (2005). "Factors of independent and identically distributed processes with non-amenable group actions". *Ergodic Theory Dynam. Systems* 25.3, pp. 711–730. MR: 2142942 (cit. on pp. 1877, 1878).

L. Bartholdi and D. Kielak (2017). "Amenability of groups is characterized by Myhill's Theorem". arXiv: 1605.09133 (cit. on p. 1880).

Lewis Bowen (2010a). "Measure conjugacy invariants for actions of countable sofic groups". *J. Amer. Math. Soc.* 23.1, pp. 217–245. MR: 2552252 (cit. on pp. 1868, 1873–1875).

– (2010b). "Non-abelian free group actions: Markov processes, the Abramov-Rohlin formula and Yuzvinskii's formula". *Ergodic Theory Dynam. Systems* 30.6, pp. 1629–1663. MR: 2736889 (cit. on pp. 1876, 1877).

– (2010c). "The ergodic theory of free group actions: entropy and the f-invariant". *Groups Geom. Dyn.* 4.3, pp. 419–432. MR: 2653969 (cit. on p. 1876).

– (2011a). "Entropy for expansive algebraic actions of residually finite groups". *Ergodic Theory Dynam. Systems* 31.3, pp. 703–718. MR: 2794944 (cit. on pp. 1879, 1880).

– (2011b). "Weak isomorphisms between Bernoulli shifts". *Israel J. Math.* 183, pp. 93–102. MR: 2811154 (cit. on p. 1877).

– (2012a). "Every countably infinite group is almost Ornstein". In: *Dynamical systems and group actions*. Vol. 567. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 67–78. MR: 2931910 (cit. on p. 1877).

– (2012b). "Sofic entropy and amenable groups". *Ergodic Theory Dynam. Systems* 32.2, pp. 427–466. MR: 2901354 (cit. on p. 1873).

– (2017a). "Examples in the entropy theory of countable group actions". *submitted* (cit. on pp. 1865, 1876).

– (2017b). "Finitary random interlacements and the Gaboriau-Lyons problem". *submitted* (cit. on pp. 1877, 1878).

Lewis Phylip Bowen (2010). "A measure-conjugacy invariant for free group actions". *Ann. of Math. (2)* 171.2, pp. 1387–1400. MR: 2630067 (cit. on p. 1876).

Lewis Bowen and Yonatan Gutman (2014). "A Juzvinskii addition theorem for finitely generated free group actions". *Ergodic Theory Dynam. Systems* 34.1, pp. 95–109. MR: 3163025 (cit. on pp. 1876, 1880).

Lewis Bowen and Hanfeng Li (2012). "Harmonic models and spanning forests of residually finite groups". *J. Funct. Anal.* 263.7, pp. 1769–1808. MR: 2956925 (cit. on pp. 1879, 1880).

Rufus Bowen (1971). "Entropy for group endomorphisms and homogeneous spaces". *Trans. Amer. Math. Soc.* 153, pp. 401–414. MR: 0274707 (cit. on pp. 1866, 1879).

Valerio Capraro and Martino Lupini (2015). *Introduction to Sofic and hyperlinear groups and Connes' embedding conjecture*. Vol. 2136. Lecture Notes in Mathematics. With an appendix by Vladimir Pestov. Springer, Cham, pp. viii+151. MR: 3408561 (cit. on pp. 1871, 1872).

Nhan-Phu Chung (2013). "Topological pressure and the variational principle for actions of sofic groups". *Ergodic Theory Dynam. Systems* 33.5, pp. 1363–1390. MR: 3103087 (cit. on p. 1873).

Laura Ciobanu, Derek F. Holt, and Sarah Rees (2014). "Sofic groups: graph products and graphs of groups". *Pacific J. Math.* 271.1, pp. 53–64. MR: 3259760 (cit. on p. 1871).

Amir Dembo and Andrea Montanari (2010). "Gibbs measures and phase transitions on sparse random graphs". *Braz. J. Probab. Stat.* 24.2, pp. 137–211. MR: 2643563 (cit. on p. 1876).

Christopher Deninger (2006). "Fuglede-Kadison determinants and entropy for actions of discrete amenable groups". *J. Amer. Math. Soc.* 19.3, 737–758 (electronic). MR: 2220105 (cit. on p. 1879).

Christopher Deninger and Klaus Schmidt (2007). "Expansive algebraic actions of discrete residually finite amenable groups and their entropy". *Ergodic Theory Dynam. Systems* 27.3, pp. 769–786. MR: 2322178 (cit. on p. 1879).

Ken Dykema, David Kerr, and Mikaël Pichot (2014). "Sofic dimension for discrete measured groupoids". *Trans. Amer. Math. Soc.* 366.2, pp. 707–748. MR: 3130315 (cit. on p. 1871).

Gábor Elek and Endre Szabó (2004). "Sofic groups and direct finiteness". *J. Algebra* 280.2, pp. 426–434. MR: 2089244 (cit. on p. 1872).

– (2005). "Hyperlinearity, essentially free actions and L^2-invariants. The sofic property". *Math. Ann.* 332.2, pp. 421–441. MR: 2178069 (cit. on p. 1871).

– (2006). "On sofic groups". *J. Group Theory* 9.2, pp. 161–171. MR: 2220572 (cit. on p. 1871).

– (2011). "Sofic representations of amenable groups". *Proc. Amer. Math. Soc.* 139.12, pp. 4285–4291. MR: 2823074 (cit. on p. 1871).

Damien Gaboriau (2016). "Sofic entropy, after Lewis Bowen, David Kerr and Hanfeng Li". arXiv: 1607.06454 (cit. on p. 1865).

Damien Gaboriau and Russell Lyons (2009). "A measurable-group-theoretic solution to von Neumann's problem". *Invent. Math.* 177.3, pp. 533–540. MR: 2534099 (cit. on p. 1878).

Damien Gaboriau and Brandon Seward (2015). "Cost, ℓ^2-Betti numbers, and the sofic entropy of some algebraic actions". *preprint* (cit. on p. 1870).

Walter Gottschalk (1973). "Some general dynamical notions", 120–125. Lecture Notes in Math., Vol. 318. MR: 0407821 (cit. on p. 1871).

M. Gromov (1999). "Endomorphisms of symbolic algebraic varieties". *J. Eur. Math. Soc. (JEMS)* 1.2, pp. 109–197. MR: 1694588 (cit. on p. 1871).

Ben Hayes (2016a). "Doubly quenched convergence and the entropy of algebriac actions of sofic groups". *submitted* (cit. on p. 1879).

– (2016b). "Fuglede-Kadison determinants and sofic entropy". *Geom. Funct. Anal.* 26.2, pp. 520–606. MR: 3513879 (cit. on pp. 1879, 1880).

– (2016c). "Relative entropy and the Pinsker product formula for sofic groups". *submitted* (cit. on p. 1880).

Ben Hayes and Andrew Sale (2016). "The Wreath Product of Two Sofic Groups is Sofic". *submitted* (cit. on p. 1871).

S. A. Juzvinskiĭ (1965a). "Metric properties of automorphisms of locally compact commutative groups". *Sibirsk. Mat. Ž.* 6, pp. 244–247. MR: 0201560 (cit. on pp. 1879, 1880).

– (1965b). "Metric properties of the endomorphisms of compact groups". *Izv. Akad. Nauk SSSR Ser. Mat.* 29, pp. 1295–1328. MR: 0194588 (cit. on pp. 1879, 1880).

A. Katok (1980). "Lyapunov exponents, entropy and periodic orbits for diffeomorphisms". *Inst. Hautes Études Sci. Publ. Math.* 51, pp. 137–173. MR: 573822 (cit. on p. 1867).

David Kerr and Hanfeng Li (2011a). "Bernoulli actions and infinite entropy". *Groups Geom. Dyn.* 5.3, pp. 663–672. MR: 2813530 (cit. on p. 1875).

– (2011b). "Entropy and the variational principle for actions of sofic groups". *Invent. Math.* 186.3, pp. 501–558. MR: 2854085 (cit. on pp. 1867, 1868, 1871, 1873, 1879).

– (2013). "Soficity, amenability, and dynamical entropy". *Amer. J. Math.* 135.3, pp. 721–761. MR: 3068400 (cit. on p. 1873).

– (2016). *Ergodic theory*. Springer Monographs in Mathematics. Independence and dichotomies. Springer, Cham, pp. xxxiv+431. MR: 3616077 (cit. on pp. 1865, 1868, 1873).

Hanfeng Li (2012). "Compact group automorphisms, addition formulas and Fuglede-Kadison determinants". *Ann. of Math. (2)* 176.1, pp. 303–347. MR: 2925385 (cit. on pp. 1879, 1880).

Hanfeng Li and Bingbing Liang (2016). "Sofic mean length". *preprint* (cit. on p. 1872).

Hanfeng Li and Andreas Thom (2014). "Entropy, determinants, and L^2-torsion". *J. Amer. Math. Soc.* 27.1, pp. 239–292. MR: 3110799 (cit. on p. 1880).

Douglas Lind, Klaus Schmidt, and Tom Ward (1990). "Mahler measure and entropy for commuting automorphisms of compact groups". *Invent. Math.* 101.3, pp. 593–629. MR: 1062797 (cit. on pp. 1879, 1880).

A. Malcev (1940). "On isomorphic matrix representations of infinite groups". *Rec. Math. [Mat. Sbornik] N.S.* 8 (50), pp. 405–422. MR: 0003420 (cit. on p. 1871).

Donald Ornstein (1970a). "Bernoulli shifts with the same entropy are isomorphic". *Advances in Math.* 4, 337–352 (1970). MR: 0257322 (cit. on p. 1877).

– (1970b). "Factors of Bernoulli shifts are Bernoulli shifts". *Advances in Math.* 5, 349–364 (1970). MR: 0274717 (cit. on p. 1881).

Donald S. Ornstein and Benjamin Weiss (1980). "Ergodic theory of amenable group actions. I. The Rohlin lemma". *Bull. Amer. Math. Soc. (N.S.)* 2.1, pp. 161–164. MR: 551753 (cit. on p. 1877).

– (1987). "Entropy and isomorphism theorems for actions of amenable groups". *J. Analyse Math.* 48, pp. 1–141. MR: 910005 (cit. on p. 1866).

Liviu Păunescu (2011). "On sofic actions and equivalence relations". *J. Funct. Anal.* 261.9, pp. 2461–2485. MR: 2826401 (cit. on p. 1871).

Vladimir G. Pestov (2008). "Hyperlinear and sofic groups: a brief guide". *Bull. Symbolic Logic* 14.4, pp. 449–480. MR: 2460675 (cit. on p. 1871).

Vladimir G Pestov and Alexsandra Kwiatkowska (2012). "An introduction to hyperlinear and sofic groups". arXiv: 0911.4266 (cit. on p. 1871).

Karl Petersen (1989). *Ergodic theory*. Vol. 2. Cambridge Studies in Advanced Mathematics. Corrected reprint of the 1983 original. Cambridge University Press, Cambridge, pp. xii+329. MR: 1073173 (cit. on p. 1865).

Sorin Popa and Roman Sasyk (2007). "On the cohomology of Bernoulli actions". *Ergodic Theory Dynam. Systems* 27.1, pp. 241–251. MR: 2297095 (cit. on p. 1881).

V. A. Rohlin (1967). "Lectures on the entropy theory of transformations with invariant measure". *Uspehi Mat. Nauk* 22.5 (137), pp. 3–56. MR: 0217258 (cit. on p. 1878).

Klaus Schmidt (1995). *Dynamical systems of algebraic origin*. Modern Birkhäuser Classics. [2011 reprint of the 1995 original] [MR1345152]. Birkhäuser/Springer Basel AG, Basel, pp. xviii+310. MR: 3024809 (cit. on pp. 1879–1881).

Brandon Seward (2014a). "A subgroup formula for f-invariant entropy". *Ergodic Theory Dynam. Systems* 34.1, pp. 263–298. MR: 3163033 (cit. on p. 1876).

– (2014b). "Finite entropy actions of free groups, rigidity of stabilizers, and a Howe–Moore type phenomenon". *To appear in Journal d'Analyse Mathématique* (cit. on p. 1876).

– (2014c). "Krieger's finite generator theorem for ergodic actions of countable groups I". arXiv: 1405.3604 (cit. on p. 1878).

– (2015a). "Krieger's finite generator theorem for ergodic actions of countable groups II". arXiv: 1501.03367 (cit. on p. 1878).

– (2015b). "Positive entropy actions of countable groups factor onto Bernoulli shifts". *preprint, version 30* (cit. on p. 1878).

Brandon Seward and Robin D. Tucker-Drob (2016). "Borel structurability on the 2-shift of a countable group". *Ann. Pure Appl. Logic* 167.1, pp. 1–21. MR: 3413491 (cit. on p. 1878).

A. M. Stepin (1975). "Bernoulli shifts on groups". *Dokl. Akad. Nauk SSSR* 223.2, pp. 300–302. MR: 0409769 (cit. on p. 1877).

Andreas Thom (2008). "Sofic groups and Diophantine approximation". *Comm. Pure Appl. Math.* 61.8, pp. 1155–1171. MR: 2417890 (cit. on p. 1871).

Dan Voiculescu (1995). "Free probability theory: random matrices and von Neumann algebras". In: *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*. Birkhäuser, Basel, pp. 227–241. MR: 1403924 (cit. on p. 1873).

Benjamin Weiss (2000). "Sofic groups and dynamical systems". *Sankhyā Ser. A* 62.3. Ergodic theory and harmonic analysis (Mumbai, 1999), pp. 350–359. MR: 1803462 (cit. on p. 1871).

– (2015). "Entropy and actions of sofic groups". *Discrete Contin. Dyn. Syst. Ser. B* 20.10, pp. 3375–3383. MR: 3411529 (cit. on p. 1865).

Guohua Zhang (2012). "Local variational principle concerning entropy of sofic group action". *J. Funct. Anal.* 262.4, pp. 1954–1985. MR: 2873866 (cit. on p. 1873).

Received 2017-11-26.

LEWIS P. BOWEN
UNIVERSITY OF TEXAS AT AUSTIN
lpbowen@math.utexas.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1885–1904)

CRITICAL ORBITS AND ARITHMETIC EQUIDISTRIBUTION

Laura DeMarco

Abstract

These notes present recent progress on a conjecture about the dynamics of rational maps on $\mathbb{P}^1(\mathbb{C})$, connecting critical orbit relations and the structure of the bifurcation locus to the geometry and arithmetic of postcritically finite maps within the moduli space M_d. The conjecture first appeared in a 2013 publication by Baker and DeMarco. Also presented are some related results and open questions.

1 The critical orbit conjecture

These lecture notes are devoted to a conjecture presented in Baker and DeMarco [2013] and the progress made over the past five years. The setting for this problem is the dynamics of rational maps

$$f : \mathbb{P}^1(\mathbb{C}) \to \mathbb{P}^1(\mathbb{C})$$

of degree $d > 1$. Such a map has exactly $2d - 2$ critical points, when counted with multiplicity, and it is well known in the study of complex dynamical systems that the critical orbits of f play a fundamental role in understanding its general dynamical features. For example, hyperbolicity on the Julia set, linearizability near a neutral fixed point, and stability in families can all be characterized in terms of critical orbit behavior. The postcritically finite maps – those for which each of the critical points has a finite forward orbit – play a special role within the family of all maps of a given degree d.

The critical orbit conjecture, in its most basic form, is the following:

Conjecture 1.1. *Let $f_t : \mathbb{P}^1 \to \mathbb{P}^1$ be a nontrivial algebraic family of rational maps of degree $d > 1$, parametrized by t in a quasiprojective complex algebraic curve X. There are infinitely many $t \in X$ for which f_t is postcritically finite if and only if the family f_t has at most one independent critical orbit.*

MSC2010: primary 37F45; secondary 37P30, 11G05, 11G50.

Remark 1.2. Favre and Gauthier have recently announced a proof of Conjecture 1.1 for all families f_t of polynomials, building on the series of works Baker and DeMarco [2011], Ghioca, Hsia, and Tucker [2013], Baker and DeMarco [2013], Ghioca and Ye [2016], Favre and Gauthier [2016]. Other forms of this conjecture appear as Baker and DeMarco [2013, Conjecture 1.10], Ghioca, Hsia, and Tucker [2015, Conjecture 2.3], DeMarco [2016a, Conjecture 6.1], DeMarco [2016b, Conjecture 4.8], treating also the higher-dimensional parameter spaces X, where much less is known.

An algebraic family is, by definition, one for which the coefficients of f_t are meromorphic functions of t on a compactification $\overline{X}$. We also assume that f_t is a holomorphic family on the Riemann surface X, in the sense that it determines a holomorphic map $f : X \times \mathbb{P}^1 \to \mathbb{P}^1$. The family is said to be trivial if all f_t, for $t \in X$, are Möbius conjugate rational maps.

The notion of having "at most one independent critical orbit" is a bit subtle to define. I will give two candidate definitions of this notion in Section 2, so Conjecture 1.1 is actually two distinct conjectures. But, roughly speaking, if $c_i : X \to \mathbb{P}^1$, $i = 1, \ldots, 2d-2$, parametrize the critical points of f_t, then "at most one independent critical orbit" should mean that, for every pair $i \neq j$, either (1) at least one of c_i or c_j is persistently preperiodic, so that $f_t^n(c_i(t)) = f_t^m(c_i(t))$ or $f_t^n(c_j(t)) = f_t^m(c_j(t))$ for some $n > m$ and all $t \in X$; or (2) there is an orbit relation of the form $f_t^n(c_i(t)) = f_t^m(c_j(t))$ holding for all t. (Assuming condition (1) or (2) for every pair $\{i, j\}$ easily implies that there are infinitely many postcritically finite maps in the family f_t, but this assumption is too strong for a characterization: these conditions do not capture the possible symmetries in the family f_t.)

Let us put this conjecture into context. From a complex-dynamical point of view, the independent critical orbits in a holomorphic family f_t induce bifurcations. Indeed, a holomorphic family f_t with holomorphically-parametrized critical points $c_i(t)$ (for t in a disk $D \subset \mathbb{C}$) is structurally stable on its Julia set if and only if each of the critical orbits determines a normal family of holomorphic functions $\{t \mapsto f_t^n(c(t))\}_{n \geq 0}$ from D to $\mathbb{P}^1$ Mañé, Sad, and Sullivan [1983], Lyubich [1983]. For nontrivial algebraic families as in Conjecture 1.1, McMullen proved that the family is stable on all of X if and only if all of the critical points are persistently preperiodic McMullen [1987]; in other words, the family will be postcritically finite for *all* $t \in X$. Thurston's Rigidity Theorem states that the only nontrivial families of postcritically finite maps are the flexible Lattès maps, meaning that they are quotients of holomorphic maps on a family of elliptic curves Douady and Hubbard [1993]; thus, we obtain a complete characterization of stable algebraic families. From this perspective, then, Conjecture 1.1 is an attempt to characterize a slight weakening of stability, where the number of independent critical orbits is allowed to be equal to the dimension of the parameter space. One then expects interesting geometric consequences:

for example, the postcritically finite maps should be uniformly distributed with respect to the bifurcation measure (defined in DeMarco [2001] when $\dim X = 1$ and Bassanelli and Berteloot [2007] in general) on any such parameter space X.

But Conjecture 1.1 was in fact motivated more from the perspective of arithmetic geometry and the principle of unlikely intersections, as exposited in Zannier [2012]. The moduli space M_d of rational maps on $\mathbb{P}^1$ of degree $d > 1$ is naturally an affine scheme defined over $\mathbb{Q}$ Silverman [1998]. From Thurston's Rigidity Theorem, we may deduce that the postcritically finite maps lie in $\mathrm{M}_d(\overline{\mathbb{Q}})$, except for the 1-parameter families of flexible Lattès examples. Furthermore, the postcritically finite maps form

(a) a Zariski dense subset of M_d DeMarco [2016b, Theorem A], and

(b) a set of bounded Weil height in $\mathrm{M}_d(\overline{\mathbb{Q}})$, after excluding the flexible Lattès families Benedetto, Ingram, Jones, and Levy [2014, Theorem 1.1].

It is then natural to ask which algebraic subvarieties V of M_d also contain a Zariski-dense subset of postcritically finite maps. The general form of the conjecture states that this is a very special property of the variety V: it should hold if and only if V is itself defined by critical orbit relations.

This type of question is reminiscent of some famous questions and conjectures in algebraic and arithmetic geometry. To name a few, we may consider the Manin-Mumford Conjecture about abelian varieties (now theorems of Raynaud [1983a,b]) or the multiplicative version due to Lang [1960] – where a subvariety contains "too many" torsion points if and only if it is itself a subgroup (or closely related to one) – and the André–Oort conjecture which is a moduli-space analogue Klingler and Yafaev [2014], Pila [2011], and Tsimerman [2018]. In fact, our conjecture has been called the "Dynamical André–Oort Conjecture" in the literature; however, unlike for the "Dynamical Manin–Mumford Conjecture" of Zhang (S.-W. Zhang [2006] and Ghioca, Tucker, and S. Zhang [2011]), there is no overlap between the original conjecture and its dynamical analogue, at least not in the setting presented here for critical orbits and the moduli space M_d. On the other hand, there are generalizations of each – of our critical orbit conjecture and of these geometric conjectures – which *do* have overlap, and some of this is discussed in Section 5. Here I state a sample result, from my recent joint work with N. M. Mavraki, extending the work of Masser and Zannier [2010, 2014] and closely related to that of Ullmo [1998] and S.-W. Zhang [1998a].

Theorem 1.3. *DeMarco and Mavraki [2017] Let B be a quasiprojective algebraic curve defined over $\overline{\mathbb{Q}}$. Suppose $A \to B$ is a family of abelian varieties defined over $\overline{\mathbb{Q}}$ which is isogenous to a fibered product of $m \geq 2$ elliptic surfaces over B. Let $\mathfrak{L}$ be a line bundle on A which restricts to an ample and symmetric line bundle on each fiber A_t, and let $\hat{h}_t$ be the induced Néron-Tate canonical height on A_t, for each $t \in B(\overline{\mathbb{Q}})$. Finally, let $P : B \to A$*

be a section defined over $\overline{\mathbb{Q}}$. Then there exists an infinite sequence of points $t_n \in B$ for which

$$\hat{h}_{t_n}(P_{t_n}) \to 0$$

if and only if P is special.

Of course, I have not given any of the definitions of the words in this statement, so it is perhaps meaningless at first glance. My goal is merely to illustrate the breadth of concepts that connect back to the dynamical statement of Conjecture 1.1 and the existing proofs of various special cases. To make the analogy explicit: fixing a section P in $A = E_1 \times_B \cdots \times_B E_m$ would correspond, in a dynamical setting, to marking m critical points of a family of rational functions; the parameters $t \in B(\overline{\mathbb{Q}})$ where $\hat{h}_t(P_t) = 0$ correspond to the postcritically finite maps in the family; and the "specialness" of P corresponds to the family f_t having at most one independent critical orbit. In Theorem 1.3, however, the novelty is the treatment of parameters t with small (positive) height and not only height 0.

Outline. I begin by defining critical orbit relations in Section 2. Section 3 contains the sketch of a proof of a theorem from Baker and DeMarco [2011] that inspired our formulation of Conjecture 1.1 and many of the proofs that appeared afterwards, especially the cases of treated in Baker and DeMarco [2013]. Section 4 brings us up to date with what is now known about Conjecture 1.1. Finally, in Section 5, I discuss a generalization of the conjecture which motivated Theorem 1.3 and related results.

Acknowledgments. Special thanks go to Matt Baker, Myrto Mavraki, Xiaoguang Wang, and Hexi Ye for many interesting discussions and collaborations which led to the formulation of Conjecture 1.1 and the work that I have done related to it. I would also like to thank Khashayar Filom, Thomas Gauthier, Holly Krieger, and Nicole Looper for their assistance in the preparation of these lecture notes.

2 Critical orbit relations

In this section we formalize the notion of dependent critical orbits to make Conjecture 1.1 precise.

Let f_t be a nontrivial algebraic family of rational maps of degree $d > 1$, parameterized by t in a quasiprojective, complex algebraic curve X. By passing to a branched cover of X, we may assume that each of the critical points of f_t can be holomorphically parameterized by $c_i : X \to \mathbb{P}^1$, $i = 1, \ldots, 2d-2$. A critical point c_i is *persistently preperiodic* if it satisfies a relation of the form $f_t^n(c_i(t)) = f_t^m(c_i(t))$, with $n \neq m$, for all t. A pair of non-persistently-preperiodic critical points (c_i, c_j) is said to be *coincident* if we have

$$c_i(t) \text{ is preperiodic for } f_t \iff c_j(t) \text{ is preperiodic for } f_t \tag{2-1}$$

for all but finitely many $t \in X$; see DeMarco [2016a, Section 6]. If the relation (2-1) holds for every pair of non-persistently-preperiodic critical points, then the bifurcation locus of the family f_t – in the sense of Mañé, Sad, and Sullivan [1983], Lyubich [1983] – is determined by the orbit of a single critical point. That is, choosing any $i \in \{1, \ldots, 2d-2\}$ for which c_i is not persistently preperiodic, the sequence of holomorphic maps

$$\{t \mapsto f_t^n(c_i(t)) : n \geq 1\}$$

forms a normal family on an open set $U \subset X$ if and only if the family $\{f_t\}$ is J-stable on U. (See McMullen [1994, Chapter 4], Dujardin and Favre [2008a, Lemma 2.3].)

Definition 2.1 (One independent critical orbit: weak notion)**.** We say that an algebraic family f_t has at most one independent critical orbit if every pair of non-persistently-preperiodic critical points is coincident.

The relation (2-1) is implied by a more traditional notion of critical orbit relation, namely that there exist integers $n, m \geq 0$, so that

$$f_t^n(c_i(t)) = f_t^m(c_j(t)) \tag{2-2}$$

for all t. Because of the possibility of symmetries of f_t, we cannot expect (2-1) to be equivalent to (2-2). Examples are given in Baker and DeMarco [2013]. In that article, we formulated a more general notion of orbit relation that accounts for these symmetries and still implies coincidence. To define this, we let $\overline{X}$ be a smooth compactification of X, and consider the family f_t as one rational map defined over the function field $k = \mathbb{C}(\overline{X})$; it acts on $\mathbb{P}^1_{\bar{k}}$. A pair $a, b \in \mathbb{P}^1(k)$ is *dynamically related* if the point $(a, b) \in \mathbb{P}^1_{\bar{k}} \times \mathbb{P}^1_{\bar{k}}$ lies on an algebraic curve

$$V \subset \mathbb{P}^1_{\bar{k}} \times \mathbb{P}^1_{\bar{k}} \tag{2-3}$$

which is forward invariant for the product map

$$(f, f) : (\mathbb{P}^1_{\bar{k}})^2 \to (\mathbb{P}^1_{\bar{k}})^2.$$

For example, if the point a is persistently preperiodic, then it is dynamically related to any other point b, taking $V = \{(x, y) : f^n(x) = f^m(x)\}$ to depend only on one coordinate. The relation (2-2) implies that (c_i, c_j) are dynamically related, taking $V = \{(x, y) : f^n(x) = f^m(y)\}$. But also, if f commutes with a rational function A of degree ≥ 1, then points a and $b = A(a)$ are dynamically related by the invariant curve $V = \{(x, y) : y = A(x)\}$.

Definition 2.2 (One independent critical orbit: strong notion)**.** We say that an algebraic family f_t has at most one independent critical orbit if every pair of critical points is dynamically related.

I expect the two notions of "one independent critical orbit" to be equivalent DeMarco [2016a, Conjecture 6.1], but we can easily show:

Lemma 2.3. *The strong notion implies the weak notion.*

Proof. Let f_t be a nontrivial algebraic family of rational maps, for $t \in X$, and assume that it has at most one independent critical orbit, in the strong sense. Let $F_t = (f_t, f_t)$ on $\mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$ for all $t \in X$. Assume that neither c_i nor c_j is persistently preperiodic. Then there exists an algebraic curve $V \subset \mathbb{P}^1_{\bar{k}} \times \mathbb{P}^1_{\bar{k}}$ (defined over $k = \mathbb{C}(\overline{X})$, or perhaps a finite extension) so that the specializations satisfy

$$F_t^n(c_i(t), c_j(t)) \in V_t$$

for all n and all but finitely many t. Note that V_t cannot contain the vertical component $\{c_i(t)\} \times \mathbb{P}^1$ for infinitely many t: indeed, the bi-degree of the specialization V_t within $\mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$ is equal to the bi-degree (k, ℓ) of V in $\mathbb{P}^1_{\bar{k}} \times \mathbb{P}^1_{\bar{k}}$ for all but finitely many t, the curve V_t is invariant under (f_t, f_t) so a vertical component through $c_i(t)$ implies a vertical component through $f_t^n(c_i(t))$ for all n, and there are only finitely many t where the orbit of c_i has length $\leq \ell$. Thus, for all but finitely many t, if $c_i(t)$ is preperiodic, then the orbit of $c_j(t)$ is confined to lie in a finite subset, and so $c_j(t)$ will also be preperiodic. By symmetry, the same holds when $c_j(t)$ is preperiodic, and the proof is complete. □

One might ask why we only consider dynamical relations between *pairs* of critical points and not arbitrary tuples of critical points (as was first formulated in Baker and DeMarco [2013] and DeMarco [2016a]). In fact, the model-theoretic approach of Medvedev [2007] and Medvedev and Scanlon [2014] implies that it is sufficient to consider only the relations between two points.

Theorem 2.4. *Medvedev [2007, Theorem 10], Medvedev and Scanlon [2014, Fact 2.25] Suppose that f is a rational map of degree > 1, defined over a field k of characteristic 0, and assume that it is not conjugate to a monomial map, $\pm$ a Chebyshev polynomial, or a Lattès map. Let*

$$V \subset \mathbb{P}^1_{\bar{k}} \times \cdots \times \mathbb{P}^1_{\bar{k}}$$

be forward-invariant by the action of $(f, \ldots, f)$. Then each component of V is a component of the intersection

$$\bigcap_{1 \leq i \leq j \leq n} \pi_{i,j}^{-1} \pi_{i,j}(V),$$

where $\pi_{i,j}$ is the projection to the product of the i-th and j-th factors in $(\mathbb{P}^1_{\bar{k}})^n$.

In the setting of non-trivial algebraic families f_t, as in Conjecture 1.1, the monomials and Chebyshev polynomials do not arise because they would be trivial, and the flexible Lattès maps have all their critical points persistently preperiodic. Thus, we may apply Theorem 2.4 and restrict our attention to dynamical relations among critical points that depend only on two of the critical points at a time.

Even having narrowed our concept of dynamical relation to (2-3), depending on only two points, we still do not have an explicit description of all possible relations. The article Medvedev and Scanlon [ibid.] provides a careful and complete treatment of the case of polynomials f, building on the work of Ritt [1922]. Their classification of invariant curves for polynomial maps of the form (f, f) appears as a key step in the proof of the main theorem of my work with Baker and DeMarco [2013], where we prove special cases of Conjecture 1.1.

The classifcation of invariant curves in $\mathbb{P}^1 \times \mathbb{P}^1$ for a product of rational maps is still an open problem. Works by Pakovich and Zieve (see, e.g., Pakovich [2016] and Zieve [2007]) take steps towards such a classification. I posed the following question in DeMarco [2016b, Conjecture 4.8], as this represents the form of all relations I know (including for pairs of points that are not necessarily critical):

Question 2.5. *Let f_t be a non-trivial algebraic family of rational maps of degree >1, parameterized by $t \in X$, and suppose that $a, b \in \mathbb{P}^1(k)$ are two non-persistently-preperiodic points, for $k = \mathbb{C}(\overline{X})$. If a and b are dynamically related in the sense of* (2-3)*, then do there exist rational functions A, B of degrees ≥ 1 defined over $\bar{k}$ and an integer $\ell \geq 1$ so that*

$$f^\ell \circ A = A \circ f^\ell, \quad f^\ell \circ B = B \circ f^\ell, \quad \text{and} \quad A(a) = B(b)?$$

Note that A and B might themselves be iterates of f. It is known that if two rational maps of degree > 1 commute, and if they aren't monomial, Chebyshev, or Lattès, then they must share an iterate Ritt [1923].

3 Proof strategy: heights and equidistribution

In this section, I present the sketch of a proof of a closely related result from Baker and DeMarco [2011], one which initially inspired the formulation of Conjecture 1.1 and its generalizations. The ideas in the proof given here have gone into the proofs of all of the successive results related to Conjecture 1.1, though of course distinct technical issues and complications arise in each new setting.

Before getting started, I need to introduce one important tool, the canonical height of a rational function Call and Silverman [1993]. If $f : \mathbb{P}^1 \to \mathbb{P}^1$ is a rational map of degree

$d > 1$, defined over a number field, then its canonical height function

$$\hat{h}_f : \mathbb{P}^1(\overline{\mathbb{Q}}) \to \mathbb{R}_{\geq 0}$$

is defined by

$$\hat{h}_f(\alpha) = \lim_{n\to\infty} \frac{1}{d^n}\, h(f^n(\alpha))$$

where h is the usual logarithmic Weil height on $\mathbb{P}^1(\overline{\mathbb{Q}})$. It is characterized by the following two important properties: (1) there exists a constant $C = C(f)$ so that $|h - \hat{h}_f| < C$ and (2) $\hat{h}_f(f(\alpha)) = d\,\hat{h}_f(\alpha)$ for all $\alpha \in \mathbb{P}^1(\overline{\mathbb{Q}})$. As a consequence, we have $\hat{h}_f(\alpha) = 0$ if and only if α has finite forward orbit for f Call and Silverman [1993, Corollary 1.1.1].

Theorem 3.1. *Baker and DeMarco [2011] Fix points $a_1, a_2 \in \mathbb{C}$. Let $f_t(z) = z^2 + t$ be the family of quadratic polynomials, for $t \in \mathbb{C}$, and define*

$$S(a_i) := \{t \in \mathbb{C} : a_i \textit{ is preperiodic for } f_t\}.$$

Then the intersection $S(a_1) \cap S(a_2)$ is infinite if and only if $a_1 = \pm a_2$.

Sketch of Proof. Step 1 is to treat the easy implication: assume that $a_1 = \pm a_2$ and deduce that $S(a_1) \cap S(a_2)$ is an infinite set. This uses a standard argument from complex dynamics. For any given point a, we first observe that the family of functions $\{t \mapsto f_t^n(a)\}$ cannot be normal on all of $\mathbb{C}$. Indeed, for all t large, we find that $f_t^n(a) \to \infty$ as $n \to \infty$, while for $t = a - a^2$, the point a is fixed by f_t. In fact, via Montel's Theorem on normal families, there must be infinitely many values of t for which a is preperiodic, and therefore $S(a)$ is infinite. If $a_1 = \pm a_2$, then $f_t(a_1) = f_t(a_2)$ for all t, and therefore $S(a_1) = S(a_2)$.

The goal of Step 2 is to show that $S(a_1) \cap S(a_2)$ being infinite implies that a_1 and a_2 are coincident: we will see that $S(a_1) = S(a_2)$. First assume that a_1 and a_2 are algebraic numbers, and suppose K is a number field containing both a_1 and a_2. We define a height function on $\mathbb{P}^1(\overline{K})$ associated to each a_i. Indeed, for each $t \in \overline{K}$, we set

$$h_i(t) := \hat{h}_{f_t}(a_i)$$

where $\hat{h}_{f_t}$ is the canonical height function of f_t. In particular, we see that

$$h_i(t) = 0 \iff t \in S(a_i).$$

It turns out that h_i is the height associated to a continuous adelic metric of non-negative curvature on the line bundle $\mathcal{O}(1)$ on $\mathbb{P}^1$ (in the sense of S. Zhang [1995]) and an adelic measure (in the sense of Baker and Rumely [2006] and Favre and Rivera-Letelier [2006]).

Therefore, we may apply the equidistribution theorems of Baker and Rumely [2010], Favre and Rivera-Letelier [2006], and Chambert-Loir [2006] to see that the elements of $S(a_i)$ are uniformly distributed with respect to a natural measure μ_i on $\mathbb{P}^1(\mathbb{C})$. More precisely, given any sequence of finite subsets $S_n \subset S(a_i)$ which are $\mathrm{Gal}(\overline{K}/K)$-invariant and with $|S_n| \to \infty$, the discrete probablity measures

$$\mu_{S_n} = \frac{1}{|S_n|} \sum_{s \in S_n} \delta_s$$

on $\mathbb{P}^1(\mathbb{C})$ will converge weakly to μ_i. In particular, when $S(a_1) \cap S(a_2)$ is infinite, this set – because it is $\mathrm{Gal}(\overline{K}/K)$ invariant – will be uniformly distributed with respect to *both* μ_1 and μ_2, allowing us to deduce that $\mu_1 = \mu_2$. Even more, by the nondegeneracy of a pairing between heights of this form, we can also conclude that $h_1 = h_2$. Therefore $S(a_1) = S(a_2)$.

Step 2 will be complete if we can also treat the case where at least one of the a_i is not algebraic. In this setting, the smallest field K containing the a_i and $\mathbb{Q}$ will have a nonzero transcendence degree over $\mathbb{Q}$, so we treat K as a function field. The arithmetic equidistribution theorems (Baker and Rumely [2010], Favre and Rivera-Letelier [2006], and Chambert-Loir [2006]) work just as well in this setting. However, the equidistribution in question – of Galois-invariant subsets of $S(a_i)$ becoming uniformly distributed with respect to a natural measure μ_i – is no longer taking place on the Riemann sphere $\mathbb{P}^1(\mathbb{C})$. Instead, we obtain a geometric convergence statement on a family of Berkovich projective lines $\mathbf{P}^1_v$, one for each place v of the function field K. Nevertheless, one concludes from equidistribution that if $S(a_1) \cap S(a_2)$ is infinite, then the heights h_1 and h_2 on $\mathbb{P}^1(\overline{K})$ must coincide, and therefore

$$a_1(t) \text{ is preperiodic for } f_t \iff a_2(t) \text{ is preperiodic for } f_t$$

for all $t \in \overline{K}$. But if there is some element t of $\mathbb{C}$ where $a_1(t)$ is preperiodic, then that t can be identified with an element of $\overline{K}$, so in fact

$$a_1(t) \text{ is preperiodic for } f_t \iff a_2(t) \text{ is preperiodic for } f_t$$

for all $t \in \mathbb{C}$.

Step 3 is to show that coincidence implies an explicit relation between the two points a_1 and a_2. When a_1 and a_2 are both algebraic, a closer look at the definitions of the heights h_i reveals that the measure μ_i on $\mathbb{P}^1(\mathbb{C})$ is the equilibrium measure on the boundary of a "Mandelbrot-like" set associated to the point a_i. That is, we consider the sets

$$M_i = \{t \in \mathbb{C} : \sup_n |f_t^n(a_i)| < \infty\},$$

and μ_i is the harmonic measure for the domain $\hat{\mathbb{C}} \setminus M_i$ centered at ∞. (Note that if $a_i = 0$, then M_i is the usual Mandelbrot set M, and μ_i is the bifurcation measure for the family f_t.) But even for non-algebraic points a_i, having concluded from Step 2 that $S(a_1) = S(a_2)$, we see that $M_1 = M_2$; this is because the set M_i is obtained from the closure $\overline{S(a_i)}$ by filling in the bounded complementary components. Now, just as in the original proof that the Mandelbrot set M is connected, which uses a dynamical construction of the Riemann map to $\hat{\mathbb{C}} \setminus M$, we investigate the uniformizing map near ∞ for the sets $M_1 = M_2$. The injectivity of that map – built out of the Böttcher coordinates near ∞ for the maps f_t with t large – allows us to deduce that $f_t(a_1) = f_t(a_2)$ for all t large. Therefore, we have $a_1 = \pm a_2$. □

It is worth observing at this point, as was observed in Baker and DeMarco [2011], that the proof of Theorem 3.1 gives a stronger statement. The arithmetic equidistribution theorems allow us to treat intersections of points of small height and not only those of height 0 (for the heights h_i introduced in the proof). For example, the proof provides:

Theorem 3.2. *Baker and DeMarco [ibid.] Fix points $a_1, a_2 \in \overline{\mathbb{Q}}$. Let $f_t(z) = z^2 + t$ be the family of quadratic polynomials, for $t \in \mathbb{C}$, and define*

$$S(a_i) := \{t \in \mathbb{C} : a_i \text{ is preperiodic for } f_t\}.$$

Then $S(a_i) \subset \overline{\mathbb{Q}}$, and the following are equivalent:

1. *there exists an infinite sequence $t_n \in \overline{\mathbb{Q}}$ for which $h_1(t_n) \to 0$ and $h_2(t_n) \to 0$;*
2. *the intersection $S(a_1) \cap S(a_2)$ is infinite;*
3. *$S(a_1) = S(a_2)$;*
4. *$\mu_1 = \mu_2$; and*
5. *$a_1 = \pm a_2$.*

The original motivation for statements like Theorem 3.2, and specifically the inclusion of condition (1), includes the Bogomolov Conjecture, proved by Ullmo and Zhang Ullmo [1998] and S.-W. Zhang [1998a], building on the equidistribution theorem of Szpiro, Ullmo, and S. Zhang [1997]. I will return to this theme in Section 5.

4 What is known

A good deal of work has gone into proving Conjecture 1.1 and its generalizations in various settings. Here, I mention some of the key recent developments. Most progress has been made in the context of polynomial dynamics. One important advantage of working with polynomials is that the conjecture itself is easier to state in a more precise form: the critical orbit relations, in the sense of (2-3), have been classified, as discussed in Section 2. But also, we have the advantage of extra tools: the uniformizing Böttcher coordinates of a complex polynomial near ∞ have proved immensely useful (as in Step 3 of the proof of Theorem 3.1), and the height functions (as defined in Step 2 of Theorem 3.1) are easier to work with. For example, the main theorem of Favre and Gauthier [2017] addresses an important property of the dynamically-defined local height functions, used in their proof of Conjecture 1.1 for families of polynomials; the analogous result fails for general families of rational maps DeMarco and Okuyama [2017].

Conjecture 1.1 is trivially satisfied for polynomials in degree 2, where the moduli space has dimension 1 and can be parameterized by the family $f_t(z) = z^2 + t$ with $t \in \mathbb{C}$ with exactly one independent critical point at $z = 0$. There are infinitely many postcritically finite polynomials in this family. Furthermore, it is known that the postcritically finite maps are uniformly distributed with respect to the bifurcation measure in this family (which is equal to the equilibrium measure μ_M on the boundary of Mandelbrot set) Levin [1989]. In addition, Baker and Hsia proved an arithmetic version of the equidistribution theorem, deducing that any sequence of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-invariant subsets of the postcritically finite parameters is also uniformly distributed with respect to μ_M Baker and Hsia [2005].

In Baker and DeMarco [2013], we proved Conjecture 1.1 for families f_t of polynomials of arbitrary degree, parametrized by $t \in \mathbb{C}$ with coefficients that were polynomial in t. In degree 3, but for arbitrary algebraic families, the following result was obtained a few years ago, independently by Favre-Gauthier and Ghioca-Ye.

Theorem 4.1. *Favre and Gauthier [2016] and Ghioca and Ye [2016] Let X be an irreducible complex algebraic curve in the space $\mathcal{P}_3 \simeq \mathbb{C}^2$ of polynomials of the form*

$$f_{a,b}(z) = z^3 - 3a^2 z + b,$$

with critical points at $\pm a$. There are infinitely many $t \in X$ for which f_t is postcritically finite if and only if one of the following holds:

1. *either a or $-a$ is persistently preperiodic on X;*
2. *there is a symmetry of the form*

$$f_t(-z) = -f_t(z)$$

for all $t \in X$ and all $z \in \mathbb{C}$, and $X = \{(a,b) \in \mathbb{C}^2 : b = 0\}$; or

3. *there exist non-negative integers n and m so that*

$$f_t^n(a(t)) = f_t^m(-a(t))$$

for all $t \in X$.

Furthermore, in each of these cases, the postcritically finite maps will be (arithmetically) equidistributed with respect to the bifurcation measure on X.

Remark 4.2. Favre and Gauthier have recently announced a proof of Conjecture 1.1 for families f_t of polynomials in arbitrary degree, extending Theorem 4.1.

The proof of Theorem 4.1 has the same outline as the proof of Theorem 3.1. The idea is to consider the two points a_t and $-a_t$ for $t \in X$ and follow the same three steps. Step 1, the "easy" implication, follows as before, with the additional input that any non-persistently-preperiodic critical point along the curve X must be undergoing bifurcations Dujardin and Favre [2008a, Theorem 2.5].

For Step 2, we assume that neither critical point is persistently preperiodic and that there are infinitely many postcritically finite maps on X, and we aim to show coincidence of the two points. There is one simplifying condition in the setting of Theorem 4.1: the postcritically finite maps are algebraic points in $\mathcal{P}_3$, so X must itself be defined over $\overline{\mathbb{Q}}$. Thus we can avoid the arguments needed for the transcendental case of Theorem 3.1. Nevertheless, there are new difficulties that arise; for example, it is not obvious that the height functions

$$h_{\pm a}(t) := \hat{h}_{f_t}(\pm a_t)$$

defined on $X(\overline{\mathbb{Q}})$ will satisfy the hypotheses of the existing arithmetic equidistribution theorems. This is checked with some careful estimates near the cusps of X. Then one can apply the equidistribution theorems of Yuan [2008], Thuillier [2005], and Chambert-Loir [2006] and conclude that $h_a = h_{-a}$ on X. In particular, we then have that $a(t)$ is preperiodic if and only if $-a(t)$ is preperiodic, for all $t \in X$.

Finally, one needs to deduce the explicit algebraic relations on the critical points, as in Step 3 of Theorem 3.1. One strategy is provided in my work with Baker and DeMarco [2013], to first produce an *analytic* relation between the critical points a and $-a$ via the Böttcher coordinates at infinity (similar to what was done for Theorem 3.1 but in a more general setting). From there, we used iteration to promote the analytic relation to an algebraic relation which is invariant under the dynamics. This strategy is followed in Ghioca and Ye [2016]; an alternative approach is given in Favre and Gauthier [2016]. To obtain the explicit form of the relation, Baker and I used results of Medvedev and Scanlon [2014], classifying the invariant curves for (f, f) acting on $\mathbb{P}^1 \times \mathbb{P}^1$ (over the function field $\mathbb{C}(X)$). The results of Favre-Gauthier and Ghioca-Ye for cubic polynomials simplify the relations further to give the possibilities appearing in Theorem 4.1.

One can also formulate a version of Conjecture 1.1 for tuples of polynomials or rational maps, rather than a single family of rational maps. The following result answered a question posed by Patrick Ingram, inspired by the result of André about complex-multiplication pairs in the moduli space $\mathfrak{M}_1 \times \mathfrak{M}_1 \simeq \mathbb{C}^2$ of pairs of elliptic curves André [1998].

Theorem 4.3. *Ghioca, Krieger, Nguyen, and Ye [2017] Let X be an irreducible complex algebraic curve in the space $\mathcal{P}_2 \times \mathcal{P}_2 \simeq \mathbb{C}^2$ of pairs of quadratic polynomials of the form $f_t(z) = z^2 + t$. If X contains infinitely many pairs (t_1, t_2) for which both f_{t_1} and f_{t_2} are postcritically finite, then X is*

1. *a vertical line $\{t_1\} \times \mathbb{C}$ where f_{t_1} is postcritically finite;*
2. *a horizontal line $\mathbb{C} \times \{t_2\}$ where f_{t_2} is postcritically finite; or*
3. *the diagonal $\{(t,t) : c \in \mathbb{C}\}$.*

By contrast, in the case of pairs of elliptic curves, there is an infinite collection of modular curves in $\mathfrak{M}_1 \times \mathfrak{M}_1$, all of which contain infinitely many CM pairs. Thus, Theorem 4.3 tells us that there is no analogue of these modular curves in the quadratic family. To see this, the authors prove an important rigidity property of the Mandelbrot set M: it is not invariant under nontrivial algebraic correspondences. This rigidity was recently extended to a local, analytic rigidity statement in Luo [2017]: Luo proved that any conformal isomorphism between domains $U, V \subset \mathbb{C}$ intersecting the boundary ∂M, sending $U \cap \partial M$ to $V \cap \partial M$, must be the identity.

For non-polynomial rational maps, Conjecture 1.1 is only known for some particular families. For example, in the moduli space of quadratic rational maps $\mathrm{M}_2 \simeq \mathbb{C}^2$, for each $\lambda \in \mathbb{C}$, one may consider the dynamically-defined subvariety

$$\mathrm{Per}_1(\lambda) = \{f \text{ in } \mathrm{M}_2 \text{ with a fixed point of multiplier } \lambda\},$$

where the multiplier of a fixed point is simply the derivative of f at that point Milnor [1993]. (Similarly, one can define the algebraic curves $\mathrm{Per}_n(\lambda)$ for maps with a cycle of period n and multiplier λ, but one should take care in the definition when $\lambda = 1$.) Observe that the curve $\mathrm{Per}_1(0)$ is defined by a critical orbit relation, the existence of a fixed critical point. Thus, the curve $\mathrm{Per}_1(0)$ coincides with the family of polynomials within M_2; in particular, it contains infinitely many postcritically finite maps.

Theorem 4.4. *DeMarco, Wang, and Ye [2015] Fix $\lambda \in \mathbb{C}$. The curve $\mathrm{Per}_1(\lambda)$ in the moduli space of quadratic rational maps contains infinitely many postcritically finite maps if and only if $\lambda = 0$.*

Remark 4.5. The analogous result for $\mathrm{Per}_1(\lambda)$ in the space of cubic polynomials was obtained in Baker and DeMarco [2013]. The theorem is proved for curves $\mathrm{Per}_n(\lambda)$, for every n, in the space of cubic polynomials in Favre and Gauthier [2016].

All of these theorems are closely related to questions about the geometry of the bifurcation locus the family f_t of rational maps, as seen in Step 3 in the proof of Theorem 3.1, or the statement about the rigidity of the Mandelbrot set used to prove Theorem 4.3. In the case of $\mathrm{Per}_1(\lambda)$, we should first pass to a double cover $\widehat{\mathrm{Per}}_1(\lambda)$ where the two critical points can be holomorphically and independently parameterized by c_1 and c_2. The bifurcation measure μ_i^λ of the critical point c_i reflects the failure of the family $\{t \mapsto f_t^n(c_i(t))\}$ to be normal on $\widehat{\mathrm{Per}}_1(\lambda)$. Let $S(c_i)$ denote the set of parameters t where the critical point c_i is preperiodic. It is known that $S(c_i)$ is uniformly distributed with respect to this measure μ_i^λ for all $\lambda \in \mathbb{C}$ Dujardin and Favre [2008a,b]. The proof of Theorem 4.4 uses the following two strengthenings of this equidistribution statement:

Theorem 4.6. *DeMarco, Wang, and Ye [2015] For each $\lambda \in \mathbb{C} \setminus \{0\}$, we have $\mu_1^\lambda \neq \mu_2^\lambda$ on $\widehat{\mathrm{Per}}_1(\lambda)$.*

Theorem 4.7. *DeMarco, Wang, and Ye [2015] and Mavraki and Ye [2015] For each $\lambda \in \overline{\mathbb{Q}} \setminus \{0\}$, we have arithmetic equidistribution of $S(c_1)$ and $S(c_2)$. That is, for any infinite sequence t_n in $S(c_i)$, the discrete measures*

$$\frac{1}{|\operatorname{Gal}(\overline{\mathbb{Q}(\lambda)}/\mathbb{Q}(\lambda)) \cdot t_n|} \sum_{t \in \operatorname{Gal}(\overline{\mathbb{Q}(\lambda)}/\mathbb{Q}(\lambda)) \cdot t_n} \delta_t$$

converge weakly to the bifurcation measure μ_i^λ on $\widehat{\mathrm{Per}}_1(\lambda)$.

Remark 4.8. The height functions $h_i(t) := \hat{h}_{f_t}(c_i(t))$ on $\widehat{\mathrm{Per}}_1(\lambda)$ provided the first examples of this type that are *not* adelic – in the sense of S. Zhang [1995], Baker and Rumely [2010], and Favre and Rivera-Letelier [2006] – and therefore did *not* satisfy the hypotheses of the existing equidistribution theorems. The article Mavraki and Ye [2015] extends the equidistribution theorems of Baker and Rumely [2010], Favre and Rivera-Letelier [2006], and Chambert-Loir [2006] for heights on $\mathbb{P}^1(\overline{\mathbb{Q}})$ to the setting of quasi-adelic heights.

Remark 4.9. Despite Theorem 4.6, it is not yet known if $\operatorname{supp}\mu_1^\lambda \neq \operatorname{supp}\mu_2^\lambda$ for all $\lambda \in \mathbb{C}$; see DeMarco, Wang, and Ye [2015, Question 2.4]. One can ask, much more generally, about the bifurcation loci associated to independent critical points in algebraic families f_t in every degree and if they can ever coincide; see DeMarco [2016b, Question 2.5].

An assortment of results is known for other families of rational functions. For example, in Ghioca, Hsia, and Tucker [2015], the authors treat maps of the form

$$f_t(z) = g(z) + t$$

for $t \in \mathbb{C}$, where $g \in \overline{\mathbb{Q}}(z)$ is a rational function of degree > 2 with a superattracting fixed point at ∞. They show the weaker form of the conjecture, deducing coincidence of the critical orbits if there are infinitely many postcritically finite maps; this follows from an equidistribution result associated to the dynamically-defined height functions on $\mathbb{P}^1(\overline{\mathbb{Q}})$ (similar to Step 2 in the proofs of Theorems 3.1 and 4.1).

Almost nothing is known about Conjecture 1.1 in the context of higher-dimensional parameter spaces X, apart from the "easy" implication of the conjecture, as in Step 1 of Theorem 3.1; see DeMarco [2016a]. A general form of the arithmetic equidistribution theorem exists for higher-dimensional arithmetic varieties Yuan [2008], but the challenge lies in understanding when a dynamically-defined height will satisfy the stated hypotheses; see, e.g., Favre and Gauthier [2015] and Remark 4.8 above.

5 Arbitrary points

In this final section, I present some results about a generalization of Conjecture 1.1 that connects with interesting results and questions about elliptic curves (or more general families of abelian varieties). As in Theorem 3.1, one can study the orbits of arbitrary points, not only the critical orbits. This may seem less motivated in the context of studying complex dynamical systems, as the critical points are the ones that induce bifurcations (in the traditional dynamical sense), but the problem is quite natural from another point of view.

As an example, the following result was proved by Masser and Zannier, motivated by a conjecture of Pink [2005]:

Theorem 5.1. *Masser and Zannier [2010, 2012] Let E_t be a non-isotrivial family of elliptic curves over a quasiprojective curve B, so defining an elliptic curve E over the function field $k = \mathbb{C}(\overline{B})$. Suppose that P and Q are non-torsion elements of $E(k)$. There are infinitely many $t \in B(\mathbb{C})$ for which P_t and Q_t are both torsion on E_t if and only if there exist nonzero integers n and m so that $nP + mQ = 0$ on E.*

Because the multiplication-by-m maps on an elliptic curve descend to rational maps on $\mathbb{P}^1$, and because the torsion points on the elliptic curve project to the preperiodic points on $\mathbb{P}^1$, Theorem 5.1 has a direct translation into a dynamical statement:

Theorem 5.2. *Let f_t be a family of flexible Lattès maps on $\mathbb{P}^1$ parameterized by $t \in B$, induced from an endomorphism of a non-isotrivial elliptic curve E over $k = \mathbb{C}(\overline{B})$. Fix non-persistently-preperiodic points $P, Q \in \mathbb{P}^1(k)$. Then there are infinitely many $t \in B$ for which both P_t and Q_t are preperiodic for f_t if and only if there exist Lattès maps g_t and h_t (also induced by endomorphisms of E) for which*

$$g_t(P_t) = h_t(Q_t)$$

for all t.

In fact, it was Theorem 5.1 that inspired Theorem 3.1 in the first place: Baker and I were answering a question posed by Umberto Zannier. And compare the conclusion of Theorem 5.2 to that of Question 2.5: note that the Lattès maps g_t and h_t will commute with f_t. Conjecture 1.1 is just a special case of conjectures presented in Ghioca, Hsia, and Tucker [2015] or DeMarco [2016a], addressing algebraic families f_t and arbitrary (non-critical) pairs of marked points, and for which Theorems 5.1 and 5.2 are also a special case:

Conjecture 5.3. *Let $f_t : \mathbb{P}^1 \to \mathbb{P}^1$ be a nontrivial algebraic family of rational maps of degree $d > 1$, parametrized by t in a quasiprojective, complex algebraic curve X. Suppose that $a, b : X \to \mathbb{P}^1$ are meromorphic functions on a compactification $\overline{X}$. There are infinitely many $t \in X$ for which both $a(t)$ and $b(t)$ are preperiodic for f_t if and only if a and b are dynamically related.*

When E is the Legendre family of elliptic curves, and the points P and Q lie in $\mathbb{P}^1(\mathbb{C})$, X. Wang, H. Ye, and I gave a dynamical proof of Theorem 5.2, building on the same ideas that went into the proof of Theorem 3.1. As in Theorem 3.2, we obtain the stronger statement about parameters of small height, which does not follow from the proofs given in Masser and Zannier [2010, 2012]. For algebraic points, our theorem can be stated as:

Theorem 5.4. DeMarco, Wang, and Ye [2016] *Let $E_t = \{(x, y) : y^2 = x(x-1)(x-t)\}$ be the Legendre family of elliptic curves, with $t \in \mathbb{C} \setminus \{0, 1\}$. Fix $a, b \in \overline{\mathbb{Q}} \setminus \{0, 1\}$. The following are equivalent:*

1. $|\mathrm{Tor}(a) \cap \mathrm{Tor}(b)| = \infty$;
2. $\mathrm{Tor}(a) = \mathrm{Tor}(b)$;
3. *there is an infinite sequence $\{t_n\} \subset \overline{\mathbb{Q}}$ so that $\hat{h}_a(t_n) \to 0$ and $\hat{h}_b(t_n) \to 0$;*
4. *$\mu_a = \mu_b$ on $\mathbb{P}^1(\mathbb{C})$; and*
5. *$a = b$.*

Here, $\mathrm{Tor}(a) = \{t \in \mathbb{C} : (a, \sqrt{a(a-1)(a-t)})$ is torsion on $E_t\}$. The height $\hat{h}_a(t)$ is the Néron-Tate canonical height of the point $(a, \sqrt{a(a-1)(a-t)})$ in E_t for $t \in \overline{\mathbb{Q}} \setminus \{0, 1\}$, so that $\hat{h}_a(t) = 0$ if and only if $t \in \mathrm{Tor}(a)$. As in the theorems of the previous section, the geometry of the "bifurcation locus" associated to the marked points a and b plays a key role; the $\mathrm{Gal}(\overline{\mathbb{Q}(a)}/\mathbb{Q}(a))$-invariant subsets of $\mathrm{Tor}(a)$ are uniformly distributed with respect to a natural measure μ_a. As in the dynamical examples, Theorems 3.2 or 4.1, the measures μ_a at the *archimedean places* (the limiting distributions on the underlying complex curve) are sufficient to characterize the existence of a dynamical relation. It is not known if this will be the case for all families of rational maps in Conjecture 5.3.

And this returns us, finally, to the statement of Theorem 1.3 from the Introduction. That result proves a special case of a conjecture of S.-W. Zhang [1998b] from his ICM lecture notes from exactly 20 years ago, which was posed as an extension of the Bogomolov Conjecture to non-trivial families of abelian varieties. The notion of a "special" section is carefully defined in citeDM:variation, building on the work of Masser and Zannier [2012, 2014]. Our proof was inspired by the combination of ideas presented here, connecting dynamical orbit relations and equidistribution theorems with the geometry of abelian varieties. These ideas are, in turn, closely related to the original proofs of Ullmo and Zhang of the Bogomolov Conjecture Ullmo [1998] and S.-W. Zhang [1998a], relying on the (arithmetic) equidistribution of the torsion points within an abelian variety defined over $\overline{\mathbb{Q}}$ Szpiro, Ullmo, and S. Zhang [1997]. We have not yet been able to give a purely dynamical proof of Theorem 1.3, in the flavor of Theorems 3.1 and 5.4. Instead, we used the work of Silverman [1992, 1994a,b] to provide the technical statements needed to show that our height functions on $B(\overline{\mathbb{Q}})$ satisfy all the hypotheses needed to apply the arithmetic equidistribution theorems of Thuillier and Yuan Thuillier [2005] and Yuan [2008]. Via the equidistribution theorem, we were able to reduce the statement of Theorem 1.3 to a more general form of Theorem 5.1 proved by Masser and Zannier [2014].

References

Yves André (1998). "Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire". *J. Reine Angew. Math.* 505, pp. 203–208. MR: 1662256 (cit. on p. 1897).

Matthew H. Baker and Laura DeMarco (2011). "Preperiodic points and unlikely intersections". *Duke Math. J.* 159.1, pp. 1–29. MR: 2817647 (cit. on pp. 1886, 1888, 1891, 1892, 1894).

– (2013). "Special curves and postcritically finite polynomials". *Forum Math. Pi* 1, e3, 35. MR: 3141413 (cit. on pp. 1885, 1886, 1888–1891, 1895–1897).

Matthew H. Baker and Liang-Chung Hsia (2005). "Canonical heights, transfinite diameters, and polynomial dynamics". *J. Reine Angew. Math.* 585, pp. 61–92. MR: 2164622 (cit. on p. 1895).

Matthew H. Baker and Robert Rumely (2006). "Equidistribution of small points, rational dynamics, and potential theory". *Ann. Inst. Fourier (Grenoble)* 56.3, pp. 625–688. MR: 2244226 (cit. on p. 1892).

Matthew H. Baker and Robert Rumely (2010). *Potential theory and dynamics on the Berkovich projective line*. Vol. 159. Mathematical Surveys and Monographs. American

Mathematical Society, Providence, RI, pp. xxxiv+428. MR: 2599526 (cit. on pp. 1893, 1898).

Giovanni Bassanelli and François Berteloot (2007). "Bifurcation currents in holomorphic dynamics on $\mathbb{P}^k$". *J. Reine Angew. Math.* 608, pp. 201–235. MR: 2339474 (cit. on p. 1887).

Robert Benedetto, Patrick Ingram, Rafe Jones, and Alon Levy (2014). "Attracting cycles in p-adic dynamics and height bounds for postcritically finite maps". *Duke Math. J.* 163.13, pp. 2325–2356. MR: 3265554 (cit. on p. 1887).

Gregory S. Call and Joseph H. Silverman (1993). "Canonical heights on varieties with morphisms". *Compositio Math.* 89.2, pp. 163–205. MR: 1255693 (cit. on pp. 1891, 1892).

Antoine Chambert-Loir (2006). "Mesures et équidistribution sur les espaces de Berkovich". *J. Reine Angew. Math.* 595, pp. 215–235. MR: 2244803 (cit. on pp. 1893, 1896, 1898).

Laura DeMarco (2001). "Dynamics of rational maps: a current on the bifurcation locus". *Math. Res. Lett.* 8.1-2, pp. 57–66. MR: 1825260 (cit. on p. 1887).

– (2016a). "Bifurcations, intersections, and heights". *Algebra Number Theory* 10.5, pp. 1031–1056. MR: 3531361 (cit. on pp. 1886, 1889, 1890, 1899, 1900).

– (2016b). "KAWA 2015: Dynamical moduli spaces and elliptic curves". To appear in Ann. Fac. Sci. Toulouse Math. arXiv: 1607.04471 (cit. on pp. 1886, 1887, 1891, 1898).

Laura DeMarco and Niki Myrto Mavraki (Jan. 2017). "Variation of canonical height and equidistribution". arXiv: 1701.07947 (cit. on p. 1887).

Laura DeMarco and Yûsuke Okuyama (Oct. 2017). "Discontinuity of a degenerating escape rate". arXiv: 1710.01660 (cit. on p. 1895).

Laura DeMarco, Xiaoguang Wang, and Hexi Ye (2015). "Bifurcation measures and quadratic rational maps". *Proc. Lond. Math. Soc. (3)* 111.1, pp. 149–180. MR: 3404779 (cit. on pp. 1897, 1898).

– (2016). "Torsion points and the Lattès family". *Amer. J. Math.* 138.3, pp. 697–732. MR: 3506383 (cit. on p. 1900).

Adrien Douady and John H. Hubbard (1993). "A proof of Thurston's topological characterization of rational functions". *Acta Math.* 171.2, pp. 263–297. MR: 1251582 (cit. on p. 1886).

Romain Dujardin and Charles Favre (2008a). "Distribution of rational maps with a preperiodic critical point". *Amer. J. Math.* 130.4, pp. 979–1032. MR: 2427006 (cit. on pp. 1889, 1896, 1898).

– (2008b). "Distribution of rational maps with a preperiodic critical point". *Amer. J. Math.* 130.4, pp. 979–1032. arXiv: math/0601612. MR: 2427006 (cit. on p. 1898).

C. Favre and T. Gauthier (2015). "Distribution of postcritically finite polynomials". *Israel J. Math.* 209.1, pp. 235–292. MR: 3430241 (cit. on p. 1899).

Charles Favre and Thomas Gauthier (2016). "Classification of special curves in the space of cubic polynomials". arXiv: 1603.05126 (cit. on pp. 1886, 1895–1897).

– (June 2017). "Continuity of the Green function in meromorphic families of polynomials". arXiv: 1706.04676 (cit. on p. 1895).

Charles Favre and Juan Rivera-Letelier (2006). "Équidistribution quantitative des points de petite hauteur sur la droite projective". *Math. Ann.* 335.2, pp. 311–361. MR: 2221116 (cit. on pp. 1892, 1893, 1898).

D. Ghioca, L.-C. Hsia, and T. J. Tucker (2015). "Preperiodic points for families of rational maps". *Proc. Lond. Math. Soc. (3)* 110.2, pp. 395–427. MR: 3335283 (cit. on pp. 1886, 1898, 1900).

D. Ghioca, H. Krieger, K. D. Nguyen, and H. Ye (2017). "The dynamical André-Oort conjecture: unicritical polynomials". *Duke Math. J.* 166.1, pp. 1–25. MR: 3592687 (cit. on p. 1897).

Dragos Ghioca, Liang-Chung Hsia, and Thomas J. Tucker (2013). "Preperiodic points for families of polynomials". *Algebra Number Theory* 7.3, pp. 701–732. MR: 3095224 (cit. on p. 1886).

Dragos Ghioca, Thomas J. Tucker, and Shouwu Zhang (2011). "Towards a dynamical Manin-Mumford conjecture". *Int. Math. Res. Not. IMRN* 22, pp. 5109–5122. MR: 2854724 (cit. on p. 1887).

Dragos Ghioca and Hexi Ye (Mar. 2016). "The Dynamical Andre-Oort Conjecture for cubic polynomials". arXiv: 1603.05303 (cit. on pp. 1886, 1895, 1896).

Bruno Klingler and Andrei Yafaev (2014). "The André-Oort conjecture". *Ann. of Math. (2)* 180.3, pp. 867–925. MR: 3245009 (cit. on p. 1887).

Serge Lang (1960). "Integral points on curves". *Inst. Hautes Études Sci. Publ. Math.* 6, pp. 27–43. MR: 0130219 (cit. on p. 1887).

G. M. Levin (1989). "On the theory of iterations of polynomial families in the complex plane". *Teor. Funktsiĭ Funktsional. Anal. i Prilozhen.* 51, pp. 94–106. MR: 1009151 (cit. on p. 1895).

Yusheng Luo (2017). *On the inhomogeneity of the Mandelbrot set*. Preprint (cit. on p. 1897).

M. Yu. Lyubich (1983). "Some typical properties of the dynamics of rational mappings". *Uspekhi Mat. Nauk* 38.5(233), pp. 197–198. MR: 718838 (cit. on pp. 1886, 1889).

R. Mañé, P. Sad, and D. Sullivan (1983). "On the dynamics of rational maps". *Ann. Sci. École Norm. Sup. (4)* 16.2, pp. 193–217. MR: 732343 (cit. on pp. 1886, 1889).

D. Masser and U. Zannier (2010). "Torsion anomalous points and families of elliptic curves". *Amer. J. Math.* 132.6, pp. 1677–1691. MR: 2766181 (cit. on pp. 1887, 1899, 1900).

– (2012). "Torsion points on families of squares of elliptic curves". *Math. Ann.* 352.2, pp. 453–484. MR: 2874963 (cit. on pp. 1899–1901).

D. Masser and U. Zannier (2014). "Torsion points on families of products of elliptic curves". *Adv. Math.* 259, pp. 116–133. MR: 3197654 (cit. on pp. 1887, 1901).

Niki Myrto Mavraki and Hexi Ye (Feb. 2015). "Quasi-adelic measures and equidistribution on $\mathbb{P}^1$". arXiv: 1502.04660 (cit. on p. 1898).

Curtis T. McMullen (1987). "Families of rational maps and iterative root-finding algorithms". *Ann. of Math. (2)* 125.3, pp. 467–493. MR: 890160 (cit. on p. 1886).

– (1994). *Complex dynamics and renormalization*. Vol. 135. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, pp. x+214. MR: 1312365 (cit. on p. 1889).

Alice Medvedev (2007). *Minimal sets in ACFA*. Thesis (Ph.D.)–University of California, Berkeley. ProQuest LLC, Ann Arbor, MI, p. 96. MR: 2710587 (cit. on p. 1890).

Alice Medvedev and Thomas Scanlon (2014). "Invariant varieties for polynomial dynamical systems". *Ann. of Math. (2)* 179.1, pp. 81–177. MR: 3126567 (cit. on pp. 1890, 1891, 1896).

John Milnor (1993). "Geometry and dynamics of quadratic rational maps". *Experiment. Math.* 2.1. With an appendix by the author and Lei Tan, pp. 37–83. MR: 1246482 (cit. on p. 1897).

Fedor Pakovich (2016). "On semiconjugate rational functions". *Geom. Funct. Anal.* 26.4, pp. 1217–1243. MR: 3558309 (cit. on p. 1891).

Jonathan Pila (2011). "O-minimality and the André-Oort conjecture for $\mathbb{C}^n$". *Ann. of Math. (2)* 173.3, pp. 1779–1840. MR: 2800724 (cit. on p. 1887).

Richard Pink (2005). "A combination of the conjectures of Mordell-Lang and André-Oort". In: *Geometric methods in algebra and number theory*. Vol. 235. Progr. Math. Birkhäuser Boston, Boston, MA, pp. 251–282. MR: 2166087 (cit. on p. 1899).

M. Raynaud (1983a). "Courbes sur une variété abélienne et points de torsion". *Invent. Math.* 71.1, pp. 207–233. MR: 688265 (cit. on p. 1887).

– (1983b). "Sous-variétés d'une variété abélienne et points de torsion". In: *Arithmetic and geometry, Vol. I*. Vol. 35. Progr. Math. Birkhäuser Boston, Boston, MA, pp. 327–352. MR: 717600 (cit. on p. 1887).

J. F. Ritt (1922). "Prime and composite polynomials". *Trans. Amer. Math. Soc.* 23.1, pp. 51–66. MR: 1501189 (cit. on p. 1891).

– (1923). "Permutable rational functions". *Trans. Amer. Math. Soc.* 25.3, pp. 399–448. MR: 1501252 (cit. on p. 1891).

Joseph H. Silverman (1992). "Variation of the canonical height on elliptic surfaces. I. Three examples". *J. Reine Angew. Math.* 426, pp. 151–178. MR: 1155751 (cit. on p. 1901).

– (1994a). "Variation of the canonical height on elliptic surfaces. II. Local analyticity properties". *J. Number Theory* 48.3, pp. 291–329. MR: 1293864 (cit. on p. 1901).

– (1994b). "Variation of the canonical height on elliptic surfaces. III. Global boundedness properties". *J. Number Theory* 48.3, pp. 330–352. MR: 1293865 (cit. on p. 1901).

– (1998). "The space of rational maps on $\mathbf{P}^1$". *Duke Math. J.* 94.1, pp. 41–77. MR: 1635900 (cit. on p. 1887).

L. Szpiro, E. Ullmo, and S. Zhang (1997). "Équirépartition des petits points". *Invent. Math.* 127.2, pp. 337–347. MR: 1427622 (cit. on pp. 1894, 1901).

Amaury Thuillier (2005). "Théorie du potentiel sur les courbes en géométrie analytique non archimédienne. Applications à la théorie d'Arakelov". PhD thesis. Université Rennes 1 (cit. on pp. 1896, 1901).

Jacob Tsimerman (2018). "The André-Oort conjecture for $\mathcal{A}_g$". *Ann. of Math. (2)* 187.2, pp. 379–390. MR: 3744855 (cit. on p. 1887).

Emmanuel Ullmo (1998). "Positivité et discrétion des points algébriques des courbes". *Ann. of Math. (2)* 147.1, pp. 167–179. MR: 1609514 (cit. on pp. 1887, 1894, 1901).

Xinyi Yuan (2008). "Big line bundles over arithmetic varieties". *Invent. Math.* 173.3, pp. 603–649. MR: 2425137 (cit. on pp. 1896, 1899, 1901).

Umberto Zannier (2012). *Some problems of unlikely intersections in arithmetic and geometry*. Vol. 181. Annals of Mathematics Studies. With appendixes by David Masser. Princeton University Press, Princeton, NJ, pp. xiv+160. MR: 2918151 (cit. on p. 1887).

Shou-Wu Zhang (1998a). "Equidistribution of small points on abelian varieties". *Ann. of Math. (2)* 147.1, pp. 159–165. MR: 1609518 (cit. on pp. 1887, 1894, 1901).

– (1998b). "Small points and Arakelov theory". In: *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*. Extra Vol. II, pp. 217–225. MR: 1648072 (cit. on p. 1901).

– (2006). "Distributions in algebraic dynamics". In: *Surveys in differential geometry. Vol. X*. Vol. 10. Surv. Differ. Geom. Int. Press, Somerville, MA, pp. 381–430. MR: 2408228 (cit. on p. 1887).

Shouwu Zhang (1995). "Small points and adelic metrics". *J. Algebraic Geom.* 4.2, pp. 281–300. MR: 1311351 (cit. on pp. 1892, 1898).

Michael E. Zieve (Oct. 2007). "Decompositions of Laurent polynomials". arXiv: 0710.1902 (cit. on p. 1891).

Received 2017-11-30.

LAURA DEMARCO
demarco@northwestern.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1905–1926)

NONHYPERBOLIC ERGODIC MEASURES

Lorenzo J. Díaz

Abstract

We discuss some methods for constructing nonhyperbolic ergodic measures and their applications in the setting of nonhyperbolic skew-products, homoclinic classes, and robustly transitive diffeomorphisms.

to Wellington de Melo in memoriam

1 The transitive and nonhyperbolic setting

Irrational rotations of the circle $\mathbb{T}^1$ and Anosov maps of the two-torus $\mathbb{T}^2$ are emblematic examples of *transitive* systems (existence of a dense orbit). Small perturbations of Anosov systems are also transitive. This property fails however for irrational rotations. Anosov diffeomorphisms are also paradigmatic examples of hyperbolic maps and, by definition, hyperbolicity persists by small perturbations. Our focus are systems which are *robustly transitive*. In dimension three or higher, there are important examples of those systems that fail to be hyperbolic. They are one of the main foci of this paper. A second focus is on nonhyperbolic elementary pieces of dynamics. We discuss how their lack of hyperbolicity is reflected at the ergodic level by the existence of nonhyperbolic ergodic measures. We also study how this influences the structure of the space of measures. In this discussion, we see how this sort of dynamics gives rise to robust cycles and blenders.

This research has been partially supported by CNE-Faperj and CNPq-grants (Brazil). The author warmly thanks J. Bochi, Ch. Bonatti, S. Crovisier, K. Gelfert, A. Gorodetski, Y. Ilyashenko, D. Kwietniak, J. Palis, M. Rams, A. Tahzibi, C. Vásques, and J. Yang for their useful comments and conversations.
MSC2010: primary 37D25; secondary 37D30, 28D20, 28D99.
Keywords: blender domination entropy, ergodic measure, heterodimensional cycle, homoclinic class, hyperbolicity, Lyapunov exponents, partial hyperbolicity, transitivity.

1.1 Paradigmatic examples. Let us start with an important class of examples, which are also the simplest ones, called *skew-products*. Consider the space $\Sigma_N = \{0, \dots, N-1\}^{\mathbb{Z}}$, $N \geq 2$, of bi-infinite sequences $\xi = (\xi_i)_{i \in \mathbb{Z}}$ endowed with the usual metric $d(\xi, \eta) \stackrel{\text{def}}{=} 2^{-n(\xi,\eta)}$, where $n(\xi, \eta) \stackrel{\text{def}}{=} \inf\{|\ell| : \xi_i \neq \eta_i \text{ for } i = -\ell, \dots, \ell\}$, and the *shift map* $\sigma : \Sigma_N \to \Sigma_N$, $\sigma(\xi_i) \stackrel{\text{def}}{=} (\xi'_i)$, $\xi'_i \stackrel{\text{def}}{=} \xi_{i+1}$. This map is transitive and has a dense subset of periodic points. Consider now a compact manifold K and a family of diffeomorphisms $f_\xi : K \to K$, $\xi \in \Sigma_N$, depending "nicely" on ξ. Associated to these maps we consider the *skew-product*

$$F : \Sigma_N \times K \to \Sigma_N \times K, \quad F(\xi, x) \stackrel{\text{def}}{=} (\sigma(\xi), f_\xi(x)). \tag{1-1}$$

The maps f_ξ are called *fiber maps*. The simplest case occurs when $f_\xi = f_{\xi_0}$ and then the system is called a *(one-)step skew-product*. There is a differentiable version of this model. Take (for instance) an Anosov diffeomorphism $A : \mathbb{T}^2 \to \mathbb{T}^2$ and fiber maps $f_X : K \to K$, $X \in \mathbb{T}^2$, depending "nicely" on X, and define

$$\Phi : \mathbb{T}^2 \times K \to \mathbb{T}^2 \times K, \quad \Phi(X, x) \stackrel{\text{def}}{=} (A(X), f_X(x)). \tag{1-2}$$

In many cases, these systems are an important source of nonhyperbolic and transitive dynamics. It may happen that these systems fail to be transitive, for instance, when f_ξ is the identity for all ξ. However, their appropriate perturbations are robustly transitive and nonhyperbolic examples, see Bonatti and Díaz [1996].

To continue with our discussion, we will introduce notions related to hyperbolicity and ergodicity. In what follows, M denotes a closed compact Riemannian manifold and $\mathrm{Diff}^1(M)$ the space of C^1-diffeomorphisms of M equipped with the C^1-uniform metric. Given $f \in \mathrm{Diff}^1(M)$, a closed set Λ is invariant if $f(\Lambda) = \Lambda$. A property is called *generic* if there is a residual subset of diffeomorphisms satisfying it. The phrase *"for generic diffeomorphisms it holds"* means *"there is a residual subset of diffeomorphisms such that..."*.

1.2 Weak forms of hyperbolicity. Let $f \in \mathrm{Diff}^1(M)$ and Λ be an f-invariant set. A Df-invariant splitting over Λ, $T_\Lambda M = E \oplus F$, is *dominated* if there are constants $C > 0$ and $\lambda < 1$ with[1]

$$||Df^{-n}|_{F_{f^n(x)}}|| \, ||Df^n|_{E_x}|| < C\lambda^n, \quad \text{for all } x \in \Lambda \text{ and } n \in \mathbb{N},$$

where $||\cdot||$ stands for the norm. The dimension of E is called the *index* of the splitting. A special type of dominated splitting is the hyperbolic one, when E is *uniformly contracting*

[1] The order of the bundles is relevant: the first one is the "most contracting" one.

($\|Df^n|_{E_x}\| < C\lambda^n$) and F is *uniformly expanding* ($\|Df^{-n}|_{F_{f^n(x)}}\| < C\lambda^n$). In such a case, we write $E = E^{\mathrm{s}}$ and $F = E^{\mathrm{u}}$ and call these bundles *stable* and *unstable* ones, respectively.

A Df-invariant splitting $T_\Lambda M = E_1 \oplus \cdots \oplus E_r$ with several bundles is *dominated* if for all $j \in \{1, \ldots, r-1\}$ the splitting $TM = E_1^j \oplus E_{j+1}^r$ is dominated, where $E_i^j \stackrel{\text{def}}{=} E_i \oplus \cdots \oplus E_j$, $i \le j$. When studying dominated splittings it is important to consider those with *undecomposable* bundles (i.e., the bundles cannot be decomposed further in a dominated way). Such a splitting is uniquely defined and called the *finest dominated splitting* of Λ, for details see Bonatti, Díaz, and Pujals [2003].

The set Λ is *hyperbolic* if there is a dominated splitting $T_\Lambda M = E^{\mathrm{s}} \oplus E^{\mathrm{u}}$ (one of these bundles may be trivial). The *index* of a transitive hyperbolic set Λ is the dimension of its stable bundle, denoted by $\mathrm{ind}(\Lambda)$. The set Λ is *partially hyperbolic* if there is a dominated splitting $T_\Lambda M = E^{\mathrm{s}} \oplus E \oplus E^{\mathrm{u}}$ (at most one of the bundles E^{s}, E^{u} may be trivial). For instance, if in (1-2) the rates of expansion and contraction of the maps f_X are "appropriate" then $\mathbb{T}^2 \times K$ is a partially hyperbolic set of Φ with a partially hyperbolic splitting with three nontrivial directions whose intermediate direction E has dimension $\dim(K)$.

1.3 Oseledecs' theorem and nonhyperbolicity. A measure μ is f-invariant if $\mu(A) = \mu(f^{-1}(A))$ for every Borel set A. We denote by $\mathcal{M}(f)$ the set of f-invariant probability measures and equip it with the weak$*$ topology. A measure $\mu \in \mathcal{M}(f)$ is *ergodic* if for every set B with $B = f^{-1}(B)$ it holds $\mu(B) \in \{0, 1\}$. We denote by $\mathcal{M}_{\mathrm{erg}}(f)$ the subset of $\mathcal{M}(f)$ of ergodic measures.

Given $\mu \in \mathcal{M}_{\mathrm{erg}}(f)$, by the Oseledecs Theorem (Oseledec [1968]) there are numbers $k = k(\mu) \in \{1, \ldots, \dim(M)\}$ and $\chi_1(\mu) < \chi_2(\mu) < \cdots < \chi_k(\mu)$, called the *Lyapunov exponents* of μ, and a Df-invariant splitting $F_1 \oplus F_2 \oplus \cdots \oplus F_k$, called the *Oseledets splitting of* μ, such that for μ-almost every point $x \in M$ it holds

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|Df_x^n(v)\| = \chi_i(\mu), \quad \text{for every } i \in \{1, \ldots, k\} \text{ and } v \in F_i \setminus \{\bar{0}\}.$$

The dimension of F_j is the *multiplicity* of the exponent $\chi_j(\mu)$. The number of negative exponents, counted with multiplicity, is the *index of* μ, denoted $\mathrm{ind}(\mu)$. The measure μ is *hyperbolic* if $\chi_j(\mu) \neq 0$ for every $j \in \{1, \ldots, k\}$. Otherwise, μ is called *nonhyperbolic*. If $\chi_{j(\mu)}(\mu) = 0$ then the dimension of $F_{j(\mu)}$ is the *number of zero exponents* of μ. Let j_{s} be the largest i with $\chi_i(\mu) < 0$ and j_{u} the smallest i with $\chi_i(\mu) > 0$. Note that either $j_{\mathrm{u}} = j_{\mathrm{s}} + 1$ (if the measure is hyperbolic) or $j_{\mathrm{u}} = j_{\mathrm{s}} + 2$ (otherwise). In this latter case, we let $j_{\mathrm{c}} = j_{\mathrm{s}} + 1$. We let $E^{\mathrm{cs}} \stackrel{\text{def}}{=} F_1 \oplus \cdots \oplus F_{j_{\mathrm{s}}}$, $E^{\mathrm{c}} \stackrel{\text{def}}{=} F_{j_{\mathrm{c}}}$, and $E^{\mathrm{cu}} \stackrel{\text{def}}{=} F_{j_{\mathrm{u}}} \oplus \cdots \oplus F_k$. In general, E^{cs} is not uniformly contracting and E^{cu} is not uniformly expanding.

Consider now μ ergodic and the finest dominated splitting $T_{\mathrm{supp}(\mu)} M = E_1 \oplus \cdots \oplus E_k$ over $\mathrm{supp}(\mu)$. By definition of domination, vectors in different bundles E_i have different exponents. Thus, every bundle F_j of the Oseledets splitting is contained in some E_{i_j}. The latter inclusion may be proper and then the Oseledets splitting is not dominated. We will see in Section 2.4 that the domination of the splitting and its type have dynamical consequences.

1.4 Nonhyperbolic settings. The *nonwandering set* $\Omega(f)$ of f is the set of points x such that for every neighborhood U of x there is some $n > 0$ with $f^n(U) \cap U \neq \varnothing$. The set $\Omega(f)$ is closed and f-invariant. When $\Omega(f)$ is hyperbolic we say that f is *hyperbolic*. A hyperbolic set is *nontrivial* if it contains some non-periodic orbit.

In the late 60s, Abraham and Smale exhibited open sets of diffeomorphisms consisting of nonhyperbolic ones, thus proving the non-density of hyperbolic systems, Abraham and Smale [1970]. Recall that the Kupka-Smale genericity theorem claims that periodic points of generic diffeomorphisms are hyperbolic and their invariant manifolds meet transversely. Note that the stable and unstable sets of nontrivial hyperbolic sets, in general, are not manifolds and hence we cannot speak of general position of these sets. The construction in Abraham and Smale [ibid.] shows an open set of diffeomorphisms such that the hyperbolic structures of two hyperbolic sets do not fit together nicely: invariant stable and unstable sets of (nontrivial) hyperbolic sets may not intersect "coherently". These non-coherent intersections are the germ of the notion of a *robust heterodimensional cycle* to be discussed in Section 2.2.

1.5 Robustly nonhyperbolic transitive diffeomorphisms. The construction in Abraham and Smale [ibid.] was followed by a series of examples of transitive diffeomorphisms which fail to be hyperbolic, see Shub [1971] and Mañé [1978] which fit into the class of systems nowadays called *robustly nonhyperbolic transitive diffeomorphisms.*

We say that $f \in \mathrm{Diff}^1(M)$ is C^1*-robustly transitive* if it has a C^1-neighborhood $\mathfrak{N}(f)$ such that every $g \in \mathfrak{N}(f)$ is transitive. We denote by $\mathbf{RTN}(M) \subset \mathrm{Diff}^1(M)$ the (open) set of robustly transitive and nonhyperbolic diffeomorphisms. A typical feature of these systems is the coexistence of saddles of different *indices* (dimension of the stable direction). These systems always exhibit a dominated splitting, Mañé [1982], Díaz, Pujals, and Ures [1999], and Bonatti, Díaz, and Pujals [2003], but they may fail to be partially hyperbolic, Bonatti and Viana [2000]. These findings showed the necessity of weaker notions of hyperbolicity such as partial hyperbolicity and dominated splitting, among others.

1.6 Hyperbolic flavors in nonhyperbolic dynamics. Although the examples described above are nonhyperbolic they do exhibit some "hyperbolic features". To start this discussion, recall that the set of f invariant probability measures $\mathcal{M}(f)$ is a Choquet simplex whose extremal elements are the ergodic measures. Density of ergodic measures in $\mathcal{M}(f)$

implies that either $\mathcal{M}(f)$ is a singleton or a nontrivial simplex whose extreme points are dense (*the* Poulsen simplex). Sigmund addressed the natural questions of the density of the ergodic measures in $\mathcal{M}(f)$ and the properties of generic invariant measures. Assuming that f is *Axiom A* (i.e., $\Omega(f)$ is hyperbolic and the periodic points are dense in $\Omega(f)$) he proved that the periodic measures (and thus the ergodic ones) are dense in $\mathcal{M}(f)$, Sigmund [1970]. Here a measure is *periodic* if it is the invariant probability measure supported on a periodic orbit. Moreover, the sets of ergodic measures and of measures with entropy zero are both residual in $\mathcal{M}(f)$. For an updated discussion and more references, see Gelfert and Kwietniak [n.d.].

It is now pertinent to recall the foundational talk by Mañé about ergodic properties of C^1-generic diffeomorphisms at the International Congress of Mathematicians of 1983, Mañé [1984]. Mañé proved that ergodic measures of C^1-generic diffeomorphisms are approached in the weak$*$ topology by periodic (and hence hyperbolic) measures, Mañé [1982]. Mañé's view of generic measures of C^1-diffeomorphisms was completed and substantially expanded in Abdenur, Bonatti, and Crovisier [2011] (see Abdenur, Bonatti, and Crovisier [ibid., Theorem 3.8]). In this context, another important result is Abdenur, Bonatti, and Crovisier [ibid., Theorem 3.5]: for every isolated transitive invariant set Λ of a C^1-generic diffeomorphism every generic measure supported in Λ is ergodic, hyperbolic, and its support is Λ. These results support the principle that, in the ergodic level, under a C^1-perspective hyperbolicity is somewhat widespread also in nonhyperbolic settings. However, it may be not necessarily ubiquitous if the viewpoint is changed. For instance, in the conservative setting nonhyperbolic dynamics can be robust (elliptic behavior in KAM theory) or locally generic (dichotomies all Lyapunov exponents are zero versus hyperbolicity, Bochi [2002] and Bochi and Viana [2002]). This discussion leads the following question

To what extent is the behavior of a generic dynamical system hyperbolic?

posed in Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [2005] and reformulated with different flavors in the literature (see, for example, the program in Palis [2008] and the conjectures in Pesin [2007]). In Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [2005] there is taken an ergodic point of view and "hyperbolicity" means that *all* ergodic measures are hyperbolic. The results in Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [ibid.] (see Section 3.1) inaugurated a fertile line of research about the construction of nonhyperbolic ergodic measures. Note that, by the Kupka-Smale genericity theorem, nonhyperbolic ergodic measures of generic diffeomorphisms have uncountable support. Also note that the hyperbolicity of $\Omega(f)$ implies the hyperbolicity of the ergodic measures. However, the converse is false in general as there are examples of nonhyperbolic diffeomorphisms whose all ergodic measures are hyperbolic. Nevertheless, all these examples are very specific and easily breakable. Thus, one hopes that the "big majority" of nonhyperbolic systems must exhibit nonhyperbolic ergodic measures which "truly" detect the nonhyperbolic behavior: capturing the whole dynamics (large support), the entropy of the system (large or positive entropy), and the number of nonhyperbolic directions (number of zero exponents). We will discuss these points in the next sections.

The generic measures investigated in Abdenur, Bonatti, and Crovisier [2011] were not studied in terms of their entropy[2]. There are several settings of nonhyperbolic chaotic systems which show "hyperbolic-like features from the entropy point of view". To justify this assertion, let us consider three-dimensional robustly nonhyperbolic transitive diffeomorphisms (see Section 1.5) with a partially hyperbolic splitting with one-dimensional central direction. There are three types of such diffeomorphisms: (compact case) having a global foliation consisting of circles tangent to the center direction (as the ones in Shub [1971], corresponding to systems as in (1-2)), (mixed case) having at least one invariant (or periodic) circle tangent to the center direction, Bonatti and Díaz [1996], and (non-compact case) without any invariant circle (certain derived from Anosov (DA) diffeomorphisms in Mañé [1978]). By Cowieson and Young [2005], these diffeomorphisms always have (ergodic) measures of maximal entropy (equal to the topological one).

We just discuss the "compact case". For skew-products as in (1-2) where $K = \mathbb{T}^1$, the entropy of the diffeomorphism is equal to the entropy of the base map and there is a C^1-open and dense subset of such systems having finitely many (ergodic) measures of maximal entropy, all hyperbolic, F. Rodriguez Hertz, M. A. Rodriguez Hertz, Tahzibi, and Ures [2012]. In some robustly transitive cases, there are exactly two such measures, Ures, Viana, and J. Yang [n.d.]. The spirit of these results is summarized in the following rigidity result in Tahzibi and J. Yang [n.d.] for partially hyperbolic diffeomorphisms with a central foliation by circles: *if there are high-entropy invariant measures with central Lyapunov exponent arbitrarily close to zero then the dynamics is conjugate to an isometric extension of an Anosov homeomorphism.* This result, based on the invariance principle in Avila and Viana [2010], holds for C^2-diffeomorphisms and involves some natural conditions such as dynamical coherence, transitive Anosov diffeomorphism in the base, and existence of global holonomies. See Díaz, Gelfert, and Rams [2017a] for results in the same spirit, just assuming C^1-regularity, in the skew-product setting,

To conclude, we see how hyperbolic features can also be found in the topology of $\mathcal{M}_{\mathrm{erg}}(f)$. For instance, under certain conditions (isolation and homoclinic relations, see Section 2.3) there are certain elementary components of the space of invariant measures with the same index which are Poulsen simplices and in which ergodic measures are *entropy-dense* (i.e., any measure can be approximated also in entropy by ergodic ones), see

[2]Quoting Mañé, Mañé [1984], "the generic elements of $\mathfrak{M}_{\mathrm{erg}}(f)$ fail to reflect the dynamic complexity of f". More precisely, generic measures of C^1-generic diffeomorphisms supported on a transitive isolated set have zero entropy, see Gelfert and Kwietniak [n.d., Theorem 8.1] and also Abdenur, Bonatti, and Crovisier [2011, Theorem 3.1].

Gorodetski and Pesin [2017]. Such elementary components are also studied in Bochi, Bonatti, and Gelfert [n.d.]. For an example where such components can be easier described, assume that there is a partially hyperbolic splitting $E^{s} \oplus E^{c} \oplus E^{u}$, where $\dim(E^{c}) = 1$ (as in (1-1) when $K = \mathbb{T}^1$). Then the index of any $\mu \in \mathcal{M}_{\text{erg}}(f)$ is either $\dim(E^{s})$ or $\dim(E^{s}) + 1$ and the exponent $\chi_{E^{c}}(\mu)$ is the only exponent of μ that can be zero. Thus, the space $\mathcal{M}_{\text{erg}}(f)$, splits as

$$\mathcal{M}_{\text{erg}}(f) = \mathcal{M}^{c}_{\text{erg},<0}(f) \,\dot{\cup}\, \mathcal{M}^{c}_{\text{erg},0}(f) \,\dot{\cup}\, \mathcal{M}^{c}_{\text{erg},>0}(f),$$

corresponding to the ergodic measures whose exponent $\chi_{E^{c}}(\mu)$ is negative, zero, and positive, respectively. Under additional "transitive-like" hypotheses, the components $\mathcal{M}^{c}_{\text{erg},<0}(f)$ and $\mathcal{M}^{c}_{\text{erg},>0}(f)$ both have the above mentioned properties and, moreover, any element in $\mathcal{M}^{c}_{\text{erg},0}(f)$ is approximated weak$*$ and in entropy by measures in either of the other two, see Díaz, Gelfert, and Rams [2017b]. Investigations in this direction can be also found in Bonatti and Zhang [n.d.(b)].

2 Robustly nonhyperbolic dynamics

In this section, we review the main ingredients and tools that appear in our nonhyperbolic setting. One general underlying theme is how, in our setting, nonhyperbolic dynamics forces the existence of robust cycles. To establish this, the main object is the blender. Recall first that, given $f \in \text{Diff}^1(M)$ and a hyperbolic set Λ_f of f, there is a C^1-neighborhood $\mathcal{N}(f) \subset \text{Diff}^1(M)$ such that every $g \in \mathcal{N}(f)$ has a hyperbolic set Λ_g (called the *continuation* of Λ_f) such that $g|_{\Lambda_g}$ is is conjugate to $f|_{\Lambda_f}$ and such that the map $g \mapsto \Lambda_g$ is continuous (this map is also uniquely defined). A special case occurs when the set is a periodic orbit.

2.1 Blenders. In very rough terms, a *blender* is a "semi-local plug" providing a hyperbolic set whose stable set "behaves" as a manifold of dimension greater than its index. Blenders were introduced in Bonatti and Díaz [1996], formalizing the arguments in Díaz [1995], to construct new types of robustly transitive diffeomorphisms. Blenders were also used in several contexts, for example generation of robust heterodimensional cycles and tangencies, stable ergodicity, construction of nonhyperbolic measures, among others. Each of these cases involves a specific type of blender. Here we follow the definition of a *geometrical blender* introduced in Bochi, Bonatti, and Díaz [2016].

We need some preliminary ingredients. The set $\mathfrak{D}^i(M)$ of i-dimensional (closed) discs C^1-embedded in M is endowed with a natural distance, Bochi, Bonatti, and Díaz [ibid., Section 3.1]. Given a family of disks $\mathfrak{D}$ in $\mathfrak{D}^i(M)$ we denote by $\mathcal{V}_\eta(\mathfrak{D})$ its η-neighborhood. The (non-empty) family $\mathfrak{D}$ is *strictly f-invariant* if for every $D \in \mathfrak{D}$ there is $\eta > 0$ such that the image $f(D)$ of any disk $D \in \mathcal{V}_\eta(\mathfrak{D})$ contains a disk of $\mathfrak{D}$.

We consider a transitive set Γ that is locally maximal in an open neighborhood V of it and is simultaneously hyperbolic and also partially hyperbolic,

$$\Gamma = \bigcap_{n\in\mathbb{Z}} f^n(\overline{V}), \qquad T_\Gamma M = E^{\mathrm{s}} \oplus E^{\mathrm{wu}} \oplus E^{\mathrm{uu}}, \qquad E^{\mathrm{u}} = E^{\mathrm{wu}} \oplus E^{\mathrm{uu}},$$

where each of the bundles E^{s}, E^{wu}, E^{uu} is nontrivial. The set Γ is a *dynamical* cu-*blender* if there are a strictly f-invariant family of discs $\mathfrak{D} \subset \mathfrak{D}^i(M)$, $i = \dim(E^{\mathrm{uu}})$, $\epsilon > 0$, and a strong unstable cone field $\mathcal{C}^{\mathrm{uu}}$ around E^{uu} such that every disc in $\mathcal{V}_\epsilon(\mathfrak{D})$ is contained in V and tangent to $\mathcal{C}^{\mathrm{uu}}$.

An important property of blenders is that they are C^1-robust: If Γ_f is a dynamical blender of f then for every g C^1-close to f the continuation Γ_g is also a blender. An important consequence of the definition of a blender is that its local stable manifold $W^{\mathrm{s}}_{\mathrm{loc}}(\Gamma_f)$ (i.e., the set of points whose forward orbits are contained in V) intersects every disk of the family $\mathfrak{D}$, see Bochi, Bonatti, and Díaz [2016, Section 3.4]. This property corresponds to the assertion above "a blender is a hyperbolic set whose local stable manifold behaves as a manifold of dimension $\dim(E^{\mathrm{s}}) + 1$".

2.2 Heterodimensional cycles. A pair of saddles of different indices, p_f and q_f, of a diffeomorphism f are related by a *heterodimensional cycle* if their invariant manifolds intersect cyclically. Suppose that $\mathrm{ind}(p_f) > \mathrm{ind}(q_f)$. Then, due to dimension deficiency, the intersection $W^{\mathrm{u}}(p_f, f) \cap W^{\mathrm{s}}(q_f, f)$ cannot be transverse, while, due to dimension sufficiency (the sum of the dimensions of these manifolds is bigger than $\dim(M)$), the intersection $W^{\mathrm{s}}(p_f, f) \cap W^{\mathrm{u}}(q_f, f)$ can be transverse (indeed, this is what happens "typically"). A (heterodimensional) cycle associated with transitive hyperbolic sets of different indices is defined similarly, just replacing the saddles by the corresponding transitive hyperbolic sets. By the Kupka-Smale genericity theorem, cycles associated with saddles cannot be robust, in our nonhyperbolic context they are ubiquitous and play a fundamental role (see comments below). In the spirit of Abraham and Smale [1970], we aim to get "robust cycles" and for that we need to consider cycles associated to nontrivial hyperbolic sets. Two transitive hyperbolic sets of different indices of f, Λ_f and Γ_f, have a *C^1-robust cycle* if there is a neighborhood $\mathcal{N}(f) \subset \mathrm{Diff}^1(M)$ such that for every $g \in \mathcal{N}(f)$ the invariant sets Λ_g and Γ_g intersect cyclically.

In the case when the saddles p_f and q_f satisfy $\mathrm{ind}(p_f) = \mathrm{ind}(q_f) + 1$, there are g C^1-arbitrarily close to f having a pair of transitive hyperbolic sets Λ_g and Γ_g with a robust cycle, see Bonatti and Díaz [2008]. However, these sets may be not related to the saddles in the initial cycle. If some of these saddles have a nontrivial homoclinic class then one can take these sets such that $\Lambda_g \ni p_g$ and $\Gamma_g \ni q_g$ and say that the cycle is C^1-*stabilized*, see Bonatti, Díaz, and Kiriki [2012, Theorem 1].

Let us explain the mechanism for robust cycles when $\dim(M) = 3$, $\mathrm{ind}(p_f) = 2$, and $\mathrm{ind}(q_f) = 1$. The unfolding of the cycle generates a blender Γ_g and the saddle p_g is "related" to Γ_g. This relation has two parts: first, $W^{\mathrm{s}}(p_g, g)$ intersects transversely $W^{\mathrm{u}}(\Gamma_g, g)$ (this is possible by dimension sufficiency), second $W^{\mathrm{u}}(p_g, g)$ contains a disk of the family $\mathfrak{D}$ and hence it intersects $W^{\mathrm{s}}_{\mathrm{loc}}(\Gamma_g, g)$.

2.3 Homoclinic relations and classes. A pair of hyperbolic periodic saddles p_f and q_f of a diffeomorphism f are *homoclinically related* if the invariant manifolds of the orbits $\mathcal{O}(p_f)$ and $\mathcal{O}(q_f)$ intersect transversely in a cyclic way. In particular, two saddles that are homoclinically related have the same index and their continuations (in a small neighborhood) are also homoclinically related.

The *homoclinic class* of a saddle p_f of f, denoted by $H(p_f, f)$, is the closure of the transverse intersections of the stable and unstable manifolds of $\mathcal{O}(p_f)$. The hyperbolic periodic points of f form a dense subset of $H(p_f, f)$. A homoclinic class is also a transitive set. A homoclinic class $H(p_f, f)$ may contain saddles of indices different from the one of p_f (thus these saddles cannot be homoclinically related to p_f) and hence may fail to be hyperbolic. In many relevant cases (as in the Axiom A case) the homoclinic classes are the "elementary pieces of dynamics", see Bonatti [2011, Sections 3 and 5] for an in-depth discussion.

Let us recall some properties of homoclinic classes. The map $p_f \mapsto H(p_f, f)$ is upper semi-continuous and hence this map is generically continuous. Also, C^1-generically, two homoclinic classes are either disjoint or equal, Carballo, Morales, and Pacifico [2003]. To recall an even stronger version of this result, consider saddles p_f and q_f of f, then there is a C^1-neighborhood $\mathfrak{N}(f)$ of f such that either it holds $H(p_g, g) = H(q_g, g)$ for every for C^1-generic $g \in \mathfrak{N}(f)$ or $H(p_g, g) \cap H(q_g, g) = \varnothing$ for every C^1-generic $g \in \mathfrak{N}(f)$, see for instance Abdenur, Bonatti, Crovisier, Díaz, and Wen [2007, Lemma 2.1]. In the first case, we say that p_f and q_f *are C^1-persistently in the same homoclinic class*.

Consider an open set $\mathcal{N} \subset \mathrm{Diff}^1(M)$ so that there is a saddle p_f such that $H(p_f, f)$ contains a saddle q_f with $\mathrm{ind}(p_f) = \mathrm{ind}(q_f) \pm 1$ for every $f \in \mathcal{N}$. The *connecting lemma,* Hayashi [1997], provides a dense subset $\mathfrak{D}$ of $\mathcal{N}$ such that p_f and q_f are involved in a cycle for every $f \in \mathfrak{D}$. Since $H(p_f, f)$ is nontrivial, the stabilization of cycles above gives an open and dense subset $\mathcal{O}$ of $\mathcal{N}$ such that every $f \in \mathcal{O}$ has a robust cycle associated to hyperbolic sets containing p_f and q_f.

2.4 Weak forms of hyperbolicity and homoclinic classes. The existence of a dominated splitting and its type provide important dynamical information. For instance, for homoclinic classes of C^1-generic diffeomorphisms there is the dichotomy "existence of a dominated splitting versus accumulation of the class by sinks or sources", Bonatti, Díaz, and Pujals [2003]. We highlight two results in this spirit.

The Oseledets splitting of an ergodic measure μ may be non-dominated, but if it is dominated then it has an extension to the whole support of μ. Recall that any hyperbolic measures whose Oseledets splitting is dominated is supported on the homoclinic class of a point of $\operatorname{ind}(\mu)$ (this is a version of the Anosov closing lemma, see Crovisier [2011, Proposition 1.4] based on Abdenur, Bonatti, and Crovisier [2011, Lemma 8.1]).

Proposition 2.1 (Cheng, Crovisier, Gan, Wang, and D. Yang [n.d., Proposition 1.4]). *There are the following possibilities for a nonhyperbolic homoclinic class $H(p_f, f)$ of a C^1-generic diffeomorphism:*

(i) Index-variability: *$H(p_f, f)$ contains saddles with different indices.*

(ii) *All periodic points of $H(p_f, f)$ have the same index k. There are two cases:*

(a) *There is a dominated splitting $T_{H(p_f,f)}M = E \oplus F$ of index k. Moreover, either $E = E^{\mathrm{s}}$ is uniformly contracting or E has a dominated splitting $E = E^{\mathrm{s}} \oplus E^{\mathrm{c}}$ where E^{s} is uniformly contracting and E^{c} is one-dimensional.*

(b) *There is no dominated splitting of $T_{H(p_f,f)}M$ of index k.*

Let us observe that all known examples of C^1-generic nonhyperbolic homoclinic classes fall into item i). In case ii.b) there are different sub-cases according to the different types of dominated splittings, for details see Cheng, Crovisier, Gan, Wang, and D. Yang [ibid., Proposition 1.4].

3 Tools to build nonhyperbolicity

In this section, we present two methods for constructing nonhyperbolic ergodic measures with uncountable support. In the first one, the measure is obtained as a limit of periodic measures. The second one provides a set of positive entropy supporting only nonhyperbolic measures. We also discuss "mixed" methods.

3.1 GIKN method. We present the GIKN-method introduced in Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [2005] guaranteeing the ergodicity (and non-triviality) of accumulation points of a sequence of periodic measures.

Given a periodic point p of period $\pi(p)$ of f, consider the periodic measure $\mu_{\mathcal{O}(p)}$ supported on the orbit $\mathcal{O}(p)$ of p. Consider a sequence of periodic points $(p_n)_n$ of f with increasing periods $\pi(p_n)$ and the measures $\mu_n = \mu_{\mathcal{O}(p_n)}$. Assume that in this process for every n there is a "large proportion" of points of $\mathcal{O}(p_{n+1})$ which are close to the previous orbit $\mathcal{O}(p_n)$ ("shadowing part"). A careful selection of these proportion times assures the convergence of μ_n, but without extra assumptions, there is a risk of obtaining a periodic measure as a limit. Thus, in the construction, it is also assumed that at each step there is also proportion of the orbit $\mathcal{O}(p_{n+1})$ that is far from the previous one ("tail part"). A careful choice of the proportion of the "shadowing" and "tail" parts forces that the limit measure is non-periodic. This involves some quantitative estimates. Given $\gamma, \tau > 0$, we say that the periodic orbit $\mathcal{O}(p)$ is a (γ, τ)-*good approximation* of $\mathcal{O}(q)$ is there are a subset $\Gamma \subset \mathcal{O}(p)$ and a (surjective) projection $\varrho \colon \Gamma \to \mathcal{O}(q)$ such that: (i) $\mathrm{dist}(f^i(x), f^i(\varrho(x)) < \gamma$ for every $x \in \Gamma$ and every $0 \le i \le \pi(q)$, (ii) $\#(\Gamma) \ge \tau \cdot \pi(p)$, and (iii) $\#(\varrho^{-1}(x))$ is independent of $x \in \mathcal{O}(q)$. Here $\#(A)$ denotes the cardinality of A. The next result corresponds to Bonatti, Díaz, and Gorodetski [2010, Lemmas 2.3 and 2.5], which reformulate Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [2005, Lemma 2 and Section 8].

Lemma 3.1 (Nontrivial ergodic limit of periodic measures). *Consider a sequence of periodic orbits $(p_n)_n$ of f with increasing periods. Suppose that there are sequences of strictly positive numbers $(\gamma_n)_n$ and $(\tau_n)_n$ such that for each n the orbit $\mathcal{O}(p_{n+1})$ is (γ_n, τ_n)-good approximation of $\mathcal{O}(p_n)$, where*

$$\sum_{n=1}^{\infty} \gamma_n < \infty \quad \text{and} \quad \prod_{n=1}^{\infty} \tau_n > 0.$$

Then $\mu_{\mathcal{O}(p_n)} \to \mu$ in the weak$*$ topology, where μ is ergodic, non-periodic, and

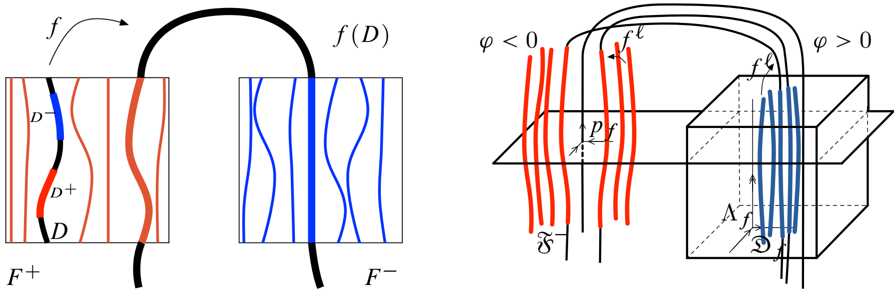
$$\mathrm{supp}(\mu) = \bigcap_{k=1}^{\infty} \left(\overline{\bigcup_{\ell=k}^{\infty} \mathcal{O}(p_\ell)} \right).$$

Assume now that there is a one-dimensional, Df-invariant, and continuous direction field E defined on the whole space such that $|\chi_E(\mu_{n+1})| < \alpha |\chi_E(\mu_n)|$, for some $\alpha \in (0,1)$. Then the limit measure satisfies $\chi_E(\mu) = 0$ and thus is nonhyperbolic, see Bonatti, Díaz, and Gorodetski [2010, Proposition 2.7]. Note that if the periodic orbits are constructed scattered throughout the whole manifold then the limit measure has full support. Note that the constructed measures have a somewhat repetitive pattern, as a result the limit measure has zero entropy, see Kwietniak and Łącka [n.d.] where this type of limit measures are studied and general results are obtained.

3.2 The flip-flop method.

Given a compact subset K of M and continuous map $\varphi \colon K \to \mathbb{R}$ we consider its *Birkhoff averages* defined by

$$\varphi_n(x) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(f^i(x)), \quad n \in \mathbb{N}, \quad \text{and} \quad \varphi_\infty(x) \stackrel{\text{def}}{=} \lim_{n \to \infty} \varphi_n(x),$$



A) Flip-flop family B) Flip-flop configuration

Figure 1: Flip-flops

provided this limit exists. In our setting, $\varphi = \log ||Df|_E||$, where E is a continuous Df-invariant line-field defined on K. Hence, $\varphi_\infty(x) = 0$ corresponds to $\chi_E(x) = 0$. We strive for a criterion implying that φ_∞ is zero uniformly on some compact set of positive entropy. For that we consider *controlled averages* and *flip-flop families*.

Let $\beta > 0$, $t \in \mathbb{N}$, and $T \in \mathbb{N} \cup \{\infty\}$. A point $x \in K$ is (β, t, T)-*controlled* if $\bigcup_{i=1}^{T} f^i(x) \subset K$ and if there is a subset $\mathcal{C} \subset \mathbb{N}$ of *control times* such that

- $0 \in \mathcal{C}$, $T \in \mathcal{C}$ if $T < \infty$, and $\mathcal{C}$ is infinite if $T = \infty$, and
- given $k < \ell$ two consecutive times in $\mathcal{C}$, then $\ell - k \leq t$ and $|\varphi_{\ell-k}(f^k(x))| \leq \beta$. The point x is *controlled at all scales* if there are sequences (t_i) of natural numbers and (β_i) of positive numbers, with $t_i \nearrow \infty$ and $\beta_i \searrow 0$, such that x is (β_i, t_i, T)-controlled for every i. The following holds, Bochi, Bonatti, and Díaz [2016, Lemma 2.2]: Assume that $x \in K$ is controlled at all scales and denote by $\omega(x)$ the ω-limit set of x. Then $\varphi_\infty(y) = 0$ for every $y \in \omega(x)$ and this limit is uniform on $\omega(x)$.

We now introduce the main ingredient to get "controlled" orbits. A family $\mathfrak{F} = \mathfrak{F}^+ \dot\cup \mathfrak{F}^-$ of compact subsets of K is called *flip-flop* if it satisfies the following properties (see Figure 1 A)):

- Let $F^+ \stackrel{\text{def}}{=} \bigcup_{D \in \mathfrak{F}^+} D$ and $F^- \stackrel{\text{def}}{=} \bigcup_{D \in \mathfrak{F}^-} D$. There is $\alpha > 0$ such that

$$\varphi(x^-) < -\alpha < 0 < \alpha < \varphi(x^+), \quad \text{for every } x^+ \in F^+ \text{ and } x^- \in F^-.$$

- For every $D \in \mathfrak{F}$ there are compact sets $D^+, D^- \subset D$ such that $f(D^+) \in \mathfrak{F}^+$ and $f(D^-) \in \mathfrak{F}^-$. Moreover, there is a constant $\lambda > 1$ such that

$$d(f(x), f(y)) \geq \lambda\, d(x, y) \quad \text{for every } x, y \in D^\pm.$$

The following result relates flip-flop families to zero Birkhoff averages: Given any D in a flip-flop family, there is $x \in D$ that is controlled at all scales and such that the restriction of f to $\omega(x)$ has positive topological entropy, Bochi, Bonatti, and Díaz [ibid., Theorem 2.1]. By the variational principle for entropy, the set $\omega(x)$ supports an ergodic measure of positive entropy. Let us explain some ingredients of this result.

An $\mathfrak{F}$*-segment of length* T is a sequence $\mathbf{D} = \{D_i\}_{i=0}^{T}$ such that $f(D_i) = D_{i+1}$, each D_i is contained in an element of $\mathfrak{F}$, and $D_T \in \mathfrak{F}$. Given $\beta > 0$ and $t \leq T$, we say that $\mathbf{D}$ is (β, t)*-controlled* if there exists a set of control times $\mathcal{C} \subset \{0, \dots, T\}$ containing 0 and T such that $|\varphi_{\ell-k}(f^k(x))| \leq \beta$ for every $x \in D_0$ and every pair $k < \ell$ of consecutive control times in $\mathcal{C}$.

We now relate flip-flop families, Birkhoff averages, and entropy. First, we encode orbits using their "itineraries". Let $\tau \in \mathbb{N}$. Given x in $F^+ \cup F^-$, $\mathbf{s} = (s_n) \in \{+,-\}^{\mathbb{N}}$, and $T \in \mathbb{N} \cup \{\infty\}$, the point x *follows the* τ*-pattern* $\mathbf{s}$ *up to time* T if $f^{n+1}(x) \in F^{s_n}$ for every $0 \leq n < T$ with $n = 0 \pmod{\tau}$. Fix $D \in \mathfrak{F}$, Bochi, Bonatti, and Díaz [ibid., Lemma 2.12], gives a sequence of $\mathfrak{F}$-segments $(\mathbf{D}_k^+)$, $\mathbf{D}_k^+ = \{D_i^k\}_{i=0}^{T_k^+}$, such that $D_0^k \subset D$ and every point in D_0^k is (β_i, t_i, T_k^+)-controlled for every i with $1 \leq i \leq k$, where $\beta_i \to 0$ and $t_i \to \infty$. For each k pick x_k in D_0^k and let $x_\infty \in D$ be any accumulation point of (x_k). By Bochi, Bonatti, and Díaz [ibid., Section 2.5.1], the point x_∞ is (β_i, t_i, ∞)-controlled for every i.

Fix any $\mathbf{s} \in \{+,-\}^{\mathbb{N}}$ with dense σ-orbit. For each k we select an $\mathfrak{F}$-segment $\mathbf{D}_k^+$ whose τ-pattern is $\mathbf{s}$, consider $x_k \in D_0^k$, and a limit point $x = x_\infty$. As $\omega(x) \subset \bigcap_{i \geq 0} f^{-\tau i}(F^+ \cup F^-)$, we can define the projection $\pi : \omega(x) \to \{-,+\}^{\mathbb{N}}$, $\pi(y) = \mathbf{s}(y)$, where $\mathbf{s}(y)$ is the itinerary of y (i.e., $f^{\tau i}(y) \in F^{s_i(y)}$). The map π is continuous and satisfies $\pi \circ f^\tau = \sigma \circ \pi$. Using that $\mathbf{s}$ has a dense orbit one gets that π is onto. Hence, the restriction of f^τ to $\omega(x)$ is semi-conjugate to the one-sided full-shift on 2 symbols and thus has positive entropy.

Finally, we observe that Bonatti, Díaz, and Bochi [n.d.] introduces a criterion relaxing the control at all scales-one, inspired by the GIKN-method and called *control at any scale with a long sparse tail.* It guarantees that any weak$*$ limit measure μ of a sequence of measures of the form $\nu_n(x_0) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=0}^{n-1} \delta_{f^i(x_0)}$, where δ_z is the Dirac measure at z, is such that μ-almost every point y has a dense orbit in the ambient space and $\varphi_\infty(y) = 0$. Again, the "role of the tail" is to spread the support of the measure. The construction implies that almost every ergodic measure η of the ergodic decomposition of μ satisfies $\int \varphi \, d\eta = 0$, see Bonatti, Díaz, and Bochi [ibid., Theorem 1].

4 Building nonhyperbolic measures and sets

We discuss some applications of the tools provided in Section 3.

4.1 Applications of the GIKN-method. We will see that there exist large (locally residual) sets in the space of C^1-diffeomorphisms of any compact manifold of dimension greater than 2 such that any diffeomorphisms in it has a nonhyperbolic ergodic measure (with zero entropy) that is the limit of periodic ones. In some cases, these measures have a zero exponent with multiplicity greater than 1. We start by discussing skew-products and see how the ingredients there pass on to the differentiable setting.

4.1.1 Step skew-products. Consider a skew-product as in (1-1) with $N = 2$, $F : \Sigma_2 \times \mathbb{T}^1 \to \Sigma_2 \times \mathbb{T}^1$, whose fiber maps are of the form $f_\xi = f_{\xi_0}$. The *iterated function system* associated to f_0 and f_1, $\mathrm{IFS}(f_0, f_1)$, is the set of maps g of the form $g = f_{\xi_{k-1}} \circ \cdots \circ f_{\xi_0}$ for some $(\xi_0, \ldots, \xi_{k-1}) \in \{0,1\}^k$ and some $k \geq 0$. The (forward) orbit of a point $x \in \mathbb{T}^1$ for $\mathrm{IFS}(f_0, f_1)$ is $\mathcal{O}^+(x) \stackrel{\text{def}}{=} \{g(x),\ g \in \mathrm{IFS}(f_0, f_1)\}$.

Theorem 4.1 (Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [2005, Theorem 2])**.** *Let $F : \Sigma_2 \times \mathbb{T}^1 \to \Sigma_2 \times \mathbb{T}^1$ be a step skew-product with fiber maps f_0 and f_1. Suppose that the following hypotheses hold:*
(i) (minimality) *for every $x \in \mathbb{T}^1$ the orbit $\mathcal{O}^+(x)$ is dense in $\mathbb{T}^1$,*
(ii) (expansion) *for every $x \in \mathbb{T}^1$ there is $g \in \mathrm{IFS}(f_0, f_1)$ with $|g'(x)| > 1$,*
(iii) (attracting fixed point) *there are $g \in \mathrm{IFS}(f_0, f_1)$ and $p \in \mathbb{T}^1$ with $g(p) = p$ and $|g'(p)| < 1$.*
Then F has a nonhyperbolic ergodic measure with full support.

Note that conditions (ii) and (iii) are open conditions and hence persistent by small perturbations. Although condition (i) is, *a priori*, non-open, in Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [ibid.] the authors provide an open set of pairs satisfying the conditions in the theorem.

To prove this theorem it is constructed a sequence of periodic orbits $\mathcal{O}(p_n)$ satisfying Lemma 3.1 and whose fiber Lyapunov exponents goes to zero. Hence the limit measure μ, $\mu_{\mathcal{O}(p_n)} \to \mu$, is ergodic and $\chi^c(\mu) = \lim_n \chi^c(\mu_{\mathcal{O}(p_n)}) = 0$ (as the Lyapunov exponents are given by integrals $\chi^c(\mu_{\mathcal{O}(p_n)}) = \int \log |f'_{\xi_i}|\, d\mu_{\mathcal{O}(p_n)}$).

The systems in Theorem 4.1 admit smooth realizations and thus provide diffeomorphisms with nonhyperbolic measures with uncountable support. In Kleptsyn and Nalskiĭ [2007] the methods above are used to get an open set of such systems.

To get some extra dynamical information behind the hypotheses of Theorem 4.1, let us replace condition (iii) by the slightly more restrictive condition (iii') f_0 is Morse-Smale with exactly two fixed points (say s attracting and n repelling). By minimality, the unstable set of the fixed point $S = (0^{\mathbb{Z}}, s)$ of F accumulates to the stable set of $N = (0^{\mathbb{Z}}, n)$. We also have that the stable set of $S = (0^{\mathbb{Z}}, s)$ meets the unstable set of $N = (0^{\mathbb{Z}}, n)$. Thinking of S and N as "hyperbolic saddles" of different indices, we have that S and N

are involved in an "heterodimensional quasi-cycle". This is the key observation to adapt the constructions in Gorodetskiĭ, Ilyashenko, Kleptsyn, and Nalskiĭ [2005] to the setting of homoclinic classes, see Section 4.1.3.

4.1.2 Multiple zero exponents in step skew-products. Theorem 4.1 was generalized in Bochi, Bonatti, and Díaz [2014] replacing the circle fiber by any compact manifold M, obtaining skew-products with nonhyperbolic ergodic measures with full support in the ambient space and whose fiber Lyapunov exponents are all equal to zero. The number of considered fiber maps is large to guarantee a condition called *maneuverability* (assuming $N \geq 2$ large). The property of full support is obtained by spreading the sequence of periodic points in the ambient space and using Lemma 3.1. Obtaining Lyapunov exponents which are all equal to zero is much more delicate. First, in higher dimensions, fiber exponents are not given by integrals (as in the one-dimensional case). Thus, the simultaneous convergence of all Lyapunov exponents along the orbits to zero does not imply the same property for the limit measure. A second problem is the loss of commutativity of products of matrices in higher dimensions.

These difficulties are bypassed in Bochi, Bonatti, and Díaz [ibid.] considering the induced skew-product on the *flag bundle* of M, $\mathfrak{F}_M$, defined as the set of the form $(x, F_1, \dots, F_{\dim(M)})$ where $x \in M$, F_i is a subspace of $T_x M$ of dimension i, and $F_1 \subset \cdots \subset F_{\dim(M)}$. To each $f \in \mathrm{Diff}^r(M)$, $r \geq 2$, there is associated the flag C^{r-1}-diffeomorphism

$$\mathfrak{F}_f : (x, F_1, \dots, F_{\dim(M)}) \mapsto \big(f(x), Df(x)(F_1), \dots, Df(x)(F_{\dim(M)})\big)$$

and to the skew-product F with the fiber maps $f_0, \dots, f_{N-1}$ the skew-product $\mathfrak{F}_F : \Sigma_N \times \mathfrak{F}_M \to: \Sigma_N \times \mathfrak{F}_M$ with the fiber maps $\mathfrak{F}_{f_0}, \dots, \mathfrak{F}_{f_{N-1}}$. The maneuverability condition implies the minimality of the flag iterated function system.

The natural projection from $\mathfrak{F}_M$ to M defines a fiber bundle. The projection of an $\mathfrak{F}_F$-invariant and ergodic probability measure ν on $\Sigma_N \times \mathfrak{F}_M$ is an F-invariant and ergodic measure μ. The fibered Lyapunov exponents of ν are given by linear functions of integrals of continuous maps (determinants). Hence these exponents vary continuously with respect to ν. Moreover, the exponents of ν and μ are related: all the Lyapunov exponents of ν are zero if and only if all the Lyapunov exponents of μ are zero. This allows to recover the continuity of the exponents, thus proving that all exponents of μ are zero.

4.1.3 Nonhyperbolic ergodic measures in homoclinic classes. We now consider nonhyperbolic homoclinic classes of C^1-generic diffeomorphisms and see that they support ergodic nonhyperbolic measures. Some of the ingredients of Section 4.1.1 will reappear:

Theorem 4.2 (Cheng, Crovisier, Gan, Wang, and D. Yang [n.d., Main Theorem]). *There is a residual subset* $\mathcal{R}$ *of* $\mathrm{Diff}^1(M)$ *such that every nonhyperbolic homoclinic class of* $f \in \mathcal{R}$ *supports ergodic nonhyperbolic measures.*

To see how the above result arises let us start by considering the simplest case where the class $H(p_f, f)$ contains a saddle q_f with $\mathrm{ind}(p_f) \neq \mathrm{ind}(q_f)$, a property which is called *index-variability (of the class)*. In view of Section 2.3, we can assume that there are an open set $\mathcal{U}$ and a residual subset $\mathcal{R} \subset \mathrm{Diff}^1(M)$ such that $H(p_f, f) = H(q_f, f)$ for all $f \in \mathcal{U} \cap \mathcal{R}$. To $H(p_f, f)$ we associate its set $\mathrm{ind}(H(p_f, f))$ of indices (i.e., $k \in \mathbb{N}$ such that here is a saddle $q_f \in H(p_f, f)$ with $\mathrm{ind}(q_f) = k$). By Abdenur, Bonatti, Crovisier, Díaz, and Wen [2007], this set is an interval in $\mathbb{N}$. Thus, we can assume that p_f and q_f have consecutive indices, say $s+1$ and s, respectively. Since the homoclinic class $H(p_f, f)$ is nontrivial one can replace these points by points homoclinically related to them such that the eigenvalues of $Df^{\pi(p_f)}(p_f)$ and $Df^{\pi(q_f)}(q_f)$ are all real and all have multiplicity one, see Abdenur, Bonatti, Crovisier, Díaz, and Wen [ibid., Proposition 2.3]. Thus, we will assume that p_f and q_f satisfy such properties.

We now review the arguments in Díaz and Gorodetski [2009] proving that every $f \in \mathcal{R} \cap \mathcal{U}$ has a non-periodic nonhyperbolic ergodic measure supported on $H(p_f, f)$. As observed in Section 2.3, there is a C^1-dense subset $\mathfrak{D}$ of $\mathcal{U}$ such that every $g \in \mathfrak{D}$ has a heterodimensional cycle associated to p_g and q_g. The fact that the saddles have real and simple eigenvalues implies that the dynamics "associated" to the cycle is partially hyperbolic with a one-dimensional central direction E. This direction is contracting in a neighborhood of $\mathcal{O}(p_f)$ and expanding in a neighborhood of $\mathcal{O}(q_f)$. The unfolding of this cycle generates new saddles r_h (also with real and simple eigenvalues) which are homoclinically related to p_h. The orbit of r_h stays most of the time close to $\mathcal{O}(p_h)$ (i.e., $\mathcal{O}(r_h)$ shadows $\mathcal{O}(p_h)$ most of the time), but it stays also some prescribed time nearby $\mathcal{O}(q_h)$ (this is the tail part of $\mathcal{O}(r_h)$). A consequence of the "tail part" is that $0 < |\chi_E(r_h)| < |\chi_E(p_h)|$. As r_h and p_h are homoclinically related, the classes of q_h and r_h coincide and after a new perturbation one can produce a cycle related to q_φ and r_φ. This gives an inductive pattern for the creation of saddles in $H(p_h, h)$ and with decreasing Lyapunov exponents. This provides a sequence of periodic orbits satisfying Lemma 3.1. In this construction, we are in a setting similar to the one in Section 4.1.1, the points p_f and q_f playing the roles of S and N, respectively. The difference between these constructions is that now the generation of periodic points involves perturbations, while in Section 4.1.1 it does not.

In the construction above, we first identified a part of the ambient space where the dynamics is partially hyperbolic and has a nonhyperbolic one-dimensional direction. This implies that if the homoclinic class has a undecomposable central bundle of dimension two then obtained measure cannot have full support in the class. On the other hand, if the homoclinic class has a dominated splitting $T_{H(p,f)}M = F_1 \oplus E \oplus F_2$ with $\dim(E) = 1$

and $H(p_f, f)$ has saddles of indices $\dim(F_1)$ and $\dim(F_1)+1$, then one can obtain nonhyperbolic measures μ with $\operatorname{supp}(\mu) = H(p_f, f)$. To get limit measures μ with $\operatorname{supp}(\mu) = H(p_f, f)$ one chooses the saddles r_h whose tails have iterates close to q_h but also iterates scattered throughout the whole class, Bonatti, Díaz, and Gorodetski [2010]. By Lemma 3.1 one gets $\operatorname{supp}(\mu) = H(p_f, f)$. This concludes the case when the class has index-variability.

We now consider the general case without *a priori* assuming index-variability. This will then complete the discussion of the proof of Theorem 4.2. The main idea is to recover the index-variability by "changing" the indices of saddles in the class by perturbation and then to fall into the previous case. Let us discuss the case (ii.a) in Proposition 2.1 where all the saddles of $H(p_f, f)$ have index k and $H(p_f, f)$ has a dominated splitting $E \oplus F$ adapted to $\operatorname{ind}(p_f) = k$.

Recall that if Λ is a compact f-invariant set and G is a one-dimensional invariant bundle over Λ that is not uniformly contracting, then there is an ergodic measure μ such that $\chi_G(\mu) \geq 0$, see Crovisier [2011, Claim 1.7]. By Proposition 2.1, there are two possibilities for the splitting $E \oplus F$ over $H(p_f, f)$, either $E = E^{\mathrm{s}} \oplus E^{\mathrm{c}}$ (with $\dim(E^{\mathrm{c}}) = 1$) or $E = E^{\mathrm{s}}$. If the first possibility holds then, by the previous comment, we get an ergodic measure μ with $\chi_E(\mu) \geq 0$. If $\chi_E(\mu) = 0$ we are done. Otherwise, the measure μ is hyperbolic with a dominated splitting $E^{\mathrm{s}} \oplus E^{\mathrm{cu}}$, $E^{\mathrm{cu}} = E^{\mathrm{c}} \oplus F$, and hence $\operatorname{supp}(\mu)$ is contained in a homoclinic class of index $k-1$, contradicting that all saddles of $H(p_f, f)$ have index k. If the second possibility $E = E^{\mathrm{s}}$ holds, we repeat the arguments above for f^{-1} (the case $F = E^{\mathrm{u}}$ does not occur as $H(p_f, f)$ is not hyperbolic).

Finally, we mention that there is the following more general version of Theorem 4.2 (see Cheng, Crovisier, Gan, Wang, and D. Yang [n.d., Theorem A]). *For a generic $f \in \operatorname{Diff}^1(M)$, if p_f is a saddle of index k and $T_{H(p_f,f)}M = E \oplus F$ is dominated, E not uniformly contracting and* $\dim(E) = k$, *and $H(p_f, f)$ does not contain saddles of index $k-1$ then there is an ergodic measure μ supported in $H(p_f, f)$ such that $\chi_k(\mu) = 0$.*

Under these conditions, Wang [n.d., Theorem A] claims that for every $\epsilon > 0$ there is a saddle q_ϵ homoclinically related to p_f and such that $|\chi_k(q_\epsilon)| < \epsilon$. Then, after a small C^1-perturbation, one can change its index and generate a cycle related the saddle q_h of index $k-1$ and p_h whose unfolding generates saddles of index $k-1$ inside $H(p_g, g)$. Thus, we recover the index-variability scenario above.

4.1.4 Ergodic measures with multiple zero exponents.

We now study generic homoclinic classes supporting nonhyperbolic ergodic measures with several zero exponents, which corresponds to the results in Section 4.1.2 in the differentiable setting.

As explained in Section 1.2, if the nonhyperbolic direction splits in a dominated way into one-dimensional sub-bundles then nonhyperbolic measures have exactly one zero exponent. Thus, in what follows, we consider homoclinic classes having a higher-dimensional and undecomposable central direction. More precisely, consider generic diffeomorphisms f having a saddle p_f whose homoclinic class have a dominated splitting $T_{H(p,f)}M = E \oplus E^{\mathrm{c}} \oplus F$, where $\operatorname{ind}(p_f) = \dim(E) = k$ and such that the class contains a saddle q_f of index $\dim(E \oplus E^{\mathrm{c}})$ and E^{c} is undecomposable. By the results in Section 4.1.3, generically, for each j with $\dim(E) < j \le \dim(E \oplus E^{\mathrm{c}})$ there is a nonhyperbolic ergodic measure μ with $\chi_j(\mu) = 0$. Furthermore, Wang and Zhang [n.d., Corollary 1.1] claims that under these conditions there is a nonhyperbolic ergodic measure such that $\chi_{E^{\mathrm{c}}}(\mu) = 0$ (i.e., with $\dim(E^{\mathrm{c}})$ zero exponents) and also points out that the index-variability conditions are necessary (see Wang and Zhang [ibid., Section 5]). The proof of this theorem is based on the GIKN-method and all the comments in Section 4.1.2 about the difficulties to pass from dimension one to higher dimensions apply here. To discuss this result, recall first the extremely useful classical Franks' lemma for C^1-dynamics: for every (small) perturbation of the derivative of a diffeomorphism along a periodic orbit there is a small local C^1-perturbation of the diffeomorphism with such a derivative along the orbit. This result shows the importance of understanding the "perturbations of the linear part" of the dynamics.

A first ingredient in Wang and Zhang [ibid.] is Bochi and Bonatti [2012, Theorem 4.7] about perturbations of linear cocycles. Consider matrices $A_1, \dots, A_n \in \mathrm{GL}(n, \mathbb{R})$ such that $B = A_n \circ \dots \circ A_1$ has no dominated splitting of index i. Then there is an arbitrarily small perturbations A_i' of these matrices such that the Lyapunov exponents of $B' = A_n' \circ \dots \circ A_1'$ satisfy $\chi_i(B') = \chi_{i+1}(B')$ and $\chi_j(B) = \chi_j(B')$ if $j \notin \{i, i+1\}$. Thus, one gets a dynamics whose linear part has exponents of multiplicity two. A second argument is the improvement of Franks' lemma presented in Gourmelon [2016]: the perturbation can be done preserving parts of stable and unstable manifolds. This allows to preserve intersections of invariant manifolds throughout the perturbations.

We emphasize two ingredients in Wang and Zhang [n.d., Corollary 1.1]: i) a version of Lemma 3.1, including a comparison of the "central" Lyapunov exponents and the periodic points "heteroclinic-like" related, and ii) similar to arguments in Section 4.1.2, the exponents of the measures are linear functions of integrals of continuous maps, providing continuity of the exponents with respect to the measures.

4.2 Applications of the flip-flop method: robust zeros. In comparison with the results presented in Section 4.1, in what follows we replace locally residual sets of C^1-diffeomorphism by open ones and also obtain measures with positive entropy. Recall the definition of the set $\mathbf{RTN}(M)$ of C^1-robustly transitive and nonhyperbolic diffeomorphisms in Section 1.5.

Theorem 4.3 (Bochi, Bonatti, and Díaz [2016, Corollary 1]). *There is a C^1-open and dense subset of $\mathbf{RTN}(M)$ of diffeomorphisms with a nonhyperbolic ergodic measure with positive entropy.*

The proof of this result uses the method of controlling Birkhoff averages on sets of positive entropy and relies on finding an appropriate flip-flop family. The main ingredients for that are heterodimensional cycles and blenders.

First, there is an open and dense subset of $\mathbf{RTN}(M)$ consisting of diffeomorphisms f having a saddle p_f (assume that $\pi(p_f) = 1$) of index $s+1$ and a blender Λ_f of index s which are "heteroclinically related": $W^{\mathrm{u}}(p_f, f)$ contains a disk of the strictly f-invariant family of disks $\mathfrak{D}_f$ of the blender and $W^{\mathrm{s}}(\Lambda_f, f)$ and $W^{\mathrm{s}}(p_f, f)$ have transverse intersections. Indeed, this means that p_f and Λ_f are involved in a robust cycle. We can also assume that the eigenvalues of $Df(p_f)$ are all real and all have multiplicity one (recall Section 4.1.3). We can now identify heteroclinic orbits $\mathcal{O}_1$ (going from Λ_f to p_f) and $\mathcal{O}_2$ (going from p_f to Λ_f) in such a way in a neighborhood of $\mathcal{O}_1 \cup \mathcal{O}_2 \cup \Lambda_f \cup \mathcal{O}(p_f)$ the dynamics is partially hyperbolic with a one-dimensional center direction E (that is contracting nearby p_f and expanding nearby Λ_f). See Figure 1 B).

The above *flip-flop* configuration yields flip-flop families associated to $\varphi = \log ||Df^m|_E||$ and some power of f, Bochi, Bonatti, and Díaz [ibid., Section 4]. The flip-flop family $\mathfrak{F} = \mathfrak{F}^- \dot\cup \mathfrak{F}^+$ is (roughly) defined as follows: $\mathfrak{F}^+$ is the family of f-invariant disks $\mathfrak{D}_f$ of the blender and $\mathfrak{F}^-$ is a small neighborhood of of $W^{\mathrm{s}}_{\mathrm{loc}}(p_f, f)$. The heteroclinic connection between the blender and the saddle implies that $\mathfrak{F}$ is a flip-flop family for some f^ℓ. Theorem 4.3 follows applying Section 3.2.

Denote by $\mathbf{RTN}^{\mathrm{c}=1}(M)$ the open subset of $\mathbf{RTN}(M)$ with a partially hyperbolic $TM = E^{\mathrm{s}} \oplus E^{\mathrm{c}} \oplus E^{\mathrm{u}}$ with three nontrivial bundles and $\dim(E^{\mathrm{c}}) = 1$. In Bonatti, Díaz, and Bochi [n.d.] and Bonatti and Zhang [n.d.(a)] it is proved that C^1-open and densely in $\mathbf{RTN}^{\mathrm{c}=1}(M)$ the diffeomorphisms have nonhyperbolic ergodic measures with full support. The method in Bonatti, Díaz, and Bochi [n.d.] uses the control at all scales with a long sparse tail criterion, while Bonatti and Zhang [n.d.(a)] uses a cocktail combining the methods in Sections 3.1 and 3.2 and a version of the shadowing lemma in Gan [2002]. Finally, in a work in progress with Bonatti and Kwietniak (Łącka announced a similar result by different methods) we prove that these measures can be obtained with full support and positive entropy.

References

Flavio Abdenur, Christian Bonatti, S. Crovisier, Lorenzo J. Díaz, and L. Wen (2007). "Periodic points and homoclinic classes". *Ergodic Theory Dynam. Systems* 27.1, pp. 1–22. MR: 2297084 (cit. on pp. 1913, 1920).

Flavio Abdenur, Christian Bonatti, and Sylvain Crovisier (2011). "Nonuniform hyperbolicity for C^1-generic diffeomorphisms". *Israel Journal of Math.* 183, pp. 1–60. MR: 2811152 (cit. on pp. 1909, 1910, 1914).

R. Abraham and S. Smale (1970). "Nongenericity of Ω-stability". In: *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*. Amer. Math. Soc., Providence, R.I., pp. 5–8. MR: 0271986 (cit. on pp. 1908, 1912).

Artur Avila and Marcelo Viana (2010). "Extremal Lyapunov exponents: an invariance principle and applications". *Invent. Math.* 181.1, pp. 115–189. MR: 2651382 (cit. on p. 1910).

Jairo Bochi (2002). "Genericity of zero Lyapunov exponents". *Ergodic Theory Dynam. Systems* 22.6, pp. 1667–1696. MR: 1944399 (cit. on p. 1909).

Jairo Bochi and Christian Bonatti (2012). "Perturbation of the Lyapunov spectra of periodic orbits". *Proc. Lond. Math. Soc. (3)* 105.1, pp. 1–48. MR: 2948788 (cit. on p. 1922).

Jairo Bochi, Christian Bonatti, and Lorenzo J. Díaz (2014). "Robust vanishing of all Lyapunov exponents for iterated function systems". *Math. Z.* 276.1-2, pp. 469–503. MR: 3150214 (cit. on p. 1919).

– (2016). "Robust criterion for the existence of nonhyperbolic ergodic measures". *Comm. Math. Phys.* 344.3, pp. 751–795. MR: 3508160 (cit. on pp. 1911, 1912, 1916, 1917, 1923).

Jairo Bochi, Christian Bonatti, and Katrin Gelfert (n.d.). "Dominated Pesin theory: convex sum of hyperbolic measures". To appear in *Israel J. Math.* (cit. on p. 1911).

Jairo Bochi and Marcelo Viana (2002). "Uniform (projective) hyperbolicity or no hyperbolicity: a dichotomy for generic conservative maps". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 19.1, pp. 113–123. MR: 1902547 (cit. on p. 1909).

Chrisitian Bonatti, Lorenzo J. Díaz, and Jairo Bochi (n.d.). "A criterion for zero averages and full support of ergodic measures". To appear in *Mosc. Math. J* (cit. on pp. 1917, 1923).

Christian Bonatti (2011). "Survey: Towards a global view of dynamical systems, for the C^1-topology". *Ergodic Theory Dynam. Systems* 31.4, pp. 959–993. MR: 2818683 (cit. on p. 1913).

Christian Bonatti and Lorenzo J. Díaz (1996). "Persistent nonhyperbolic transitive diffeomorphisms". *Ann. of Math. (2)* 143.2, pp. 357–396. MR: 1381990 (cit. on pp. 1906, 1910, 1911).

– (2008). "Robust heterodimensional cycles and C^1-generic dynamics". *J. Inst. Math. Jussieu* 7.3, pp. 469–525. MR: 2427422 (cit. on p. 1912).

Christian Bonatti, Lorenzo J. Díaz, and Anton Gorodetski (2010). "Non-hyperbolic ergodic measures with large support". *Nonlinearity* 23.3, pp. 687–705. MR: 2593915 (cit. on pp. 1915, 1921).

Christian Bonatti, Lorenzo J. Díaz, and S. Kiriki (2012). "Stabilization of heterodimensional cycles". *Nonlinearity* 25.4, pp. 931–960. MR: 2891553 (cit. on p. 1913).

Christian Bonatti, Lorenzo J. Díaz, and E. R. Pujals (2003). "A C^1-generic dichotomy for diffeomorphisms: weak forms of hyperbolicity or infinitely many sinks or sources". *Ann. of Math. (2)* 158.2, pp. 355–418. MR: 2018925 (cit. on pp. 1907, 1908, 1914).

Christian Bonatti and Marcelo Viana (2000). "SRB measures for partially hyperbolic systems whose central direction is mostly contracting". *Israel J. Math.* 115, pp. 157–193. MR: 1749677 (cit. on p. 1908).

Christian Bonatti and Jinhua Zhang (n.d.[a]). "On the existence of non-hyperbolic ergodic measures as the limit of periodic measures". To appear in *Ergodic Theory Dynam. Systems* (cit. on p. 1923).

– (n.d.[b]). "Periodic measures and partially hyperbolic homoclinic classes". arXiv: 1609.08489 (cit. on p. 1911).

C. M. Carballo, C. A. Morales, and M. J. Pacifico (2003). "Homoclinic classes for generic C^1 vector fields". *Ergodic Theory Dynam. Systems* 23.2. MR: 1972228 (cit. on p. 1913).

Cheng Cheng, Sylvain Crovisier, Shaboo Gan, Xiadong Wang, and Dawei Yang (n.d.). "Hyperbolicity versus non-hyperbolic ergodic measures inside homoclinic classes". To appear in *Ergodic Theory Dynam. Systems* (cit. on pp. 1914, 1920, 1921).

William Cowieson and Lai-Sang Young (2005). "SRB measures as zero-noise limits". *Ergodic Theory Dynam. Systems* 25.4, pp. 1115–1138. MR: 2158399 (cit. on p. 1910).

Sylvain Crovisier (2011). "Partial hyperbolicity far from homoclinic bifurcations". *Adv. Math.* 226.1, pp. 673–726. MR: 2735772 (cit. on pp. 1914, 1921).

Lorenzo J. Díaz (1995). "Robust nonhyperbolic dynamics and heterodimensional cycles". *Ergodic Theory Dynam. Systems* 15.2, pp. 291–315. MR: 1332405 (cit. on p. 1911).

Lorenzo J. Díaz, K. Gelfert, and M. Rams (2017a). "Nonhyperbolic step skew-products: Ergodic approximation". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 34.6, pp. 1561–1598. MR: 3712011 (cit. on p. 1910).

– (2017b). "Topological and ergodic aspects of partially hyperbolic diffeomorphisms and nonhyperbolic step skew products". *Tr. Mat. Inst. Steklova* 297.Order and Chaos in Dynamical Systems, pp. 113–132. MR: 3695409 (cit. on p. 1911).

Lorenzo J. Díaz and Anton Gorodetski (2009). "Non-hyperbolic ergodic measures for non-hyperbolic homoclinic classes". *Ergodic Theory Dynam. Systems* 29.5, pp. 1479–1513. MR: 2545014 (cit. on p. 1920).

Lorenzo J. Díaz, Enrique R. Pujals, and Raúl Ures (1999). "Partial hyperbolicity and robust transitivity". *Acta Math.* 183.1, pp. 1–43. MR: 1719547 (cit. on p. 1908).

Shaobo Gan (2002). "A generalized shadowing lemma". *Discrete Contin. Dyn. Syst.* 8.3, pp. 627–632. MR: 1897871 (cit. on p. 1923).

Katrin Gelfert and Dominik Kwietniak (n.d.). "On density of ergodic measures and generic points". To appear in *Ergodic Theory Dynam. Systems* (cit. on pp. 1909, 1910).

Anton Gorodetski and Yakov Pesin (2017). "Path connectedness and entropy density of the space of hyperbolic ergodic measures". In: *Modern theory of dynamical systems*. Vol. 692. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 111–121. MR: 3666070 (cit. on p. 1911).

A. S. Gorodetskiĭ, Yu. S. Ilyashenko, V. A. Kleptsyn, and M. B. Nalskiĭ (2005). "Nonremovability of zero Lyapunov exponents". *Funktsional. Anal. i Prilozhen.* 39.1, pp. 27–38, 95. MR: 2132437 (cit. on pp. 1909, 1914, 1915, 1918, 1919).

Nikolaz Gourmelon (2016). "A Franks' lemma that preserves invariant manifolds". *Ergodic Theory Dynam. Systems* 36.4, pp. 1167–1203. MR: 3492974 (cit. on p. 1922).

Shuhei Hayashi (1997). "Connecting invariant manifolds and the solution of the C^1 stability and Ω-stability conjectures for flows". *Ann. of Math. (2)* 145.1, pp. 81–137. MR: 1432037 (cit. on p. 1913).

V. A. Kleptsyn and M. B. Nalskiĭ (2007). "Stability of the existence of nonhyperbolic measures for C^1-diffeomorphisms". *Funktsional. Anal. i Prilozhen.* 41.4, pp. 30–45, 96. MR: 2411604 (cit. on p. 1918).

D. Kwietniak and M. Łącka (n.d.). "Feldman-Katok pseudometric and the GIKN construction of nonhyperbolic ergodic measures". arXiv: 1702.01962 (cit. on p. 1915).

Ricardo Mañé (1978). "Contributions to the stability conjecture". *Topology* 17.4, pp. 383–396. MR: 516217 (cit. on pp. 1908, 1910).

– (1982). "An ergodic closing lemma". *Ann. of Math. (2)* 116.3, pp. 503–540. MR: 678479 (cit. on pp. 1908, 1909).

– (1984). "Oseledec's theorem from the generic viewpoint". In: *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*. PWN, Warsaw, pp. 1269–1276. MR: 804776 (cit. on pp. 1909, 1910).

V. I. Oseledec (1968). "A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems". *Trudy Moskov. Mat. Obšč.* 19, pp. 179–210. MR: 0240280 (cit. on p. 1907).

J. Palis (2008). "Open questions leading to a global perspective in dynamics". *Nonlinearity* 21.4, T37–T43. MR: 2399817 (cit. on p. 1909).

Ya. Pesin (2007). "Existence and genericity problems for dynamical systems with nonzero Lyapunov exponents". *Regul. Chaotic Dyn.* 12.5, pp. 476–489. MR: 2350334 (cit. on p. 1909).

Federico Rodriguez Hertz, María Alejandra Rodriguez Hertz, Ali Tahzibi, and Raúl Ures (2012). "Maximizing measures for partially hyperbolic systems with compact center leaves". *Ergodic Theory Dynam. Systems* 32.2, pp. 825–839. MR: 2901373 (cit. on p. 1910).

M. Shub (1971). "Topologically transitive diffeomorphisms on $\mathbb{T}^4$". In: *Proceedings of the Symposium on Differential Equations and Dynamical Systems*. Vol. 206. Lecture Notes in Math. Springer, Berlin, p. 39 (cit. on pp. 1908, 1910).

Karl Sigmund (1970). "Generic properties of invariant measures for Axiom A diffeomorphisms". *Invent. Math.* 11, pp. 99–109. MR: 0286135 (cit. on p. 1909).

Ali Tahzibi and Jiagang Yang (n.d.). "Strong hyperbolicity of ergodic measures with large entropy". To appear in *Trans. Amer. Math. Soc.* (cit. on p. 1910).

Raúl Ures, Marcelo Viana, and Jiagang Yang (n.d.). "Maximal measures of diffeomorphisms with circle fiber bundle". In preparation (cit. on p. 1910).

Xiadong Wang (n.d.). "Hyperbolicity versus weak periodic orbits inside homoclinic classes". arXiv: 1504.03153 (cit. on p. 1921).

Xiadong Wang and Jinhua Zhang (n.d.). "Ergodic measures with multi-zero Lyapunov exponents inside homoclinic classes". arXiv: 1604.03342 (cit. on p. 1922).

Received 2017-11-30.

LORENZO J. DÍAZ
DEPARTAMENTO DE MATEMÁTICA PUC-RIO
MARQUÊS DE SÃO VICENTE 225
RIO DE JANEIRO 22451-900
BRAZIL
lodiaz@mat.puc-rio.br

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1927–1950)

SOME QUESTIONS AROUND QUASI-PERIODIC DYNAMICS

Bassam Fayad and Raphaël Krikorian

Abstract

We propose in these notes a list of some old and new questions related to quasi-periodic dynamics. A main aspect of quasi-periodic dynamics is the crucial influence of arithmetics on the dynamical features, with a strong duality in general between Diophantine and Liouville behavior. We will discuss rigidity and stability in Diophantine dynamics as well as their absence in Liouville ones. Beyond this classical dichotomy between the Diophantine and the Liouville worlds, we discuss some unified approaches and some phenomena that are valid in both worlds. Our focus is mainly on low dimensional dynamics such as circle diffeomorphisms, disc dynamics, quasi-periodic cocycles, or surface flows, as well as finite dimensional Hamiltonian systems.

In an opposite direction, the study of the dynamical properties of some diagonal and unipotent actions on the space of lattices can be applied to arithmetics, namely to the theory of Diophantine approximations. We will mention in the last section some problems related to that topic.

The field of quasi-periodic dynamics is very extensive and has a wide range of interactions with other mathematical domains. The list of questions we propose is naturally far from exhaustive and our choice was often motivated by our research involvements.

1 Arithmetic conditions

A vector $\alpha \in \mathbb{R}^d$ is *non-resonant* if it has rationally independent coordinates: for all $(k_1 \dots, k_d) \in \mathbb{Z}^d$, the identity $\sum_{i=1}^d k_i \alpha_i = 0$ implies $k_i = 0$ for $i = 1, \dots, d$; otherwise, it is called *resonant*.

BF and RK are supported by project ANR BEKAM: ANR-15-CE40-0001. RK is supported by a Chaire d'Excellence LABEX MME-DII..

MSC2010: 37C55.

For $\gamma, \sigma > 0$, we define the set $DC_d(\gamma,\sigma) \subset \mathbb{R}^d$ of *diophantine* vectors with *exponent* σ and *constant* γ as the set of $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$ such that

$$\forall (k_1, \dots, k_d) \in \mathbb{Z}^d,\ |\sum_{i=1}^{d} k_i \alpha_i| \geq \frac{\gamma}{(\sum_{i=1}^{d} |k_i|)^\sigma}; \tag{1-1}$$

we then set $DC_d(\sigma) = \bigcup_{\gamma>0} DC(\gamma,\sigma)$, $DC_d = \bigcup_{\sigma>0} DC(\sigma)$. For each fixed $\sigma > d$ and γ small enough the set $DC(\gamma,\sigma)$ has positive Lebesgue measure in the unit ball of $\mathbb{R}^d$ and the Lebesgue measure of its complement goes to zero as γ goes to zero. Thus the sets $DC(\sigma), \sigma > d$ and DC_d have full Lebesgue measure in $\mathbb{R}^d$. The set DC_d is the set of *Diophantine* vectors of $\mathbb{R}^d$ while its complement in the set of non-resonant vectors is called the set of *Liouville* vectors.

For a translation vectors of $\mathbb{T}^d$ defined as $\alpha + \mathbb{Z}^d$, $\alpha \in \mathbb{R}^d$, we say that it is resonant, Diophantine or Liouville, if the $\mathbb{R}^{d+1}$ vector $(1,\alpha)$ is resonant, Diophantine or Liouville respectively.

2 Diffeomorphisms of the circle and the torus

For $k \in \mathbb{N} \cup \{\infty, \omega\}$ we define $\mathrm{Diff}_0^k(\mathbb{T}^d)$ as the set of orientation preserving homeomorphisms of $\mathbb{T}^d$ of class C^k together with their inverse. To any $f \in \mathrm{Diff}_0^0(\mathbb{T}^d)$ one can associate its *rotation set* $\rho(f) := \{\int_{\mathbb{T}} (\bar{f} - id) d\mu, \mu \in \mathfrak{M}(f)\} \mod \mathbb{Z}^d$ where $\bar{f} : \mathbb{R}^d \to \mathbb{R}^d$ is a lift of f and $\mathfrak{M}(f)$ is the set of all f-invariant probability measures on $\mathbb{T}^d$. Let $T_\alpha : \mathbb{T}^d \to \mathbb{T}^d$ be the translation $x \mapsto x + \alpha$, $\mathfrak{F}_\alpha^k(\mathbb{T}^d) = \{f \in \mathrm{Diff}_0^k(\mathbb{T}^d),\ \rho(f) = \{\alpha\}\}$, $\mathfrak{O}_\alpha^k(\mathbb{T}^d) = \{h \circ T_\alpha \circ h^{-1}, h \in \mathrm{Diff}_0^k(\mathbb{T}^d)\}$. We say that $f \in \mathrm{Diff}_0^\infty(\mathbb{T}^d)$ is *almost reducible* if there exists a sequence $(h_n)_{n\in\mathbb{N}} \in (\mathrm{Diff}_0^\infty(\mathbb{T}^d))^{\mathbb{N}}$ such that $h_n \circ f \circ h_n^{-1}$ converges in the C^∞-topology to T_α. When $d = 1$, $\rho(f)$ is reduced to a single element and we denote by $\rho(f)$ this element. By Denjoy Theorem, any $f \in \mathrm{Diff}_0^k(\mathbb{T})$ with $k \geq 2$, is conjugated by an orientation preserving homeomorphism to T_α. If furthermore α is Diophantine and $k = \infty$ then by Herman-Yoccoz theorem Herman [1979], Yoccoz [1984] this conjugacy is smooth which amounts to $\mathfrak{F}_\alpha^\infty(\mathbb{T}) = \mathfrak{O}_\alpha^\infty(\mathbb{T})$. It is of course natural to try to extend this result to the higher dimensional situation where f is an orientation preserving diffeomorphism of the d-dimensional torus $\mathbb{T}^d$. Unfortunately, no Denjoy theorem is available in this situation and the only reasonable question to ask for is the following

Question 1. *Let $f : \mathbb{T}^d \to \mathbb{T}^d$ be a smooth diffeomorphism of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ which is topologically conjugate to a translation $T_\alpha : \mathbb{T}^d \to \mathbb{T}^d$, $x \mapsto x + \alpha$ with α Diophantine. Is the conjugacy smooth?*

Notice that when $d = 2$, even if α is Diophantine, $\mathcal{F}_\alpha^\infty(\mathbb{T}^2)$ is not equal to $\overline{\mathcal{O}_\alpha^\infty(\mathbb{T}^2)}$ or $\overline{\mathcal{O}_\alpha^\infty(\mathbb{T}^d)}$ as is shown by taking projectivization of cocyles in $SW^\infty(\mathbb{T}, SL(2,\mathbb{R}))$: such cocycles have a uniquely defined rotation number, that can be chosen Diophantine, and at the same time can have positive Lyapunov exponents (which prevents the projective action to be conjugated to a translation) (cf. Herman [1983]). Analogously, by taking projectivization of cocycles in $SW^\omega(\mathbb{T}, SL(2,\mathbb{R}))$ and using Avila's theory characterizing sub-critical/critical cocycles and the Almost Reducibility Conjecture (see Section 5) one can show that there exist elements of $\overline{\mathcal{O}_\alpha^\infty(\mathbb{T}^d)}$ which are not C^∞- almost reducible and, even if $\alpha \in \mathbb{T}^d$ is Diophantine, that the set $\mathcal{O}_\alpha^\infty(\mathbb{T}^d)$ is not closed.

In a similar vein

Question 2. *Let $f : \mathbb{T}^d \to \mathbb{T}^d$ be a smooth diffeomorphism which is topologically conjugate to the translation with α non-resonant. Is it C^∞-accumulated by elements of $\mathcal{O}_\alpha^\infty(\mathbb{T}^d)$? Is it C^∞-almost reducible?*

When $d = 1$ the first and the second part of the preceding question have a positive answer. Yoccoz proved Yoccoz [1995b] that $\mathcal{F}_\alpha^\infty(\mathbb{T}) = \overline{\mathcal{O}_\alpha^\infty(\mathbb{T})}$ and it is proved in Avila and Krikorian [n.d.(b)] that any smooth orientation preserving diffeomorphism of the circle is C^∞-almost reducible. The proof of this result uses renormalization techniques which at the present time doesn't seem to extend to the higher dimensional case. Still the situation in the semi-local case might be more accessible.

Question 3. *Same questions as in Questions 1 and 2 in the semi-local case that is for f in some neighborhood of the set of rotations, independent of α.*

If one assumes α to be Diophantine and f to be in a neighborhood of T_α that *depends* on α the answer to Question 1 is positive; this can be proved by standard KAM techniques.

3 Pseudo-rotations of the disc

A C^k ($k \in \mathbb{N} \cup \{\infty, \omega\}$) *pseudo rotation* of the disk $\mathbb{D} = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$ is a C^k orientation and *area preserving* diffeomorphism of the disk $\mathbb{D}$ that fixes the origin, leaves invariant the boundary $\partial\mathbb{D}$ of the disk and with *no other periodic point than the origin*. Like in the case of circle diffeomorphisms one can define for such pseudo-rotation a unique rotation number around the origin which is invariant by conjugation (see for example Franks [1988b, Corollary 2.6] or Franks [1988a, Theorem 3.3]). Anosov and Katok [1970] constructed in 1970, *via* approximation by periodic dynamics, *ergodic* (for the area measure) and infinitely differentiable pseudo-rotations of the disk, providing thus the first examples of pseudo-rotations which are not topologically conjugate to rigid rotations. By a theorem of Franks and Handel [2012] a *transitive* area and orientation preserving diffeomorphism of the disk fixing the origin and leaving invariant the boundary of the disk must be a pseudo-rotation.

3.1 Birkhoff rigidity conjecture. A famous question on pseudo-rotations attributed to Birkhoff is the following.

Question 4. *Is a real analytic pseudo-rotation of angle α analytically conjugated to the rotation R_α of angle α on the disc?*

Addressing this question should involve the artihmetics of α. On one hand, Rüssmann Rüssmann [1967] proved the following alternative for a Diophantine (in fact of Brjuno type is sufficient) elliptic fixed point of a real analytic area preserving surface diffeomorphism f: either the point is surrounded by a positive measure set of invariant circles with different Diophantine frequencies, or the map f is locally conjugate to a rotation in the neighborhood of the fixed point. On the other hand, when the real analytic category is relaxed to infinite differentiability, Anosov-Katok construction provides many counter-examples to the preceding question (for Liouville α's). We can thus divide the preceding question into two questions

Question 5. *Can one construct Anosov-Katok examples (viz. ergodic pseudo-rotations) in the real analytic category? If possible, can one impose the rotation number to be any non-Brjuno number?*

Question 6 (Reducibility)**.** *Is it true that every C^k, $k = \infty, \omega$, pseudo-rotation of the disk with diophantine rotation number α C^k-conjugated to a rigid rotation by angle α?*

Notice that in the smooth category the answer to Question 5 is positive: for the first part this is the existence of Anosov-Katok ergodic, even weak mixing, pseudo-rotations and for the second part one can prove that for any Liouville number α, there exists weak mixing pseudo-rotations as well as examples that are isomorphic to the rotation of frequency α on the circle Fayad and Saprykina [2005] and Fayad, Saprykina, and Windsor [2007]. Together with Herman's last geometric theorem, this gives in the C^∞-case a complete dichotomy between Diophantine and Liouville behavior.

Let $\mathcal{F}_\alpha^\infty$ be the set of C^∞ pseudo-rotations with rotation number α and $\mathcal{O}_\alpha^\infty$ be the set of $h \circ R_\alpha \circ h^{-1}$ where h is a C^∞ area and orientation preserving map of the disk fixing 0 and leaving invariant the boundary of the disk. A weaker question in the smooth case is :

Question 7. *For α diophantine is $\mathcal{O}_\alpha^\infty$ closed for the C^∞-topology?*

In fact, a more general question than Question 6 is the following:

Question 8 (Almost reducibility). *Is any C^k-pseudo-rotation $k = \infty, \omega$, f of the disk with irrational rotation number α almost reducible: there exists a sequence of area preserving smooth map h_n such that $h_n \circ f \circ h_n^{-1}$ converges in the C^k topology to R_α (in the analytic case this convergence should occur on a fixed complex neighborhood of the disk)?*

Question 6 has a positive answer in the local case (Rüssmann for $k = \omega$, Herman, Fayad and Krikorian [2009a] for $k = \infty$) that is when f is in some C^k-neighborhood of R_α (the size of this neighborhood *depending* on the *arithmetics* of α). Thus, a positive answer to Question 8 would imply a positive answer to Question 6. When $k = \infty$, Question 8 (hence Question 6) has a positive answer in the semi-local case Avila and Krikorian [n.d.(b)] that is with the extra assumption that for some k and ε *independent* of α, the C^k-norm of $Df - id$ is less than ε. In this situation one also has $\mathcal{F}_\alpha^\infty \cap W \subset \overline{\mathcal{O}_\alpha}^\infty$, where W is a neighborhood for the C^∞-topology of the set of rigid rotations. The proof of the result of Avila and Krikorian [ibid.] is based on renormalization techniques and on the fact (proved in Avila, Fayad, Calvez, Xu, and Zhang [2015]) that if one has a control on the C^1-norm of a pseudo-rotation f, the displacement $\max_{\mathbb{D}} \|f - id\|$ polynomially compares with the rotation number of f. Such a control is in general not true for diffeomorphisms of the circle. It is thus natural to ask:

Question 9. *Describe the set of smooth diffeomorphisms of the circle that are obtained as the restriction on $\mathbb{D}$ of the dynamics of pseudo-rotations?*

3.2 Rigidity times, mixing and entropy.

A diffeomorphism of class C^k, $k \in \mathbb{N} \cup \{\infty\}$, is said to admit C^k *rigidity times* (or for short is C^k-rigid) if there exists a sequence q_n such that f^{q_n} converges to the Identity map in the C^k topology. If we just know that the latter holds in a fixed neighborhood of some point p, we say that f is C^k *locally rigid* at p. All the smooth examples on the disc or the sphere obtained by the Anosov-Katok method are C^∞*-rigid* by construction. Obviously, rigidity or local rigidity precludes mixing. Hence, the following natural question was raised in Fayad and Katok [2004] in connection with the smooth realization problem and the Anosov-Katok construction method.

Question 10. *Is it true that a smooth area preserving diffeomorphism of the disc with zero metric entropy is not mixing?*

In the case of zero topological entropy, and in light of Franks and Handel result, the question becomes

Question 11. *Is it true that a smooth pseudo-rotation is not mixing?*

Bramham [2015] proved that this is true if the rotation number is sufficiently Liouville; indeed he proves in that case the existence of C^0-rigidity times. It was shown in Avila, Fayad, Calvez, Xu, and Zhang [2015] that real analytic pseudo-rotations (with no restriction on the rotation number) are never topologically mixing. By a combination of KAM results and control of recurrence for pseudo-rotations with Liouville rotation numbers, it is actually shown that real analytic pseudo-rotations are C^∞ locally rigid near their center.

Note that the following is not known, except in C^1 regularity where a positive answer is given by Bochi [2002].

Question 12. *Does there exist a smooth area preserving disc diffeomorphism that has zero metric entropy and positive topological entropy?*

The following question was raised by Bramham in Bramham [2015].

Question 13. *Does every C^k pseudo-rotation f admit C^0 rigidity times? The question can be asked for any $k \geq 1$, $k = \infty$ or $k = \omega$.*

In the case $k = \omega$ or $\rho(f)$ Diophantine and $k = \infty$, the latter question becomes an intermediate question relative to the Birkhoff-Herman problem on the conjugability of f to the rigid disc rotation of angle $\rho(f)$. In Avila, Fayad, Calvez, Xu, and Zhang [2015] it was shown that for every irrational α, if an analytic pseudo-rotation of angle α is sufficiently close to R_α then it admits C^∞-rigidity times.

Question 14. *Given a fixed analyticity strip, does there exist $\epsilon > 0$ such that if a real analytic pseudo-rotation is ϵ close to the rotation on the given analyticity strip, then it is rigid?*

An *a priori* control on the growth of $\|Df^m\|$ for a pseudo-rotation is sufficient to deduce the existence of rigidity times for larger classes of rotation numbers. If for example a polynomial bound holds on the growth of $\|Df^m\|$ for a smooth pseudo-rotation, then the existence of C^∞ rigidity times would follow for any Liouville rotation number (see Avila, Fayad, Calvez, Xu, and Zhang [ibid.]). In the case of a circle diffeomorphism f a gap in the growth of these norms is known to hold between exponential growth in the case f has a hyperbolic periodic point or a growth bounded by $O(m^2)$ if not Polterovich and Sodin [2004]. Does a similar dichotomy hold for area preserving disc diffeomorphisms?

Question 15. *Is there any polynomial bound on the growth of the derivatives of a pseudo-rotation? Does every C^∞ pseudo-rotation with Liouville rotation number admit C^0 (or even C^∞) rigidity times?*

With Herman's last geometric theorem, a positive answer to the second part of Question 15 would imply that smooth pseudo-rotations, and therefore area preserving smooth diffeomorphisms of the disc with zero topological entropy are never topologically mixing.

In the proof of absence of mixing of an analytic pseudo-rotation, Avila, Fayad, Calvez, Xu, and Zhang [2015] uses an *a priori* bound on the growth of the derivatives of the iterates of a pseudo-rotation that is obtained *via* an effective finite information version of the Katok closing lemma for an area preserving surface diffeomorphism f. This effective result provides a positive gap in the possible growth of the derivatives of f between exponential and sub-exponential.

In Fayad and Zhang [2017], an explicit finite information condition is obtained for area preserving C^2 surface diffeomorphisms, that guarantees positive topological entropy.

Question 16. *Find a finite information condition on the complexity growth of an area preserving C^2 surface diffeomorphism that insures positive metric entropy.*

Finally, inspired by Rüssmann and Herman's last geometric theorem on one hand, and the Liouville pseudo-rotations rigidity on the other, we ask the following

Question 17. *Can a smooth area preserving diffeomorphism of a surface that has an irrational elliptic fixed point be topologically mixing? Can it have an orbit that converges to the fixed point?*

4 Hamiltonian systems

A C^2 function $H : (\mathbb{R}^{2d}, 0) \to \mathbb{R}$ such that $DH(0) = 0$ defines on a neighborhood of 0 a hamiltonian vector field $X_H(x, y) = (\partial_y H(x, y), -\partial_x H(x, y))$ and its flow ϕ_H^t is a flow of symplectic diffeomorphisms preserving the origin. We shall assume that $0 \in \mathbb{R}^{2n}$ is an elliptic equilibrium point with H of the following form

$$H(x, y) = \sum_{j=1}^{d} \omega_j (x_j^2 + y_j^2)/2 + O_3(x, y), \tag{4-2}$$

where the frequency vector ω is *non-resonant*.

Alternatively we may take H a C^2 function defined on $\mathbb{T}^d \times \mathbb{R}^d$ and consider its Hamiltonian flow $X_H(\theta, r) = (\partial_r H(\theta, r), -\partial_\theta H(\theta, r))$. If

$$H(\theta, r) = \langle \omega_0, r\rangle + \mathcal{O}(r^2) \tag{4-3}$$

then the torus $\mathbb{T}^d \times \{0\}$ is invariant under the Hamiltonian flow and the induced dynamics on this torus is the translation $\phi_H^t : \theta \mapsto \theta + t\omega_0$. Moreover this torus is Lagrangian with respect to the canonical symplectic form $d\theta \wedge dr$ on $\mathbb{T}^d \times \mathbb{R}^d$. When ω is Diophantine we say that this torus is a KAM torus.

The stability of an equilibrium or of an invariant quasi-periodic torus by a Hamiltonian flow can be studied from three points of view. The usual topological or Lyapunov stability, the stability in a measure theoretic or probabilistic sense which can be addressed by KAM theory (Kolmogorov, Arnold, Moser), or the effective stability in which one is interested in quantitative stability in time.

4.1 Topological stability. Arnold conjectured that apart from two cases, the case of a sign-definite quadratic part, and generically for $d = 2$, an elliptic equilibrium point is generically unstable.

Conjecture 4.1 (Arnold)**.** *An elliptic equilibrium point of a generic analytic Hamiltonian system is Lyapounov unstable, provided $n \geq 3$ and the quadratic part of the Hamiltonian function at the equilibrium point is not sign-definite.*

Despite a rich literature and a wealth of results in the C^∞ smoothness (to give a list of contributions would exceed the scope of this presentation), this conjecture is wide open in the real analytic category, to such an extent that under our standing assumptions (real-analyticity of the Hamiltonian and a non-resonance condition on the frequency vector) not a single example of instability is known.

Question 18. *Give examples of an analytic Hamiltonian that have a non-resonant elliptic equilibrium (or a non-resonant Lagrangian quasi-periodic torus) that is Lyapunov unstable.*

Question 19. *Give examples of an analytic Hamiltonian that have a non-resonant elliptic equilibrium (or a non-resonant Lagrangian quasi-periodic torus) that attracts an orbit (distinct from the equilibrium or the torus itself).*

In Fayad, Marco, and Sauzin [n.d.] an example is given of a Gevrey regular Hamiltonien on $\mathbb{R}^6$ that has a non-resonant fixed point at the origin and that has an orbit distinct from the origin that converges to it in the future. In Kaloshin and Saprykina [2012] and Guardia and Kaloshin [2014], Arnold diffusion methods are used to yield in particular orbits that have α-limit or ω-limit sets that are non-resonant invariant Lagrangian tori instead of a single non-resonant fixed point.

Following Perez-Marco we ask:

Question 20. *Is it true that a smooth Hamiltonian flow with a non-resonant elliptic equilibrium isolated from periodic points has a hedgehog (a totally invariant compact connected set containing the origin)?*

Regarding the additional stability features of elliptic fixed points in the case of two degrees of freedom, we ask the following

Question 21. *Is the iso-energetic twist condition the optimal condition for Lyapunov stability of an irrational elliptic equilibrium in two degrees of freedom?*

A smooth example of an irrational equilibrium was constructed by F. Trujillo that satisfies the Kolmogorov non degeneracy condition in $d = 2$ degrees of freedom and that has diffusing orbits in some special energy levels.

4.2 Beyond the classical KAM theory. An equilibrium (or an invariant torus) of a Hamiltonian system is said to be KAM stable if it is accumulated by a positive measure of invariant KAM tori, and if the set of these tori has density one in the neighborhood of the equilibrium (or the invariant torus).

4.2.1 Weak transversality conditions. In classical KAM theory, an elliptic fixed point is shown to be KAM-stable under the hypothesis that the frequency vector at the fixed point is non-resonant (or just sufficiently non-resonant) and that the Hamiltonian is sufficiently smooth and satisfies a generic non degeneracy condition of its Hessian matrix at the fixed point. Further development of the theory allowed to relax the non degeneracy condition. In Eliasson, Fayad, and Krikorian [2013] KAM-stability was established for non-resonant elliptic fixed points under the (most general) Rüssmann transversality condition on the Birkhoff normal form of the Hamiltonian. Similar results were obtained for Diophantine invariant tori in Eliasson, Fayad, and Krikorian [2015].

4.2.2 Absence of transversality conditions.

Conjecture 4.2. *[Herman] Prove that an elliptic equilibrium with a diophantine frequency or a KAM torus of an analytic Hamiltonian is accumulated by a set of positive measure of KAM tori.*

Clearly, one can of course ask whether KAM stability also holds.

Conjecture 4.2 was was made by M. Herman in his ICM98 lecture (in the context of symplectomorphisms). The conjecture is known to be true in two degrees of freedom Rüssmann [1967], but remains open in general. It is shown in Eliasson, Fayad, and Krikorian [2015] that an analytic invariant torus $\mathcal{T}_0$ with Diophantine frequency ω_0 is never isolated due to the following alternative. If the Birkhoff normal form of the Hamiltonian at $\mathcal{T}_0$ satisfies a Rüssmann transversality condition, the torus $\mathcal{T}_0$ is accumulated by KAM tori of positive total measure. If the Birkhoff normal form is degenerate, there exists a subvariety of dimension at least $d+1$ that is foliated by analytic invariant tori with frequency ω_0.

For Liouville frequencies, one does not expect the conjecture to hold.

Question 22. *Give an example of an analytic Hamiltonian that has a non-resonant (Liouville) elliptic equilibrium that is not is accumulated by a set of positive measure of KAM tori.*

In the C^∞ category (or Gevrey), counter-examples to stability with positive probability can be obtained: in 2 or more degrees of freedom for Liouville frequencies; and in 3 or more degrees of freedom for any frequency vector (Eliasson, Fayad, and Krikorian [ibid.] for $d \geq 4$ and Fayad and Saprykina [2005] for $d \geq 3$). In the remaining case of

Diophantine equilibrium with $d = 2$, Herman proved stability with positive probability without any twist condition (see Fayad and Krikorian [2009a]).

4.3 Effective stability. Combining KAM theory, Nekhoroshev theory and estimates of Normal Birkhoff forms, it was proven in Bounemoura, Fayad, and Niederman [2017] that generically, both in a topological and measure-theoretical sense, an invariant Lagrangian Diophantine torus of a Hamiltonian system is doubly exponentially stable in the sense that nearby solutions remain close to the torus for an interval of time which is doubly exponentially large with respect to the inverse of the distance to the torus. It is proven there also that for an arbitrary small perturbation of a generic integrable Hamiltonian system, there is a set of almost full positive Lebesgue measure of KAM tori which are doubly exponentially stable. These results hold true for real-analytic but more generally for Gevrey smooth systems. Similar results for elliptic equilibria are obtained in Bounemoura, Fayad, and Niederman [2015].

Question 23. *Give examples of analytic or Geverey differentiable Hamiltonians that have a Diophantine elliptic equilibrium with positive definite twist, that is not more than doubly-exponentially stable in time. Show that this is generic.*

Question 24. *Give an example of an analytic Hamiltonian that has a non-resonant elliptic equilibrium with positive definite twist that is not more than exponentially stable in time.*

Question 25. *Give an example of an analytic Hamiltonian that has a Diophantine elliptic equilibrium that is not more than exponentially stable in time.*

4.4 On invariant tori of convex Hamiltonians.

4.4.1 The "last invariant curve" of annulus twist maps. A classic topic in Hamiltonian systems is that of the regularity of the invariant curves of annulus twist maps. A celebrated result of Birkhoff states that such curves (if they are not homotopic to a point) must be Lipschitz. Numerical evidence seems to indicate that invariant curves are always at least C^1. After Mather and Arnaud we ask the following.

Question 26. *Give an example of a C^r, $r \in [2, \infty) \cup \{\omega\}$, annulus twist map that has an invariant C^0 but not C^1 curve with minimal restricted dynamics.*

In Avila and Fayad [n.d.], a C^1 example is constructed, and Arnaud [2011] gives a C^1 example with an invariant C^0 but not C^1 curve having Denjoy type restricted dynamics.

Due to a result proved by Herman the problem can be reduced to finding a minimal circle homeomorphism f such that $f + f^{-1}$ is C^r but f is only C^0.

Question 27. *Give an example of a C^r, $r \in [2,\infty) \cup \{\omega\}$, annulus twist map that has an invariant C^r curve that is not accumulated by other invariant curves.*

4.4.2 On the destruction of all tori.

Given the Hamiltonian $H = \frac{1}{2}\sum r_i^2$ on $\mathbb{T}^d \times \mathbb{R}^d$.

Question 28. *What is the maximum of r for which it is possible to perturb H so that the perturbed flow has no invariant Lagrangian torus that is the graph of a C^1 function.*

By Herman, $r \geq d+2-\epsilon$, $\forall \epsilon > 0$. We also know that $r \leq 2d$ (see Pöschel [1982]). In Cheng and L. Wang [2013], given any frequency ω, a $C^{2d-\epsilon}$ perturbation of H is given that has no invariant Lagrangian torus with as unique rotation frequency vector ω.

4.5 Birkhoff Normal Forms.

Let $H : (\mathbb{R}^{2d}, 0) \to \mathbb{R}$ be a real analytic hamiltonian function admitting 0 as an elliptic non-resonant fixed point. One can always *formally* conjugate H to an *integrable* hamiltonian: there exist a *formal* (exact) symplectic germ of diffeomorphism g tangent to the identity and a *formal* series $N \in \mathbb{R}[[r_1,\dots,r_d]]$ such that $g_* X_H = X_B$ where $B(x,y) = N(x_1^2+y_1^2,\dots,x_d^2+y_d^2)$. This B is unique and is called the *Birkhoff Normal Form* (BNF). This formal object is an invariant of C^k-conjugations ($k=\infty,\omega$). Birkhoff Normal Forms can be defined for C^k ($k=\infty,\omega$) symplectic diffeomorphisms admitting an invariant elliptic fixed point or even (in the case of symplectic diffeomorphisms or hamiltonian flows) in a neighborhood of an invariant KAM torus (the frequency must be then diophantine). Siegel [1954] proved that in general the conjugating transformation could not be convergent and Eliasson asked whether the Birkhoff Normal Form itself could be convergent. In the real analytic setting Pérez-Marco [2003] proved that for any given non-resonant quadratic part one has the following dichotomy: either the BNF always converges or it generically diverges. Gong [2012] provided an example of divergent BNF with Liouville frequencies. In Krikorian [n.d.] it is proved that the BNF of a real analytic symplectic diffeomorphism admitting a diophantine elliptic fixed point (with torsion) is generally divergent.

Question 29. *Let H be a real analytic Hamiltonian admitting the origin as a diophantine elliptic fixed point and assume that its Birkhoff Normal Form defines a real analytic function. Is H real analytically conjugated to its Birkhoff Normal Form on a neighborhood of the origin?*

5 Dynamics of quasi-periodic cocycles

Let G be a Lie group (possibly infinite dimensional). A quasi-periodic cocycle of class C^k, $k \in \mathbb{N} \cup \{\infty, \omega\}$ is a map $(\alpha, A) : \mathbb{T}^d \times G \to \mathbb{T}^d \times G$ of the form $(\alpha, A) : (x, y) \mapsto (x+$

$\alpha, A(x)y)$ where $\alpha \in \mathbb{T}^d$ (we assume α to be non-resonant) and $A : \mathbb{T}^d \to G$ is of class C^k. We denote the set of such cocycles (α, A) by $SW^k(\mathbb{T}^d, G)$ (or $SW^k_\alpha(\mathbb{T}^d, G)$). The iterates $(\alpha, A)^n$ of (α, A) are of the form $(n\alpha, A^{(n)})$ where (for $n \geq 1$) $A^{(n)}$ is the fibered product $A^{(n)}(\cdot) = A(\cdot + (n-1)\alpha) \cdots A(\cdot + \alpha)A(\cdot)$. Two cocycles (α, A_1) and (α, A_2) are said to be C^l-*conjugated* if there exists a map $B : \mathbb{T}^d \to G$ (or $B : \mathbb{R}^d/N\mathbb{Z}^d \to G$ for some $N \in \mathbb{N}^*$) of class C^l such that $(\alpha, A_2) = (0, B) \circ (\alpha, A_1) \circ (0, B)^{-1}$ or equivalently $A_2 = B(\cdot + \alpha)A_1 B(\cdot)^{-1}$. The cocycle (α, A) is said to be *reducible* if it is conjugated to a constant cocycle and, when H is a subgroup of G, H-*reducible* if it is conjugated to an H-valued (not necessarily constant) cocycle. We say that the cocycle is *linear* when the group G is a group of matrices.

5.1 The case $G = SL(2, \mathbb{R})$. Quasi-periodic $SL(2, \mathbb{R})$-valued cocycles play an important role in the theory of *quasi-periodic Schrödinger operators on* $\mathbb{Z}$ of the form $H_x : l^2(\mathbb{Z}) \to l^2(\mathbb{Z})$, $H_x : (u_n)_{n\in\mathbb{Z}} \mapsto (u_{n+1} + u_{n-1} + V(x + n\alpha)u_n)_{n\in\mathbb{Z}}$; indeed, the (generalized) eigenvalue equation $H_x u = Eu$ leads naturally to studying the dynamics of a family of $SL(2, \mathbb{R})$-valued quasi-periodic cocycles depending on E, the so-called *Schrödinger cocycles*. Many spectral objects or quantities – such as, resolvent sets (complement of the spectrum), spectral measures, density of states, speed of decay of Green functions... – of the family of operators H_x, $x \in \mathbb{T}^d$, can be related to dynamical notions or invariants for the associated family of Schrödinger cocycles – namely (in that order), uniform hyperbolicity, m-functions, fibered rotation number, Lyapunov exponents... We refer to Eliasson [1998], You [2018] for more details on this topic.

There are two important quantities associated to $SL(2\,\mathbb{R})$-valued quasi-periodic cocycles which are invariant by conjugation[1]: the Lyapunov exponent $L(\alpha, A)$ which measure the exponential speed of growth of the iterates of the cocycle (α, A) and the fibered rotation number $\rho(\alpha, A)$ which measures the average speed of rotation of non-zero vectors in the plane under iteration of the cocycle. It is of course tempting to try and classify $SL(2, \mathbb{R})$-cocycles according to these two invariants.

The case of real analytic cocycles with one frequency is particularly well understood. In that situation, following A. Avila [2015], one can associate to any cocycle $(\alpha, A) \in SW^\omega(\mathbb{T}, SL(2, \mathbb{R}))$ a natural family $(\alpha, A_\varepsilon) \in SW^\omega(\mathbb{T}, SL(2, \mathbb{C}))$ (ε in some neighborhood of 0) with $A_\varepsilon(\cdot) = A(\cdot + \varepsilon\sqrt{-1})$. The function $\varepsilon \mapsto L(\alpha, A_\varepsilon)$ plays a very important role in the theory; Avila proved that it is an even convex continuous piecewise affine map with *quantized* slopes in $2\pi\mathbb{Z}$ (this is the phenomenon of "quantization of acceleration") and that the complex cocycle (α, A_ε) is *uniformly hyperbolic* if and only ε is not a break point of $\varepsilon \mapsto L(\alpha, A_\varepsilon)$. This analysis leads to the notions of *critical*, *supercritical* and *subcritical* cocycles, where this last term refers to the fact that the function $\varepsilon \mapsto L(\alpha, A_\varepsilon)$

[1] for the rotation number one has to assume the conjugating map to be homotopic to the identity

is zero on an neighborhood of $\varepsilon = 0$. A cocycle $(\alpha, A) \in SW^{\omega}(\mathbb{T}, SL(2,\mathbb{R}))$ (homotopic to the identity) can thus have four distinct possible behaviors if one adds to the three preceding ones uniform hyperbolicity. Moreover, the quantization of acceleration allows to *predict* the possible transitions between these four regimes and to draw consequences on the spectrum of Schrödinger operators such as for example the possibility of co-existence of absolutely continuous or pure point spectrum for some type of potentials (cf. Avila [ibid.] and for other examples Bjerklöv and Krikorian [n.d.]). The most striking *global* result on the dynamics of these cocycles is certainly the "Almost reducibility conjecture" proved by Avila Avila [2010], Avila [n.d.] which asserts that any *subcritical* cocycle in $SW^{\omega}(\mathbb{T}, SL(2,\mathbb{R}))$ is *almost-reducible* (in the analytic category, on a fixed complex neighborhood of the real axis). By Hou and You [2012], You and Zhou [2013] in the real analytic semi-local situation (viz. when A is close to a constant, this closeness being independent of α) a cocycle (α, A) is either uniformly hyperbolic or subcritical.

In the C^{∞} category, or for many-frequencies systems, our understanding of the dynamics of cocycles is much less complete. There are important reducibility or almost-reducibility results (Dinaburg and J. G. Sinaĭ [1975], Eliasson [1992], Krikorian [1999a], Krikorian [1999b], Krikorian [2001], Avila and Krikorian [2006],Puig [2004], Puig [2006], Fayad and Krikorian [2009b], Avila, Fayad, and Krikorian [2011], Hou and You [2012], You and Zhou [2013], Avila and Krikorian [2015]...) but they often involve diophantine conditions and/or are of perturbative nature. Moreover, the semi-local version of the Almost reducibility conjecture has no reasonable equivalent in the smooth (or even Gevrey) setting Avila and Krikorian [n.d.(c)]. Still, one can ask:

Question 30. *Is the semi-local version of the Almost reducibility conjecture true for cocyles in quasi-analytic classes?*

Let's say that a cocycle is *stable* if it is not accumulated by non-uniformly hyperbolic systems (with the same frequency vector on the base). Having in mind Avila's classification one can ask:

Question 31. *Is every stable cocycle in $SW^{k}(\mathbb{T}^{d}, SL(2,\mathbb{R}))$, $k = \infty, \omega$, almost-reducible?*

5.2 The symplectic case. Cocycles in $SW^{k}(\mathbb{T}^{d}, Sp(2n,\mathbb{R}))$ are of interst when one tries to understand the dynamics of a symplectic diffeomorphism in the neighborhood of an invariant torus (they appear as linearized dynamics) or in the study of quasi-periodic Schrödinger operators on strips $\mathbb{Z} \times \{1,\ldots,n\}$. For such cocycles one can define $2n$ Lyapunov exponents (symmetric with respect to 0) and one fibered Maslov index which plays the role of a fibered rotation number (cf. Xu [2016] and the references there).

We denote by $SO(2,\mathbb{R})$ the set of symplectic rotations $R_t = \begin{pmatrix} (\cos t) I_n & -(\sin t) I_n \\ (\sin t) I_n & (\cos t) I_n \end{pmatrix}$.

Question 32. *Let $(\alpha, A) \in SW^\infty(\mathbb{T}, Sp(2n,\mathbb{R}))$ homotopic (resp. non homotopic) to the identity where $\alpha \in \mathbb{T}$ is (recurrent) diophantine. Is it true that for Lebesgue almost all $t \in \mathbb{R}$ the following dichotomy holds: either the cocycle $(\alpha, R_t A)$ is C^∞-reducible (resp. $SO(2,\mathbb{R})$-reducible) or its upper Lyapunov exponent is positive?*

When $n = 1$ the answer is positive (Avila and Krikorian [2006] for the case homotopic to the identity, Avila and Krikorian [2015] for the case non-homotopic to the identity). The proof of this result is based on a renormalization procedure which works when the cocycle has some mild boundedness property and on a reduction to this case based on Kotani theory. In the case $n \geq 2$ such a Kotani theory was developed by Xu in Xu [2016], Xu [2015]. Following the same strategy as in Avila and Krikorian [2006] one should be then reduced to studying cocycles with values in the maximal compact subgroup of $Sp(2n,\mathbb{R})$. Unfortunately, one cannot conclude like in the case $n = 1$ since no reasonable *a priori* notion of fibered rotation number can be defined for cocycles with values in non-abelian compact groups (they can be defined *a posteriori* once one knows the cocycle is reducible; see Karaliolios [2017], Karaliolios [2016] for related results).

5.3 The case $G = \mathrm{Diff}_0^\infty(\mathbb{T})$.

A cocycle $(\alpha, A) \in SW(\mathbb{T}^d, SL(2,\mathbb{R}))$ naturally produces a *projective cocycle* $(\alpha, \bar{A}) \in SW(\mathbb{T}^d, \mathrm{Hom}(\mathbb{S}^1))$ where $\mathrm{Hom}(\mathbb{S}^1)$ is the group of homographies acting on $\mathbb{S}^1$; namely $\bar{A}(x) \cdot v = (A(x)v)/\|A(x)v\|$. It is thus natural to look at the more general case where the underlying group is the group of orientation preserving diffeomorphisms of the circle. In that case one can still define a fibered rotation number Herman [1983]. For the topological aspects of the theory of such quasi-periodically forced circle diffeomorphisms see Bjerklöv and Jäger [2009].

Question 33 (Non-linear Eliasson Theorem)**.** *Let $\alpha \in \mathbb{T}^d$ be a fixed diophantine vector and $G = \mathrm{Diff}_0^\infty(\mathbb{R}/\mathbb{Z})$. Does there exist k_0, ε_0 depending only on α such that for any $(\alpha, A) \in SW^\infty(\mathbb{T}^d, \mathrm{Diff}_0^\infty(\mathbb{R}/\mathbb{Z}))$ of the form $(\alpha, A)(x, y) = (x + \alpha, y + \beta + f(x, y))$ with $\|f\|_{C^{k_0}} \leq \varepsilon_0$ and $\rho(\alpha, A)$ diophantine, the cocycle (α, A) is C^∞-reducible?*

When $G = \mathrm{Hom}(\mathbb{S}^1)$ the answer is positive and is (the C^∞-version of) a theorem of Eliasson [1992] which has many consequences in the theory of quasi-periodic Schrödinger operators. If one allows ε_0 to depend on ρ then the result is true and is essentially a (generalization of a) theorem by Arnold. Its proof is classical KAM theory. In Krikorian, J. Wang, You, and Zhou [n.d.] a result of rotations-reducibility is proved where ε_0 depends on ρ but with considerably weaker assumption on α than KAM theory usually allows (compare with Avila, Fayad, and Krikorian [2011], Fayad and Krikorian [2009b] for stronger results in the case of linear cocycles).

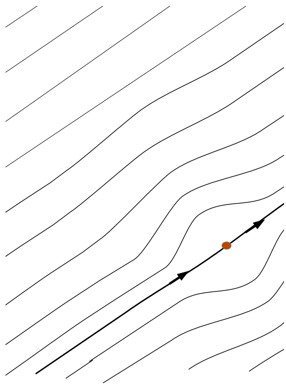


Fig 1. Degenerate saddle acting as a stopping point

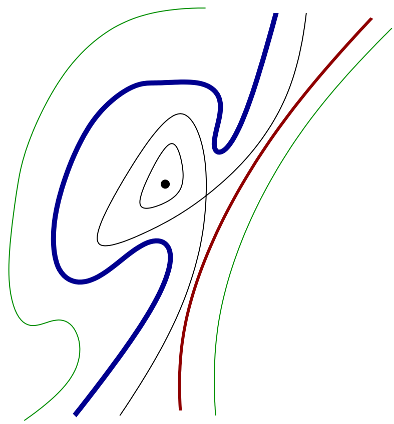


Fig 2. Non-degenerate saddle that causes asymmetry

6 Mixing surface flows

6.1 Spectral type. Area preserving surface flows provide the lowest dimensional setting in which it is interesting to study conservative systems. Such flows are sometimes called multi-valued Hamiltonian flows to emphasize their relation with solid state physics that was pointed out by Novikov [1982]. Via Poincaré sections, these flows are related to special flows above circle rotations or more generally above IETs (Interval exchange transformations). One can thus view them as time changes of translation flows on surfaces.

Katok and then Kochergin showed the absence of mixing of area preserving flows on the two torus if they do not have singularities Katok [1975] and Kočergin [1972].

The simplest mixing examples are those with one (degenerate) singularity on the two torus produced by Kochergin in the 1970s Kočergin [1975]. Kochergin flows are time changes of linear flows on the two torus with an irrational slope and with a rest point (see Figure 1).

Multi-valued Hamiltonian flows on higher genus surfaces can also be mixing (or mixing on an open ergodic component) in the presence of non-degenerate saddle type singularities that have some asymmetry (see Figure 2). Such flows are called Arnol'd flows and their mixing property, conjectured by Arnol'd in Arnold [1991], was obtained by Y. G. Sinaĭ and Khanin [1992] and in more generality by Kochergin [2003, 2004]. Note that Ulcigrai proved in Ulcigrai [2011] that area preserving flows with non-degenerate saddle singularities are generically not mixing (due to symmetry in the saddles).

Question 34. *Study the spectral type and spectral multiplicity of mixing flows on surfaces.*

By spectral type of a flow $\{T^t\}$ we mean the spectral type of the associated Koopman operator $U_t : L^2(M, \mu) : f \to f \circ T^t$.

It was proved in Fayad, Forni, and Kanigowski [2016] that Kochergin flows with a sufficiently strong power like singularity have for almost every slope a maximal spectral type that is equivalent to Lebesgue measure. The study of the spectral multiplicity of these flows is interesting in its relation to the Banach problem on the existence of a dynamical system with simple Lebesgue spectrum. It is probable however that the spectral multiplicity of Kochergin flows is infinite. Mixing reparametrizations of linear flows with simple spectrum were obtained in Fayad [2005] and it would be interesting to study their maximal spectral type following Fayad, Forni, and Kanigowski [2016].

Question 35. *Is it true that Arnol'd mixing flows have in general a purely singular spectral type?*

Arnol'd conjectured a power-like decay of correlation in the non-degenerate asymmetric case, but the decay is more likely to be logarithmic, at least between general regular observables or characteristic functions of regular sets such as balls or squares. Even a lower bound on the decay of correlations is not sufficient to preclude absolute continuity of the maximal spectral type. However, an approach based on slowly coalescent periodic approximations as in Fayad [2006] may be explored in the aim of proving that the spectrum is purely singular.

6.2 Spectral type of related systems.

Question 36. *Prove that all IET have a purely singular maximal spectral type.*

It is known that almost every IET, namely those that are not of constant type, are rigid. It follows that their maximal spectral type is purely singular. For the remaining IETs, partial rigidity was proven by Katok and used to show the absence of mixing, but proving that the spectral type is purely singular appears to be more delicate.

Question 37. *Prove that on $\mathbb{T}^3$ there exists a real analytic strictly positive reparametrization of a minimal translation flow that has a Lebesgue maximal spectral type.*

The difference with the Kochergin flows is that such flows would also be uniquely ergodic. Mixing real analytic reparametrizations of linear flows on $\mathbb{T}^3$ were obtained in Fayad [2002].

6.3 Multiple mixing. The question of multiple mixing for mixing systems is one of the oldest unsolved questions of ergodic theory.

Question 38. *Are all mixing surface flows mixing of all orders?*

Arnold and Kochergin mixing conservative flows on surfaces stand as the main and almost only natural class of mixing transformations for which higher order mixing has not been established nor disproved in full generality. Under suitable arithmetic conditions on their unique rotation vector, of full Lebesgue measure in the first case and of full Hausdorff dimension in the second, it was shown in Fayad and Kanigowski [2016] that these flows are mixing of any order, Kanigowski, Kuaaga-Przymus, and Ulcigrai [n.d.] for flows on higher genus surfaces).

7 Ergodic theory of diagonal actions on the space of lattices and applications to metric Diophantine approximation

The Diophantine properties of linear forms of one or several variables evaluated at integer points are intimately related to the divergence rates of some orbits under some diagonal actions in the space of (linear or affine) lattices of $\mathbb{R}^n$. This link is due to what can be called the Dani correspondence principle between the small values of the linear forms on one hand and the visits to the cusp of certain orbits of certain diagonal actions on the space of lattices (affine lattices in the case of inhomogeneous linear forms). The ergodic study of diagonal and unipotent actions on the space of lattices provides indeed an efficient substitute to the continued fraction algorithm that played a crucial role in the rich metric theory of Diophantine approximations in dimension 1. There is a number of important contributions to number theory related to this principle and to progress in the theory of homogeneous actions for example the surveys Dani [1994], Hasselblatt and Katok [2002], Einsiedler and Lindenstrauss [2006], Eskin [2010], and Marklof [2006, 2007]). We mention here a list of questions related to the statistical properties of Kronecker sequences that can be approached using this same principle. More details and questions can be found in Dolgopyat and Fayad [2015].

7.1 Kronecker sequences. A quantitative measure of uniform distribution of Kronecker sequences is given by the discrepancy function: for a set $\mathcal{C} \subset \mathbb{T}^d$ let

$$D(\alpha, x, \mathcal{C}, N) = \sum_{n=0}^{N-1} \mathbf{1}_{\mathcal{C}}(x + k\alpha) - N \text{volume}(\mathcal{C})$$

where $(\alpha, x) \in \mathbb{T}^d \times \mathbb{T}^d$ and $\mathbf{1}_{\mathcal{C}}$ is the characteristic function of the set $\mathcal{C}$.

Uniform distribution of the sequence $x + k\alpha$ on $\mathbb{T}^d$ is equivalent to the fact that, for regular sets $\mathcal{C}$, $D(\alpha, x, \mathcal{C}, N)/N \to 0$ as $N \to \infty$. A step further is the study of the rate of convergence to 0 of $D(\alpha, x, \mathcal{C}, N)/N$.

Already for $d = 1$, it is clear that if $\alpha \in \mathbb{T} - \mathbb{Q}$ is fixed, the discrepancy $D(\alpha, x, \mathcal{C}, N)$ displays an oscillatory behavior according to the position of N with respect to the denominators of the best rational approximations of α. A great deal of work in Diophantine approximation has been done on giving upper and lower bounds to the oscillations of the discrepancy function (as a function of N) in relation with the arithmetic properties of $\alpha \in \mathbb{T}^d$.

In particular, let

$$\overline{D}(\alpha, N) = \sup_{\Omega \in \mathbb{B}} D(\alpha, 0, \Omega, N)$$

where the supremum is taken over all sets Ω in some natural class of sets $\mathbb{B}$, for example balls or boxes.

The case of (straight) boxes was extensively studied, and properties of the sequence $\overline{D}(\alpha, N)$ were obtained with a special emphasis on their relations with the Diophantine approximation properties of α. In particular, Beck [1994] proves that when $\mathbb{B}$ is the set of straight boxes in $\mathbb{T}^d$ then for arbitrary positive increasing function $\phi(n)$

$$\sum_n \frac{1}{\phi(n)} < \infty \iff \frac{\overline{D}(\alpha, N)}{(\ln N)^d \phi(\ln \ln N)} \text{ is bounded for almost every } \alpha \in \mathbb{T}^d. \tag{7-4}$$

In dimension $d = 1$, this result is the content of Khinchine theorems obtained in the early 1920's, and it follows easily from well-known results from the metrical theory of continued fractions (see for example the introduction of Beck [ibid.]). The higher dimensional case is significantly more difficult and many questions that are relatively easy to settle in dimension 1 remain open. We mention some here and refer to Beck [1994] and Kuipers and Niederreiter [1974] for others.

Question 39. *Is it true that* $\limsup \frac{\overline{D}(\alpha,N)}{\ln^d N} > 0$ **for all** $\alpha \in \mathbb{T}^d$*?*

Question 40. *Is it true that there exists* α *such that* $\limsup \frac{\overline{D}(\alpha,N)}{\ln^d N} < +\infty$*?*

The above questions and results can be asked for balls and more general convex sets.

Question 41. *Is it true that for any* $\epsilon > 0$*, for almost every* $\alpha \in \mathbb{T}^d$ *and for any convex set* $\mathcal{C}$ *in* $\mathbb{T}^d$

$$\frac{D(\alpha, 0, \mathcal{C}, N)}{N^{\frac{d-1}{2d}+\epsilon}}$$

is bounded?

The bound in (7-4) focuses on how bad can the discrepancy become along a subsequence of N, for a fixed α in a full measure set. In a sense, it deals with the worst case scenario and do not capture the oscillations of the discrepancy.

Another point of view is to let $(\alpha, x) \in \mathbb{T}^d \times \mathbb{T}^d$ be random and have limit laws that hold for *all* N. By random we mean distributed according to a smooth density on the tori. For $d = 1$, this was done by Kesten who proved in the 1960s that the discrepancies of the number of visits of the Kronecker sequence to an interval, normalized by $\rho \ln N$ (where ρ depends on the interval but is constant if the length of the interval is irrational) converges to a Cauchy distribution.

One can ask whether Kesten's convergence remains valid for a fixed x. Another question is what happens in higher dimension? In particular :

Question 42. *Is it true that there exists $\rho > 0$, such that when $\mathcal{C}$ is a generic box in $\mathbb{T}^d$ and α is uniformly distributed on $\mathbb{T}^d$, then $\frac{D(\alpha,0,\mathcal{C},N)}{\rho(\ln N)^d}$ converges in distribution to the Cauchy law?*

In Dolgopyat and Fayad [2012] this was proved when x and the box $\mathcal{C}$ are also random (a shape is randomized by applying small deformations distributed according to a smooth measure on the space of isometries). It was shown in Dolgopyat and Fayad [2014] that in the case of a strictly convex shape $\mathcal{C} \subset \mathbb{T}^d$ one has $\frac{D(\alpha,x,r\mathcal{C},N)}{r^{\frac{d-1}{2}} N^{\frac{d-1}{2d}}}$ converges in distribution to a non standard law when $(\alpha, x) \in \mathbb{T}^d \times \mathbb{T}^d$ and $r > 0$ are random. The convex set $r\mathcal{C}$ is the rescaled set from $\mathcal{C}$ by factor r around some fixed point inside $\mathcal{C}$.

A *semialgebraic* set $\mathcal{C}$ in $\mathbb{T}^d$ is a set defined by a finite number of algebraic inequalities. This includes a diverse collection of sets such as balls, cubes, cylinders, simplexes etc. Following Dolgopyat and Fayad [ibid.] we ask

Question 43. *Assume $\mathcal{C}$ is semialgebraic. Does there exist a sequence $a_N = a_N(\mathcal{C})$ such that $\frac{D(\alpha,x,\mathcal{C},N)}{a_N}$ converges in distribution when $(x, \alpha) \in \times\mathbb{T}^d \times \mathbb{T}^d$ are random.*

One can study the fluctuations of the ergodic sums above toral translations for functions other from characteristic functions. The following is interesting for its connection with number theory as well as with the ergodic theory of some natural classes of dynamical systems such as surface flows.

Question 44. *Study the behavior of the ergodic sums $\sum_{n_1=1}^{N} A(x + n\alpha)$ for functions A that are smooth except for a finite number of singularities.*

The fluctuations can be studied for fixed α or x, as well as for random values. One should then try to classify the fluctuations according to the type of the singularities : power, fractional power, logarithmic (we refer to Marklof [2007] and Dolgopyat and Fayad [2015] for more details and questions).

7.2 Higher dimensional actions. Replacing the $\mathbb{Z}$ action by translation with $\mathbb{Z}^k$ actions (see we get following Dolgopyat and Fayad [2015]

Question 45. *Study the ergodic sums* $\sum_{j=1}^{m}\sum_{n_j=1}^{N} A(x + \sum_{j=1}^{m}\alpha_j n_j)$, *with* $(x, \alpha_1, \ldots, \alpha_m) \in (\mathbb{T}^d)^{m+1}$.

In the case where $A = \chi_I - |I|$ and χ_I the indicator of an interval we get the following possible extension of Kesten's theorem to the statistical behavior of linear forms.

Question 46. *Show that as* $x \in \mathbb{T}$ *and* $\alpha \in \mathbb{T}^m$ *are random*

$$\frac{1}{\rho(\ln N)^d}\sum_{j=1}^{m}\sum_{n_j=1}^{N} A(x + \sum_{j=1}^{m}\alpha_j n_j)$$

converges in distribution to a Cauchy law for some $\rho > 0$.

One can also investigate analogues of the Shrinking Targets Theorems of Dolgopyat, Fayad, and Vinogradov [2017] for $\mathbb{Z}^k$ actions.

Question 47. *Let* $l, \hat{l} : \mathbb{R}^d \to \mathbb{R}$, *be linear forms with random coefficients,* $Q : \mathbb{R}^d \to \mathbb{R}$ *be a positive definite quadratic form. Investigate limit theorems, after adequate renormalization, for the number of solutions to*

(a) $\{l(n)\}Q(n) \le c, |n| \le N$;
(b) $\{l(n)\}|\hat{l}(n)| \le c, |n| \le N$;
(c) $|l(n)Q(n)| \le c, |n| \le N$;
(d) $|l(n)\hat{l}(n)| < c, |n| \le N$.

References

D. V. Anosov and Anatole Katok (1970). "New examples in smooth ergodic theory. Ergodic diffeomorphisms". *Trudy Moskov. Mat. Obsc.* 23, pp. 3–36. MR: 0370662 (cit. on p. 1929).

Marie-Claude Arnaud (2011). "A nondifferentiable essential irrational invariant curve for a C^1 symplectic twist map". *J. Mod. Dyn.* 5.3, pp. 583–591. MR: 2854096 (cit. on p. 1936).

V. I. Arnold (1991). "Topological and ergodic properties of closed 1-forms with incommensurable periods". *Funktsional. Anal. i Prilozhen.* 25.2, pp. 1–12, 96. MR: 1142204 (cit. on p. 1941).

A. Avila (n.d.). "KAM, Lyapunov exponents and the spectral dichotomy for one-frequency schrödinger operators". In: In preparation (cit. on p. 1939).

A. Avila and Bassam Fayad (n.d.). "Non-differentiable minimal invariant curves for twist maps of the annulus". In preparation (cit. on p. 1936).

A. Avila, Bassam Fayad, P. Le Calvez, D. Xu, and Z. Zhang (Sept. 2015). "On mixing diffeomorphisms of the disk". arXiv: 1509.06906 (cit. on pp. 1931, 1932).

A. Avila and R. Krikorian (n.d.[a]). "Almost reducibility of circle diffeomorphisms". In preparation.

– (n.d.[b]). "Almost reducibility of pseudo-rotations of the disk". In preparation (cit. on pp. 1929, 1931).

– (n.d.[c]). "Some remarks on local and semi-local results for Schrödinger cocycles". In preparation (cit. on p. 1939).

Artur Avila (June 2010). "Almost reducibility and absolute continuity I". arXiv: 1006.0704 (cit. on p. 1939).

– (2011). "Density of positive Lyapunov exponents for $\mathrm{SL}(2,\mathbb{R})$-cocycles". *J. Amer. Math. Soc.* 24.4, pp. 999–1014. MR: 2813336.

– (2015). "Global theory of one-frequency Schrödinger operators". *Acta Math.* 215.1, pp. 1–54. MR: 3413976 (cit. on pp. 1938, 1939).

Artur Avila, Bassam Fayad, and Raphaël Krikorian (2011). "A KAM scheme for $\mathrm{SL}(2,\mathbb{R})$ cocycles with Liouvillean frequencies". *Geom. Funct. Anal.* 21.5, pp. 1001–1019. MR: 2846380 (cit. on pp. 1939, 1940).

Artur Avila and Raphaël Krikorian (2006). "Reducibility or nonuniform hyperbolicity for quasiperiodic Schrödinger cocycles". *Ann. of Math. (2)* 164.3, pp. 911–940. MR: 2259248 (cit. on pp. 1939, 1940).

– (2015). "Monotonic cocycles". *Invent. Math.* 202.1, pp. 271–331. MR: 3402800 (cit. on pp. 1939, 1940).

József Beck (1994). "Probabilistic Diophantine approximation. I. Kronecker sequences". *Ann. of Math. (2)* 140.2, pp. 449–502. MR: 1298720 (cit. on p. 1944).

K. Bjerklöv and R. Krikorian (n.d.). "On kicked quasi-periodic cocycles and the coexistence of ac and pp spectrum for 1D quasi-periodic Schrödinger operators" (cit. on p. 1939).

Kristian Bjerklöv and Tobias Jäger (2009). "Rotation numbers for quasiperiodically forced circle maps-mode-locking vs. strict monotonicity". *J. Amer. Math. Soc.* 22.2, pp. 353–362. MR: 2476777 (cit. on p. 1940).

Jairo Bochi (2002). "Genericity of zero Lyapunov exponents". *Ergodic Theory Dynam. Systems* 22.6, pp. 1667–1696. MR: 1944399 (cit. on p. 1932).

Abed Bounemoura, Bassam Fayad, and Laurent Niederman (Sept. 2015). "Double exponential stability for generic real-analytic elliptic equilibrium points". arXiv: 1509.00285 (cit. on p. 1936).

– (2017). "Superexponential stability of quasi-periodic motion in Hamiltonian systems". *Comm. Math. Phys.* 350.1, pp. 361–386. MR: 3606478 (cit. on p. 1936).

Barney Bramham (2015). "Pseudo-rotations with sufficiently Liouvillean rotation number are C^0-rigid". *Invent. Math.* 199.2, pp. 561–580. MR: 3302121 (cit. on pp. 1931, 1932).

Chong-Qing Cheng and Lin Wang (2013). "Destruction of Lagrangian torus for positive definite Hamiltonian systems". *Geom. Funct. Anal.* 23.3, pp. 848–866. MR: 3061774 (cit. on p. 1937).

S. G. Dani (1994). "Flows on Homogeneous Spaces and Diophantine Approximation". In: *Proceedings of the International Congress of Mathematicians, Zürich, Switzerland 1994* (cit. on p. 1943).

E. I. Dinaburg and Ja. G. Sinaĭ (1975). "The one-dimensional Schrödinger equation with quasiperiodic potential". *Funkcional. Anal. i Priložen.* 9.4, pp. 8–21. MR: 0470318 (cit. on p. 1939).

Dmitry Dolgopyat and Bassam Fayad (Nov. 2012). "Deviations of ergodic sums for toral translations II. Boxes". arXiv: 1211.4323 (cit. on p. 1945).

– (2014). "Deviations of ergodic sums for toral translations I. Convex bodies". *Geom. Funct. Anal.* 24.1, pp. 85–115. arXiv: 1206.4853. MR: 3177379 (cit. on p. 1945).

– (2015). "Limit theorems for toral translations". In: *Hyperbolic dynamics, fluctuations and large deviations*. Vol. 89. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 227–277. MR: 3309100 (cit. on pp. 1943, 1945).

Dmitry Dolgopyat, Bassam Fayad, and Ilya Vinogradov (2017). "Central limit theorems for simultaneous Diophantine approximations". *J. Éc. polytech. Math.* 4, pp. 1–36. MR: 3583273 (cit. on p. 1946).

Manfred Einsiedler and Elon Lindenstrauss (2006). "Diagonalizable flows on locally homogeneous spaces and number theory". In: *International Congress of Mathematicians. Vol. II*. Eur. Math. Soc., Zürich, pp. 1731–1759. MR: 2275667 (cit. on p. 1943).

L. H. Eliasson (1992). "Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation". *Comm. Math. Phys.* 146.3, pp. 447–482. MR: 1167299 (cit. on pp. 1939, 1940).

– (1998). "Reducibility and point spectrum for linear quasi-periodic skew-products". In: *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*. Extra Vol. II, pp. 779–787. MR: 1648125 (cit. on p. 1938).

L. H. Eliasson, Bassam Fayad, and R. Krikorian (2015). "Around the stability of KAM tori". *Duke Math. J.* 164.9, pp. 1733–1775. MR: 3357183 (cit. on p. 1935).

L. Hakan Eliasson, Bassam Fayad, and Raphaël Krikorian (2013). "KAM-tori near an analytic elliptic fixed point". *Regul. Chaotic Dyn.* 18.6, pp. 801–831. MR: 3146593 (cit. on p. 1935).

Alex Eskin (2010). "Unipotent flows and applications". In: *Homogeneous flows, moduli spaces and arithmetic*. Vol. 10. Clay Math. Proc. Amer. Math. Soc., Providence, RI, pp. 71–129. MR: 2648693 (cit. on p. 1943).

Bassam Fayad (2002). "Analytic mixing reparametrizations of irrational flows". *Ergodic Theory Dynam. Systems* 22.2, pp. 437–468. MR: 1898799 (cit. on p. 1942).

– (2005). "Rank one and mixing differentiable flows". *Invent. Math.* 160.2, pp. 305–340. MR: 2138069 (cit. on p. 1942).

– (2006). "Smooth mixing flows with purely singular spectra". *Duke Math. J.* 132.2, pp. 371–391. MR: 2219261 (cit. on p. 1942).

Bassam Fayad, Giovanni Forni, and Adam Kanigowski (Sept. 2016). "Lebesgue spectrum for area preserving flows on the two torus". arXiv: 1609.03757 (cit. on p. 1942).

Bassam Fayad and Adam Kanigowski (2016). "Multiple mixing for a class of conservative surface flows". *Invent. Math.* 203.2, pp. 555–614. MR: 3455157 (cit. on p. 1943).

Bassam Fayad and Anatole Katok (2004). "Constructions in elliptic dynamics". *Ergodic Theory Dynam. Systems* 24.5, pp. 1477–1520. MR: 2104594 (cit. on p. 1931).

Bassam Fayad and Raphaël Krikorian (2009a). "Herman's last geometric theorem". *Ann. Sci. Éc. Norm. Supér. (4)* 42.2, pp. 193–219. MR: 2518076 (cit. on pp. 1931, 1936).

– (2009b). "Rigidity results for quasiperiodic $\mathrm{SL}(2,\mathbb{R})$-cocycles". *J. Mod. Dyn.* 3.4, pp. 497–510. MR: 2587083 (cit. on pp. 1939, 1940).

Bassam Fayad, J.-P. Marco, and D. Sauzin (n.d.). "Non-resonant elliptic fixed points with an attracted orbit" (cit. on p. 1934).

Bassam Fayad, M. Saprykina, and A. Windsor (2007). "Non-standard smooth realizations of Liouville rotations". *Ergodic Theory Dynam. Systems* 27.6, pp. 1803–1818. MR: 2371596 (cit. on p. 1930).

Bassam Fayad and Maria Saprykina (2005). "Weak mixing disc and annulus diffeomorphisms with arbitrary Liouville rotation number on the boundary". *Ann. Sci. École Norm. Sup. (4)* 38.3, pp. 339–364. MR: 2166337 (cit. on pp. 1930, 1935).

Bassam Fayad and Zhiyuan Zhang (2017). "An effective version of Katok's horseshoe theorem for conservative C^2 surface diffeomorphisms". *J. Mod. Dyn.* 11, pp. 425–445. MR: 3668373 (cit. on p. 1933).

John Franks (1988a). "Generalizations of the Poincaré-Birkhoff theorem". *Ann. of Math. (2)* 128.1, pp. 139–151. MR: 951509 (cit. on p. 1929).

– (1988b). "Recurrence and fixed points of surface homeomorphisms". *Ergodic Theory Dynam. Systems* 8*.Charles Conley Memorial Issue, pp. 99–107. MR: 967632 (cit. on p. 1929).

John Franks and Michael Handel (2012). "Entropy zero area preserving diffeomorphisms of S^2". *Geom. Topol.* 16.4, pp. 2187–2284. MR: 3033517 (cit. on p. 1929).

Xianghong Gong (2012). "Existence of divergent Birkhoff normal forms of Hamiltonian functions". *Illinois J. Math.* 56.1, 85–94 (2013). MR: 3117019 (cit. on p. 1937).

Marcel Guardia and Vadim Kaloshin (Dec. 2014). "Orbits of nearly integrable systems accumulating to KAM tori". arXiv: 1412.7088 (cit. on p. 1934).

B. Hasselblatt and A. Katok, eds. (2002). *Handbook of dynamical systems. Vol. 1A*. North-Holland, Amsterdam, pp. xii+1220. MR: 1928517 (cit. on p. 1943).

Michael-R. Herman (1979). "Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations". *Inst. Hautes Études Sci. Publ. Math.* 49, pp. 5–233. MR: 538680 (cit. on p. 1928).

– (1983). "Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2". *Comment. Math. Helv.* 58.3, pp. 453–502. MR: 727713 (cit. on pp. 1929, 1940).

– (1998). "Some open problems in dynamical systems". In: *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*. Extra Vol. II, pp. 797–808. MR: 1648127.

Xuanji Hou and Jiangong You (2012). "Almost reducibility and non-perturbative reducibility of quasi-periodic linear systems". *Invent. Math.* 190.1, pp. 209–260. MR: 2969277 (cit. on p. 1939).

Vadim Kaloshin and Maria Saprykina (2012). "An example of a nearly integrable Hamiltonian system with a trajectory dense in a set of maximal Hausdorff dimension". *Comm. Math. Phys.* 315.3, pp. 643–697. MR: 2981810 (cit. on p. 1934).

A. Kanigowski, J. Kuaaga-Przymus, and C. Ulcigrai (n.d.). "Multiple mixing and parabolic divergence in smooth area-preserving flows on higher genus surfaces". To appear in *Journal of European Mathematical Society* (cit. on p. 1943).

Nikolaos Karaliolios (2016). "Global aspects of the reducibility of quasiperiodic cocycles in semisimple compact Lie groups". *Mém. Soc. Math. Fr. (N.S.)* 146, pp. 4+ii+200. MR: 3524104 (cit. on p. 1940).

– (2017). "Differentiable rigidity for quasiperiodic cocycles in compact Lie groups". *J. Mod. Dyn.* 11, pp. 125–142. MR: 3627120 (cit. on p. 1940).

Anatole Katok (1975). "Time change, monotone equivalence, and standard dynamical systems". *Dokl. Akad. Nauk SSSR* 223.4, pp. 789–792. MR: 0412383 (cit. on p. 1941).

Anatole Katok, Yakov Pesin, and Federico Rodriguez Hertz, eds. (2017). *Modern theory of dynamical systems*. Vol. 692. Contemporary Mathematics. A tribute to Dmitry Victorovich Anosov. American Mathematical Society, Providence, RI, pp. ix+320. MR: 3666061.

A. V. Kočergin (1972). "The absence of mixing in special flows over a rotation of the circle and in flows on a two-dimensional torus". *Dokl. Akad. Nauk SSSR* 205, pp. 515–518. MR: 0306629 (cit. on p. 1941).

– (1975). "Mixing in special flows over a rearrangement of segments and in smooth flows on surfaces". *Mat. Sb. (N.S.)* 96(138), pp. 471–502, 504. MR: 0516507 (cit. on p. 1941).

A. V. Kochergin (2003). "Nondegenerate fixed points and mixing in flows on a two-dimensional torus". *Mat. Sb.* 194.8, pp. 83–112. MR: 2034533 (cit. on p. 1941).

– (2004). "Nondegenerate fixed points and mixing in flows on a two-dimensional torus". *Mat. Sb.* 195, pp. 317–346 (cit. on p. 1941).

R. Krikorian (n.d.). “Generic divergence of Birkhoff Normal Forms”. In preparation (cit. on p. 1937).

R. Krikorian, J. Wang, J. You, and Q. Zhou (n.d.). “Linearization of quasiperiodically forced circle flow beyond Brjuno condition”. To apppear in *Comm. Math. Phys.* (cit. on p. 1940).

Raphaël Krikorian (1999a). “Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts”. *Astérisque* 259, pp. vi+216. MR: 1732061 (cit. on p. 1939).

– (1999b). “Réductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts”. *Ann. Sci. École Norm. Sup. (4)* 32.2, pp. 187–240. MR: 1681809 (cit. on p. 1939).

– (2001). “Global density of reducible quasi-periodic cocycles on $\mathbf{T}^1 \times \mathrm{SU}(2)$”. *Ann. of Math. (2)* 154.2, pp. 269–326. MR: 1865972 (cit. on p. 1939).

L. Kuipers and H. Niederreiter (1974). *Uniform distribution of sequences*. Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, pp. xiv+390. MR: 0419394 (cit. on p. 1944).

Jens Marklof (2006). “Energy level statistics, lattice point problems, and almost modular functions”. In: *Frontiers in number theory, physics, and geometry. I*. Springer, Berlin, pp. 163–181. MR: 2261094 (cit. on p. 1943).

– (2007). “Distribution modulo one and Ratner’s theorem”. In: *Equidistribution in number theory, an introduction*. Vol. 237. NATO Sci. Ser. II Math. Phys. Chem. Springer, Dordrecht, pp. 217–244. MR: 2290501 (cit. on pp. 1943, 1945).

S. P. Novikov (1982). “The Hamiltonian formalism and a multivalued analogue of Morse theory”. *Uspekhi Mat. Nauk* 37.5(227), pp. 3–49, 248. MR: 676612 (cit. on p. 1941).

Ricardo Pérez-Marco (2003). “Convergence or generic divergence of the Birkhoff normal form”. *Ann. of Math. (2)* 157.2, pp. 557–574. MR: 1973055 (cit. on p. 1937).

Leonid Polterovich and Mikhail Sodin (2004). “A growth gap for diffeomorphisms of the interval”. *J. Anal. Math.* 92, pp. 191–209. MR: 2072746 (cit. on p. 1932).

Jürgen Pöschel (1982). “Integrability of Hamiltonian systems on Cantor sets”. *Comm. Pure Appl. Math.* 35.5, pp. 653–696. MR: 668410 (cit. on p. 1937).

Joaquim Puig (2004). “Cantor spectrum for the almost Mathieu operator”. *Comm. Math. Phys.* 244.2, pp. 297–309. MR: 2031032 (cit. on p. 1939).

– (2006). “A nonperturbative Eliasson’s reducibility theorem”. *Nonlinearity* 19.2, pp. 355–376. MR: 2199393 (cit. on p. 1939).

Helmut Rüssmann (1967). “Über die Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung”. *Math. Ann.* 169, pp. 55–72. MR: 0213679 (cit. on pp. 1930, 1935).

Carl Ludwig Siegel (1954). “Über die Existenz einer Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslösung”. *Math. Ann.* 128, pp. 144–170. MR: 0067298 (cit. on p. 1937).

Ya. G. Sinaĭ and K. M. Khanin (1992). "Mixing of some classes of special flows over rotations of the circle". *Funktsional. Anal. i Prilozhen.* 26.3, pp. 1–21. MR: 1189019 (cit. on p. 1941).

Corinna Ulcigrai (2011). "Absence of mixing in area-preserving flows on surfaces". *Ann. of Math. (2)* 173.3, pp. 1743–1778. MR: 2800723 (cit. on p. 1941).

D. Xu (2016). "Lyapunov exponents and rigidity". PhD thesis. Univ. Paris Diderot, Paris 7 (cit. on pp. 1939, 1940).

Disheng Xu (June 2015). "Density of positive Lyapunov exponents for symplectic cocycles". arXiv: 1506.05403 (cit. on p. 1940).

J.-C. Yoccoz (1984). "Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne". *Ann. Sci. École Norm. Sup. (4)* 17.3, pp. 333–359. MR: 777374 (cit. on p. 1928).

Jean-Christophe Yoccoz (1995a). "Centralisateurs et conjugaison différentiable des difféomorphismes du cercle". *Astérisque* 231. Petits diviseurs en dimension 1, pp. 89–242. MR: 1367354.

– (1995b). "Centralisateurs et conjugaison différentiable des difféomorphismes du cercle". *Astérisque* 231. Petits diviseurs en dimension 1, pp. 89–242. MR: 1367354 (cit. on p. 1929).

J. You (2018). "Quantitative almost reducibility and its applications". In: *Proc. of the International Congress of Mathematicians, Rio de Janeiro 2018*. Ed. by Boyan Sirakov, Paulo Ney de Souza, and Marcelo Viana. In this volume. World Scientific, pp. 2113–2136 (cit. on p. 1938).

Jiangong You and Qi Zhou (2013). "Embedding of analytic quasi-periodic cocycles into analytic quasi-periodic linear systems and its applications". *Comm. Math. Phys.* 323.3, pp. 975–1005. MR: 3106500 (cit. on p. 1939).

Received 2018-01-13.

Bassam Fayad
bassam.fayad@gmail.com
bassam@math.jussieu.fr

Raphaël Krikorian
Université de Cergy-Pontoise
raphael.krikorian@u-cergy.fr

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (1951–1966)

SUBADDITIVE COCYCLES AND HOROFUNCTIONS

SÉBASTIEN GOUËZEL

Abstract

Subadditive cocycles are the random version of subadditive sequences. They play an important role in probability and ergodic theory, notably through Kingman's theorem ensuring their almost sure convergence. We discuss a variation around Kingman's theorem, showing that a subadditive cocycle is in fact almost additive at many times. This result is motivated by the study of the iterates of deterministic or random semicontractions on metric spaces, and implies the almost sure existence of a horofunction determining the behavior at infinity of such a sequence. In turn, convergence at infinity follows when the geometry of the space has some features of nonpositive curvature.

The aim of this text is to present and put in perspective the results we have proved with Anders Karlsson in the article Gouëzel and Karlsson [2015]. The topic of this article is the study, in an ergodic theoretic context, of some subadditivity properties, and their relationships with dynamical questions with a more geometric flavor, dealing with the asymptotic behavior of random semicontractions on general metric spaces. This text is translated from an article written in French on the occasion of the first congress of the French Mathematical Society Gouëzel, Sébastien [2017]. The proof of the main ergodic-theoretic result in Gouëzel and Karlsson [2015] has been completely formalized and checked in the computer proof assistant Isabelle/HOL Gouëzel, Sébastien [2016].

1 Iteration of a semicontraction on Euclidean space

In order to explain the problems we want to consider, it is enlightening to start with a more elementary example, showing how subadditivity techniques can be useful to understand a deterministic semicontraction. In the next section, we will see how these results can be extended to random semicontractions.

MSC2010: primary 37H15; secondary 37A30.
Keywords: subadditive cocycles, ergodic theory, Kingman theorem, horofunctions, semicontractions.

Definition 1.1. *A transformation T on a metric space X is a semicontraction if it is* 1*-Lipschitz, i.e., if $d(T(x), T(y)) \leqslant d(x, y)$ for all $x, y \in X$.*

If T is a semicontraction, its iterates also are. Hence, for any points x and y, the distance between $T^n(x)$ and $T^n(y)$ remains uniformly bounded, by $d(x, y)$ (where we write $T^n = T \circ \cdots \circ T$). As a consequence, the asymptotic behavior of $T^n(x)$ (up to bounded error) is independent of x.

In the Euclidean space $\mathbb{R}^d$, the first examples of semicontractions are given by translations (where $T^n(x)$ tends to infinity as $nv + O(1)$, where v is the translation vector) and homotheties with ratio $\leqslant 1$ (for which $T^n(x)$ remains bounded). The following theorem, proved in 1981 in Kohlberg and Neyman [1981] under slightly stronger assumptions, shows that these examples are typical since there always exists an asymptotic translation vector. The proof we give is due to Karlsson [2001].

Theorem 1.2. *Consider a semicontraction $T : X \to X$ on a subset X of Euclidean space $\mathbb{R}^d$. Then there exists a vector v such that $T^n(x)/n$ converges to v for all $x \in X$.*

Note that the asymptotic behavior of $T^n(x)/n$ does not depend on x, therefore it suffices to prove the theorem for one single point x. Translating everything if necessary, we can assume $0 \in X$ and take $x = 0$ to simplify notations.

The proof relies crucially on the subadditivity properties of the sequence $u_n = d(0, T^n(0))$.

Definition 1.3. *A sequence $(u_n)_{n\in\mathbb{N}}$ of real numbers is* subadditive *if $u_{k+\ell} \leqslant u_k + u_\ell$ for all k, ℓ.*

The main property of such a sequence is given in the next lemma, due to Fekete.

Lemma 1.4. *Let u_n be a subadditive sequence. Then u_n/n converges, to* $\operatorname{Inf}\{u_n/n, n > 0\} \in \mathbb{R} \cup \{-\infty\}$.

Proof. Fix a positive integer N. It follows from the subadditivity of u that $u_{kN+r} \leqslant ku_N + u_r$. Writing an arbitrary integer n as $kN + r$ with $r < N$, dividing by n and taking the limit, we get $\limsup u_n/n \leqslant u_N/N$. Hence, $\limsup u_n/n \leqslant \operatorname{Inf}\{u_N/N\}$. The result follows as $\liminf u_n/n \geqslant \operatorname{Inf}\{u_N/N\}$. □

Recalling the notation $u_n = d(0, T^n(0))$, we have

$$
\begin{aligned}
u_{k+\ell} &= d(0, T^{k+\ell}(0)) \leqslant d(0, T^k(0)) + d(T^k(0), T^k(T^\ell(0))) \\
&\leqslant d(0, T^k(0)) + d(0, T^\ell(0)) = u_k + u_\ell,
\end{aligned}
\tag{1-1}
$$

where we used the semicontractivity of T^k. Hence, Fekete's Lemma shows that u_n/n converges to a limit $A \geqslant 0$. At time n, the point $T^n(0)$ is close to the sphere of radius

An centered at 0. If $A = 0$, this proves Theorem 1.2. However, if $A > 0$, we should also prove the directional convergence of $T^n(0)$. For this, we will use times where the sequence u is almost additive, given by the following lemma.

Lemma 1.5. *Let $\varepsilon > 0$. Consider a subadditive sequence u_n such that $u_n/n \to A \in \mathbb{R}$. Then there exist arbitrarily large integers n such that, for all $1 \leqslant \ell \leqslant n$,*

$$u_n \geqslant u_{n-\ell} + (A-\varepsilon)\ell. \tag{1-2}$$

As u_ℓ is of magnitude $A\ell$, this inequality can informally be read as $u_n \geqslant u_{n-\ell} + u_\ell - \delta$, where δ is small. It entails additivity of the sequence at *all* intermediate times between 1 and n, up to a well controlled error.

Proof. The sequence $u_n - (A-\varepsilon)n$ is equivalent to εn, and tends therefore to infinity. In particular, there exist arbitrarily large times n which are records for this sequence, beating every previous value. For such an n, we have for $\ell \leqslant n$ the inequality $u_{n-\ell} - (A-\varepsilon)(n-\ell) \leqslant u_n - (A-\varepsilon)n$, which is equivalent to the result we claim. □

For $\varepsilon_i = 2^{-i}$, let us consider a corresponding sequence of times n_i given by Lemma 1.5, tending to infinity. Let h_i be a norm-1 linear form, equal to $-\|T^{n_i}(0)\|$ on $T^{n_i}(0)$. Then, for all $\ell \leqslant n_i$,

$$\begin{aligned} h_i(T^\ell(0)) &= h_i(T^\ell(0) - T^{n_i}(0)) + h_i(T^{n_i}(0)) \leqslant \|T^\ell(0) - T^{n_i}(0)\| - \|T^{n_i}(0)\| \\ &\leqslant \|T^{n_i-\ell}(0)\| - \|T^{n_i}(0)\| = u_{n_i-\ell} - u_{n_i} \leqslant -(A-\varepsilon_i)\ell, \end{aligned}$$

where the last inequality follows from (1-2). In the inequality $h_i(T^\ell(0)) \leqslant -(A-\varepsilon_i)\ell$ that we just obtained, it is remarkable that every mention of n_i has disappeared.

Let us now consider h a limit (weak or strong, as we are in finite dimension) of the sequence h_i, it is a norm-1 linear form. As ε_i tends to 0 with i, we deduce from the above the following inequality:

$$\text{for every integer } \ell, \quad h(T^\ell(0)) \leqslant -A\ell. \tag{1-3}$$

This inequality entails that $T^\ell(0)$ belongs to the half-space directed by h, at distance $A\ell$ from the origin. As it also has essentially norm $A\ell$, we deduce that it is essentially pointing in the direction of h (see Figure 1). This shows the convergence of $T^\ell(0)/\ell$. If one wants a more explicit argument, one can for instance consider a cluster value v of $T^\ell(0)/\ell$. It is a vector of norm A, satisfying $h(v) = -A$. As h has norm 1, this determines uniquely v thanks to the strict convexity of the Euclidean norm. Therefore, $T^\ell(0)/\ell$ has a unique cluster value, and it converges. This concludes the proof of Theorem 1.2. □

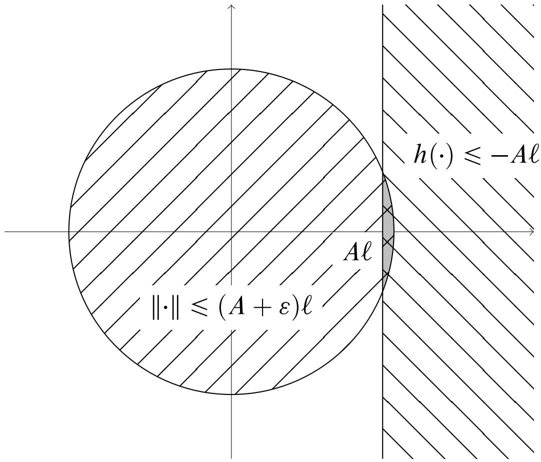


Figure 1: $T^\ell(0)$ belongs to the intersection of the dashed areas

Remark 1.6. The proof has not used finite-dimensionality (if one replaces strong limits with weak limits). Therefore, the result is still true in Hilbert spaces, or more generally in uniformly convex Banach spaces.

Remark 1.7. Most of the proof is valid in a general Banach space: there always exists a linear form h with norm at most 1 such that $h(T^\ell(0)) \leqslant -A\ell$ for all ℓ (this already implies non-trivial results, for instance the sequence $(T^\ell(0))_{\ell \geqslant 0}$ is contained in a half-space if $A > 0$, as the linear form h is necessarily nonzero in this case). The only point where the proof breaks is the last argument, relying on strict convexity of the norm.

One may wonder if this is a limitation of the proof, or if the proof captures all the relevant information. In fact, the above theorem is wrong without convexity assumptions on the norm. Let us describe quickly a counter-example due to Kohlberg and Neyman [1981], in $\mathbb{R}^2$ with the sup norm. Let us fix the two vectors $v_+ = (1,1)$ and $v_- = (1,-1)$, both of norm 1. We define a continuous path $\gamma : \mathbb{R}_+ \to \mathbb{R}^2$ starting from 0, of the form $\gamma(t) = (t, \varphi(t))$, by following the direction v_+ during a time S_0, then the direction v_- during a time $S_1 \gg S_0$, then the direction v_+ during a time $S_2 \gg S_1$, and so on. One can ensure that the angle between $\gamma(t)$ and the horizontal line fluctuates between $-\pi/4$ et $\pi/4$. As the slopes of v_+ and v_- are 1, the path γ is an isometry from $\mathbb{R}_+$ onto its image. Let $h(x_1, x_2) = x_1$ be the first coordinate. Then the map $T : x \mapsto \gamma(|h(x)| + 1)$ is a semicontraction, as a composition of 1-Lipschitz functions. One checks easily that $T^n(0) = \gamma(n)$. Therefore, by construction, $T^n(0)/n$ does not converge.

2 Horofunctions

Many interesting geometric spaces, which are not vector spaces, have semicontractions. We would like to have a version of Theorem 1.2 for these spaces. The conclusion of the theorem can not be of the form "$T^n(x)/n$ is converging" as division by n makes no sense. It is always true that $d(T^n(x), x)/n$ converges to a limit $A \geqslant 0$, by subadditivity. However, the meaning to give to directional convergence is less obvious. Such theorems already exist in different contexts. Let us mention for instance the following Denjoy-Wolff Theorem Denjoy [1926] and Wolff [1926]:

Theorem 2.1. *Let T be a holomorphic map from the unit disk $\mathbb{D}$ in $\mathbb{C}$ into itself. Then, either T has a fixed point in the disk, or $T^n(0)$ converges to a point on the unit circle.*

This statement is indeed a particular case of the previous discussion, as a holomorphic map of the unit disk if a semicontraction for the hyperbolic distance.

In the general case, the counter-example from Remark 1.7 shows that one can not hope to have convergence at infinity in a strong sense without additional assumptions of geometric nature on the space. If we follow the proof of Theorem 1.2 in the context of a general metric space, we see that it is possible to make sense of all arguments up to the inequality (1-3), in terms of horofunctions.

Definition 2.2. *Let (X, d) be a metric space with a basepoint x_0. For $x \in X$, we say that the function $h_x : y \mapsto d(x, y) - d(x, x_0)$ is an* internal horofunction. *A* horofunction *is an element of the closure of the set of internal horofunctions, for the topology of pointwise convergence.*

For every $x \in X$, the internal horofunction h_x vanishes at x_0 and it is 1-Lipschitz. Therefore, $h_x(y)$ belongs to the compact interval $[-d(y, x_0), d(y, x_0)]$. As a product of compact spaces is compact for the topology of pointwise convergence (i.e., the product topology), we deduce that the set $\overline{X}^B$ of horofunctions, endowed with the topology of pointwise convergence, is a compact space in which the set X (seen as the set of internal horofunctions) is dense. A horofunction vanishes at x_0 and is 1-Lipschitz, as these properties are preserved by pointwise limits.

In the same way that we distinguish between a point $x \in X$ and the corresponding internal horofunction, we will distinguish by the notations between an abstract point $\xi \in \overline{X}^B$ and the corresponding horofunction h_ξ.

Remark 2.3. In general, X is *not* an open subset of $\overline{X}^B$, contrary to the usual requirements for compactifications. For instance, consider for X a countable number of rays $\mathbb{R}^+$, all coming from the same point x_0, with the graph distance. If a sequence converges to infinity along one of the rays (say the ray with index i), then the sequence of corresponding internal horofunctions converges to an (external) horofunction h_i. When i tends to infinity, one checks easily that h_i tends to h_{x_0}.

On the other hand, if the space is proper (i.e., every closed ball $\overline{B}(x, r)$ is compact) and geodesic (between any two points x and y, there is a geodesic, i.e., a path isometric to the segment $[0, d(x, y)]$), then $\overline{X}^B$ is a compactification of X in the usual sense.

One should think of external horofunctions as analogues of linear forms, but on general metric spaces. In the case of Euclidean space, the two notions coincide exactly. In geometric terms, what is interesting is not so much the horofunction h itself, than the sequence of horoballs $\{x : h(x) \leqslant c\}$ it defines, for $c \in \mathbb{R}$. This is a kind of family of half-spaces, increasing with c, defining a direction at infinity when $c \to -\infty$.

The notion of horofunction is exactly the one we need to extend the above proof of Theorem 1.2 to a general metric space:

Theorem 2.4 (Karlsson [2001]). *Let T be a semicontraction on a metric space (X, d) with a basepoint x_0. Then $d(T^n(x_0), x_0)/n$ converges to a limit $A \geqslant 0$. Moreover, there exists a horofunction h such that, for all $\ell \in \mathbb{N}$, we have $h(T^\ell(x_0)) \leqslant -A\ell$.*

Proof. The proof is exactly the same as the proof of the inequality (1-3), if one replaces the notion of linear form (which relied on the linearity of the underlying space) with the notion of horofunction. Indeed, let us define A as in the proof of this theorem, by subadditivity. Let $\varepsilon_i = 2^{-i}$, and consider an increasing sequence n_i such that, for all $\ell \leqslant n_i$, holds $d(x_0, T^{n_i} x_0) \geqslant d(x_0, T^{n_i - \ell} x_0) + (A - \varepsilon_i)\ell$, thanks to Lemma 1.5. Then, we use the internal horofunction based at $T^{n_i} x_0$. It satisfies, for $\ell \leqslant n_i$,

$$\begin{aligned} h_{T^{n_i}(x_0)}(T^\ell(x_0)) &= d(T^{n_i}(x_0), T^\ell(x_0)) - d(T^{n_i}(x_0), x_0) \\ &\leqslant d(T^{n_i - \ell}(x_0), x_0) - d(T^{n_i}(x_0), x_0) \\ &\leqslant -(A - \varepsilon_i)\ell. \end{aligned}$$

This shows that the set of horofunctions satisfying $h(T^\ell(x_0)) \leqslant -(A - \varepsilon_i)\ell$ for all $\ell \leqslant n_i$ is nonempty. Moreover, it is compact, and decreases with i. As the set of horofunctions is compact, the intersection of these sets is nonempty. Any element h of this intersection satisfies $h(T^\ell(x_0)) \leqslant -A\ell$ for all ℓ, as desired. (In the case where the space X is second-countable, the topology on $\overline{X}^B$ is metrizable, and one can just take for h any cluster value of the sequence $h_{T^{n_i}(x_0)}$.) □

This theorem entails that $T^\ell(x_0)$ is in the intersection of the ball of radius $(A + \varepsilon)\ell$ and of the half-space $\{h \leqslant -A\ell\}$ for ℓ large enough, as in Figure 1, with the difference that the shapes of the ball and the half-space depend on the geometry of (X, d). Deciding if one can deduce from this statement a stronger convergence at infinity will thus depend on X. For instance, this is true in a uniformly convex Banach space, thanks to Remark 1.6, but this is false in $\mathbb{R}^2$ with the sup norm, by Remark 1.7.

One can therefore say that Theorem 2.4 decouples the dynamics from the geometry, capturing all the information about iterations of semicontractions on metric spaces, and reducing the question of convergence at infinity to a purely geometric question on the geometric shape of horofunctions.

Example 2.5. An important class of metric spaces is the CAT(0) spaces, i.e., metric spaces (they do not have to be manifolds) which have non-positive curvature in an extended sense, see Bridson and Haefliger [1999]. In such a space, there is a natural geometric notion of boundary at infinity, which turns out to be in bijection with external horofunctions. Moreover, the horofunctions can be described with sufficient precision to extend the argument given above in Euclidean space: If a sequence satisfies $d(x_n, x_0)/n \to A > 0$ and $h(x_n)/n \to -A$ where h is a horofunction, then x_n converges to the point at infinity corresponding to h. This applies to $x_n = T^n(x_0)$ when T is a semicontraction. We obtain a generalization of Theorem 1.2 to a much broader class of metric spaces.

A weakness of the previous result is that it does not give much when $A = 0$. For instance, it does not seem to reprove Theorem 2.1 of Denjoy and Wolff when $A = 0$ (while the convergence to a point on the boundary follows directly when $A > 0$, as the disk with the hyperbolic distance is CAT(0) – and even CAT(-1)). In fact, one can fully recover Theorem 2.1 from Theorem 2.4 thanks to the following lemma due to Całka Całka [1984], for which we give a direct proof.

Lemma 2.6. *Let T be a semicontraction of a proper metric space. Let $x_0 \in X$. If there exists a subsequence n_i along which $d(x_0, T^{n_i} x_0)$ stays bounded, then the whole sequence $d(x_0, T^n x_0)$ is bounded.*

Proof. Let $\mathcal{O}$ be the orbit of x_0. It has a cluster point x_1 by assumption. Let $B = \overline{\mathcal{O}} \cap \overline{B}(x_1, 1)$. By properness, B is covered by a finite number of balls $B_i = \overline{\mathcal{O}} \cap \overline{B}(x_i, 1/2)$, with $x_i \in \mathcal{O}$. For each i, choose $k_i > 0$ such that $T^{k_i}(x_i) \in \overline{B}(x_1, 1/2)$, this is possible as x_1 is a cluster point of $\mathcal{O}$. Then $T^{k_i}(B_i) \subseteq B$ as T is a semicontraction.

Consider now $n > \max k_i$. Then

$$T^n(B) \subseteq \bigcup_i T^n(B_i) = \bigcup_i T^{n-k_i}(T^{k_i} B_i) \subseteq \bigcup_i T^{n-k_i} B \subseteq \bigcup_{m<n} T^m(B).$$

We deduce by induction that $T^n(B) \subseteq \bigcup_{m \leq \max k_i} T^m(B)$. Hence, $\bigcup_n T^n(B)$ is within bounded distance of x_0. Finally, x_0 has an iterate that enters B. All its subsequent iterates remain in the above set. □

Proof of Theorem 2.1 of Denjoy-Wolff. We endow the unit disk with the hyperbolic distance, for which any holomorphic map is a semicontraction.

Assume first that $T^n(0)$ stays bounded for this distance. Then

$$K = \bigcap_n \overline{\bigcup_{m \geqslant n} \{T^m(0)\}}$$

is a nonempty compact set, satisfying $T(K) = K$. The set K is contained in a unique ball of minimal radius (this is a general property of nonpositive curvature, see Bridson and Haefliger [1999, Proposition 2.7]) that we denote by $\overline{B}(x, r)$. Then $K = T(K)$ is included in $\overline{B}(T(x), r)$ as T is a semicontraction. By uniqueness, $x = T(x)$, and T has a fixed point.

Assume now that $T^n(0)$ is unbounded. By Lemma 2.6, it tends to infinity in the hyperbolic disk, i.e., to the unit circle S^1 in $\mathbb{C}$. Moreover, Theorem 2.4 shows that the sequence $T^n(0)$ stays in a horoball $\{x : h(x) \leqslant 0\}$ for some horofunction h. In this setting, horoballs are Euclidean disks with 0 in their boundary and tangent to the unit circle. In particular, the closure of such a horoball meets S^1 at a unique point, to which $T^n(0)$ must converge. □

3 Iteration of random semicontractions

The problem of interest to us is the composition of random semicontractions. Let us describe it in the simplest case. Fix a metric space (X, d) with a basepoint x_0, consider a finite number of semicontractions $T_1, \dots, T_I$ on X, and fix a probability measure $\mathbb{P}_0$ on $\{1, \dots, I\}$, i.e., a sequence of positive real numbers $p_i > 0$ with $\sum p_i = 1$. Then we can describe a left random walk L_n on X as follows. At time 0, let $L_0 = x_0$. Then, choose randomly a semicontraction $T^{(1)}$ among $T_1, \dots, T_I$, taking T_i with probability p_i, and jump to $L_1 = T^{(1)}(x_0)$. Then, choose $T^{(2)}$ like $T^{(1)}$, independently of the choices already made, and jump to $L_2 = T^{(2)}(L_1)$. And so on. Formally,

$$L_n = T^{(n)} \circ \dots \circ T^{(1)}(x_0),$$

where the $T^{(k)}$ are random semicontractions, chosen independently according to the distribution $\mathbb{P}_0$. We should write $T^{(k)} = T^{(k)}(\omega)$ and $L_n = L_n(\omega)$ where ω is a random parameter, living in a probability space which parameterizes all objects we use (here, we can take $\Omega = \{1, \dots, I\}^{\mathbb{N}}$ with the probability measure $\mathbb{P} = \mathbb{P}_0^{\otimes \mathbb{N}}$). As usual in probability theory, we will not write explicitly the parameter ω to get simpler formulas (but it will reappear in the more general context we will describe later on).

One can also consider a right random walk R_n given by

$$R_n = T^{(1)} \circ \dots \circ T^{(n)}(x_0).$$

Its geometric meaning is less clear at first sight, but its convergence behavior is much better as we will explain now. In general, L_n can be very far away from L_{n-1}, while

$$d(R_n, R_{n-1}) = d(T^{(1)} \circ \cdots \circ T^{(n)}(x_0), T^{(1)} \circ \cdots \circ T^{(n-1)}(x_0)) \leqslant d(T^{(n)}(x_0), x_0),$$

where the last inequality follows from the fact that $T^{(1)} \circ \cdots \circ T^{(n-1)}$ is a semicontraction. Therefore, R_n is within bounded distance of R_{n-1}. The random walk R_n makes bounded jumps, contrary to L_n.

Example 3.1. The isometries of the hyperbolic disk are of three type: elliptic (with a fixed point inside the disk), parabolic (with a unique fixed point on the boundary, the dynamics is a rotation on horospheres centered at this point) and loxodromic (with two fixed points on the boundary, one attractive and one repulsive). Assume that all T_i are loxodromic isometries, with attractive fixed point ξ_i. We expect that, independently of the position of L_{n-1}, the map $T^{(n)} = T_{i_n}$ sends it close to its attractive point ξ_{i_n}. Hence, the sequence L_n should tend towards the circle at infinity, but alternate between different possible limit points since, almost surely, i_n will take every value in $\{1, \dots, I\}$ infinitely often when n tends to infinity. In particular, we should not expect L_n to converge typically. On the other hand, in R_n, the map that is applied last is always $T^{(1)}$, so that R_n should be close to ξ_{i_1}, up to an error depending on the next terms in the sequence. As the maps we are composing are contractions on the boundary (away from their repulsive fixed point), the influence of the n-th map should be exponentially small. Therefore, R_n should typically be a Cauchy sequence in $\overline{\mathbb{D}}$, and it should converge (to a random limit, that depends on the random parameter ω). In this geometric context, this heuristic description is correct (the almost sure convergence of R_n to a limit point is due to Furstenberg, in a broader context).

We will consider a more general setting, encompassing the previous one, in which the semicontractions we compose are not any more independent from each other.

Let us consider a space Ω with a probability measure $\mathbb{P}$ and a measurable map U which preserves the measure (i.e., for every measurable subset B, we have $\mathbb{P}(U^{-1}B) = \mathbb{P}(B)$). We will moreover assume that U is ergodic: any measurable set B with $U^{-1}(B) = B$ has measure 0 or 1. Finally, let us fix a map $\omega \mapsto T(\omega)$ associating to $\omega \in \Omega$ a semicontraction $T(\omega)$ on the space (X, d), in a measurable way. We also require an integrability assumption: we will always assume $\int d(x_0, T(\omega)x_0)\, \mathrm{d}\mathbb{P}(\omega) < \infty$. Then we can define "random walks" on (X, d) as follows. Writing x_0 for a basepoint in X, let $L_n(\omega) = T(U^{n-1}(\omega)) \circ \cdots \circ T(\omega)(x_0)$ and $R_n(\omega) = T(\omega) \circ \cdots \circ T(U^{n-1}\omega)(x_0)$. We will mainly be interested in $R_n(\omega)$, since this is the walk for which one can expect convergence results, as explained in Example 3.1. Therefore, let us write $T^n(\omega) = T(\omega) \circ \cdots \circ T(U^{n-1}\omega)$.

This setting is a generalization of the case of random compositions: It is recovered by taking $\Omega = \{1, \cdots I\}^{\mathbb{N}}$ and $\mathbb{P} = \mathbb{P}_0^{\mathbb{N}}$ and U the left shift (given by $U((\omega_k)_{k \in \mathbb{N}}) =$

$(\omega_{k+1})_{k\in\mathbb{N}})$ and $T((\omega_k)_{k\in\mathbb{N}}) = T_{\omega_0}$. The non-independent case is in general more delicate to study since several probabilistic tools do not apply any more (for instance, Furstenberg's proof in Example 3.1 relies on the martingale convergence theorem, which does not hold in this broader context).

The setting we have studied in Sections 1 and 2, of a single semicontraction, is also a particular case of the general setting, taking Ω reduced to a point. We can ask how much of the results proved in this particular case extend to the general situation.

The first result (asymptotic behavior of the distance to the origin) follows directly from an ergodic theorem, Kingman's Theorem, which is the analogue of Fekete's Lemma in an ergodic context. We will come back later to this statement, given below as Theorem 4.2. This theorem readily implies the following:

Proposition 3.2. *There exists $A \geqslant 0$ such that $d(x_0, T^n(\omega)x_0)/n \to A$ for almost every ω.*

To go further and obtain a directional convergence, we would like the analogue of Theorem 2.4, i.e., obtain for almost every ω a horofunction h^ω describing the asymptotic behavior of the walk. This result is considerably more delicate. We have proved it in full generality with Karlsson in Gouëzel and Karlsson [2015], after several partial results:

Theorem 3.3 (Karlsson and Margulis [1999]). *Let $\varepsilon > 0$. For almost every ω, there exists a horofunction h^ω such that all cluster values of the sequence $h^\omega(T^n(\omega)x_0)/n$ belong to the interval $[-A, -A+\varepsilon]$.*

Theorem 3.4 (Karlsson and Ledrappier [2006]). *Assume moreover that all the maps $T(\omega)$ are isometries of X. For almost every ω, there exists a horofunction h^ω such that $h^\omega(T^n(\omega)x_0)/n \to -A$.*

Theorem 3.5 (Gouëzel and Karlsson [2015]). *Without further assumptions, for almost every ω, there exists a horofunction h^ω such that $h^\omega(T^n(\omega)x_0)/n \to -A$.*

The last theorem realizes the full decoupling between dynamics and geometry that we already explained in Section 2 for the dynamics of a single semicontraction: Assume that, for $A > 0$, a sequence satisfying $d(x_n, x_0) \sim An$ and $h(x_n) \sim -An$ converges necessarily towards a point in a given geometric compactification of X (this is purely a geometric property of X and its compactification). Then we deduce that, for almost every ω, the sequence $T^n(\omega)x_0$ converges in the compactification. This is for instance the case when X is CAT(0), as explained in Example 2.5. However, we note that, in the CAT(0) case, Theorem 3.3 is sufficient to obtain this convergence (see Karlsson and Margulis [1999]), thanks to additional geometric arguments (that can be completely avoided if one uses Theorem 3.5).

These theorems have many applications in different contexts. For instance, if one applies them to isometries of the symmetric space associated to $\mathrm{GL}(d, \mathbb{R})$ (which is CAT(0), so that any of the above theorems would suffice), one can recover Oseledets' Theorem on random products of matrices. One can also obtain a random version of the theorem of Denjoy and Wolff (Theorem 2.1), or applications to operator theory, to Teichmüller theory. This note is not devoted to applications, we refer the reader to the articles cited above. We rather want to explore a little bit the proofs of these statements: contrary to the intuition, they have nothing geometric, they rely exclusively on subadditivity arguments (just like the proofs in Sections 1 and 2).

This is not completely true for the proof given by Karlsson and Ledrappier of Theorem 3.4: they take advantage of the fact that the maps are isometries by arguing that isometries act on the set of horofunctions. One can then use a cocycle on this space, which is geometric in spirit. However, this is true for the proofs of Theorems 3.3 et 3.5, that we will sketch in the next section.

Remark 3.6. Theorem 3.5 constructs a horofunction that satisfies $h^\omega(T^n(\omega)x_0) \leqslant -An + o(n)$, which is weaker than the conclusion of Theorem 2.4 giving $h(T^n x_0) \leqslant -An$ in the case of a single semicontraction. It is easy to see that it is impossible to get such a strong conclusion in the random case: it would for instance imply that the points $T^n(\omega)x_0$ would almost surely stay in a horoball. This is not the case if the $T(\omega)$ can go in every direction, for instance if one chooses on $\mathbb{R}$ uniformly between the translation of 2 and the translation of -1 (we have chosen two vectors with different norms to ensure that A is nonzero).

4 Ergodic theory and subadditivity

The analogue of subadditive sequences in a dynamical setting is given by the notion of subadditive cocycle (this terminology is very bad, as a subadditive cocycle is not a cocycle, the word subcocycle would certainly be better, but it is too late to change).

Definition 4.1. *Let $(\Omega, \mathbb{P})$ be a probability space and $U : \Omega \to \Omega$ an ergodic map preserving the measure. A measurable function $u : \mathbb{N} \times \Omega \to \mathbb{R}$ is a* subadditive cocycle *if, for all k, ℓ and for almost every ω,*

$$u(k + \ell, \omega) \leqslant u(k, \omega) + u(\ell, U^k\omega).$$

A subadditive cocycle is integrable *if $\int u^+(1, \omega)\, \mathrm{d}\mathbb{P}(\omega) < \infty$, where u^+ is the positive part of u.*

Let us consider for instance a family of semicontractions $T(\omega)$ depending measurably on $\omega \in \Omega$. Let $u(n, \omega) = d(x_0, T^n(\omega)(x_0))$. This is a subadditive cocycle: for all k, ℓ

and ω, we have

$$\begin{aligned} u(k+\ell,\omega) &= d(x_0, T^{k+\ell}(\omega)x_0) = d(x_0, T^k(\omega)(T^\ell(U^k\omega)(x_0))) \\ &\leqslant d(x_0, T^k(\omega)(x_0)) + d(T^k(\omega)(x_0), T^k(\omega)(T^\ell(U^k\omega)(x_0))) \\ &\leqslant d(x_0, T^k(\omega)(x_0)) + d(x_0, T^\ell(U^k\omega)(x_0)) = u(k,\omega) + u(\ell, U^k\omega), \end{aligned}$$

where we used the triangular inequality to go from the first to the second line, and the fact that $T^k(\omega)$ is a semicontraction to go from the second to the third line. This is precisely the same computation as for one single semicontraction in (1-1), with an additional dependency on ω that has to be written correctly.

In the same way that results on subadditive sequences (Lemmas 1.4 and 1.5) were instrumental in the proofs of Sections 1 and 2 on the behavior of one semicontraction, we will be able to analyze the behavior of random semicontractions if we have sufficiently precise tools on subadditive cocycles.

The first central result in this direction is Kingman's Theorem, replacing in this context Fekete's Lemma 1.4.

Theorem 4.2 (Kingman [1968]). *Let u be an integrable subadditive cocycle. There exists $A \in [-\infty, \infty)$ such that, almost surely, $u(n,\omega)/n \to A$. Moreover, if $A > -\infty$, the convergence also holds in L^1. Finally, A is the limit of the sequence $(\int u(n,\omega)\,\mathrm{d}\mathbb{P}(\omega))/n$, which is convergent by subadditivity.*

Since $d(x_0, T^n(\omega)(x_0))$ is a subadditive cocycle when the $T(\omega)$ are semicontractions, this result implies Proposition 3.2, i.e., the almost sure convergence of $d(x_0, T^n(\omega)(x_0))/n$.

There are many proofs of Kingman's Theorem in the literature. The simplest one is probably the proof of Steele [1989], that we will sketch now.

Proof sketch. Consider the measurable function $f(\omega) = \liminf u(n,\omega)/n$. The subadditivity of u implies that $f(\omega) \leqslant f(U\omega)$ almost surely. We deduce thanks to Poincaré recurrence theorem that $f(\omega) = f(U\omega)$ almost everywhere. Indeed, by this theorem, almost every point in $V_a = f^{-1}([-\infty, a])$ comes back infinitely often to V_a under the iteration of U. A point with $f(\omega) < f(U\omega)$ would belong to V_a for each rational a in $(f(\omega), f(U\omega))$ but can only come back to it if it belongs to a 0 measure set.

The function f, which is almost everywhere invariant, is almost everywhere constant by ergodicity, equal to some $A \in [-\infty, +\infty)$. Assume for instance $A > -\infty$, and take $\varepsilon > 0$. Fix also $N > 0$. For almost every ω, there exists an integer $n(\omega)$ with $u(n(\omega),\omega) \leqslant n(\omega)(A+\varepsilon)$, by definition of the inferior limit. Given a point ω, we define a sequence of times as follows. Start from $n_0 = 0$. If $n(U^{n_0}\omega) > N$ (i.e., we have to wait too long to see the almost realization of the liminf), we are not patient enough and we let $n_1 = n_0 + 1$.

Otherwise, set $n_1 = n_0 + n(U^{n_0}\omega)$, so that $u(n_1 - n_0, U^{n_0}\omega) \leqslant (n_1 - n_0)(A+\varepsilon)$. We continue this construction by induction, therefore partitioning the integers into intervals $[n_i, n_{i+1} - 1]$. On most of them, the value of u is bounded by $(n_{i+1} - n_i)\cdot(A+\varepsilon)$ by construction. On the other ones, we do not have a good control, but their frequency is very small if N is large.

Combining these two estimates and using the subadditivity of u to bound $u(n_i, \omega)$ by the sum of the contributions of each individual interval, we obtain $u(n_i,\omega) \leqslant n_i \cdot (A + \varepsilon) + o_N(1) n_i$. This is bounded by $n_i \cdot (A + 2\varepsilon)$ if N is large enough. Finally, we obtain $\limsup u(n,\omega)/n \leqslant A + 2\varepsilon$ (first along the subsequence n_i, then for any integer as two consecutive terms of this sequence are separated at most by N). Letting ε tend to 0, we finally get $\limsup u(n,\omega)/n \leqslant A = \liminf u(n,\omega)/n$. This concludes the proof of the almost sure convergence. □

One can note that this proof looks very closely like the proof of Fekete's Lemma 1.4. The difference is that, instead of using subadditivity always with respect to the same time N (which almost realizes the liminf), one has to use a time which depends on the point we are currently at. Apart from this, the two proofs can be written completely in parallel.

To prove Theorem 3.5, we need a substitute for Lemma 1.5 if we want to use the proof strategy of Section 1. The direct analogue of this lemma in our context would be the following statement:

Let $\varepsilon > 0$. Let u be an integrable subadditive cocycle, such that $u(n,\omega)/n \to A > -\infty$ almost everywhere. For almost every ω, there exist arbitrarily large integers n such that, for all $1 \leqslant \ell \leqslant n$, we have $u(n,\omega) \geqslant u(n-\ell, U^\ell(\omega)) + (A-\varepsilon)\ell$.

However, this statement is wrong. Take for instance $u(n,\omega) = \sum_{k=0}^{n-1} v(U^k\omega)$ for some function v (this is an additive cocycle, whose limit A is equal to $\int v$). If the above statement holds, then taking $\ell = 1$ we get $v(\omega) \geqslant A - \varepsilon$. Letting ε tend to 0, we get $v(\omega) \geqslant A = \int v$ almost everywhere, which is wrong if v is not almost surely constant.

This argument shows that any valid statement has to allow some fluctuations for each ℓ. At the same time, it is crucial for the application to semicontractions to have a statement which controls all intermediate times between 1 and n. The main result of Gouëzel and Karlsson [2015] is the following theorem, compatible with these two constraints.

Theorem 4.3. *Let u be an integrable subadditive cocycle, such that $u(n,\omega)/n \to A > -\infty$ almost everywhere. For almost every ω, there exists a sequence $\delta_\ell \to 0$ and arbitrarily large integers n such that, for all $1 \leqslant \ell \leqslant n$, we have $u(n,\omega) \geqslant u(n-\ell, U^\ell(\omega)) + (A - \delta_\ell)\ell$.*

In a setting of random semicontractions, applying this theorem to the subadditive cocycle $u(n,\omega) = d(x_0, T^n(\omega)(x_0))$ and following the arguments of Section 2, we obtain readily Theorem 3.5.

Note that the subadditivity of u ensures that $u(n,\omega) \leqslant u(n-\ell, U^\ell(\omega)) + u(\ell,\omega)$. As $u(\ell,\omega) \sim A\ell$ by Kingman's theorem, an upper bound $u(n,\omega) \leqslant u(n-\ell, U^\ell(\omega)) + (A+\delta_\ell)\ell$ is automatic. The difficulty in Theorem 4.3 is that, instead, we are looking after a *lower* bound, ensuring that the subadditive cocycle u is in fact almost additive at all intermediate times between 1 and n, for some good times n.

To prove this theorem, a first idea is to try to use the concept of records, at the heart of the proof of Lemma 1.5. It would work very well to prove the existence of infinitely many times n for which $u(n,\omega) \geqslant u(n-\ell,\omega) + (A-\delta_\ell)\ell$ for all intermediate time ℓ. Unfortunately, this is not the statement we are interested in: we do not want a statement involving $u(n-\ell,\omega)$, but rather $u(n-\ell, U^\ell(\omega))$ since this is the quantity that is relevant for the application to semicontractions. We need a different argument.

The proof of Theorem 3.3 by Karlsson and Margulis in Karlsson and Margulis [1999] relied on a statement which is slightly weaker than Theorem 4.3. In the same context, they show that, given $\varepsilon > 0$, there exist almost surely a time $k(\omega)$ and arbitrarily large integers n such that, for all $k(\omega) \leqslant \ell \leqslant n$, we have $u(n,\omega) \geqslant u(n-\ell, U^\ell(\omega)) + (A-\varepsilon)\ell$. This is enough to prove Theorem 3.3 by following the proof in Section 2. At first, one could think that this statement is very close to Theorem 4.3: a strategy to prove this theorem could be to start from the statement of Karlsson and Margulis for $\varepsilon_i = 2^{-i}$, and then apply some kind of diagonal argument to obtain times n that work simultaneously for $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_N$ (with N arbitrarily large). The problem with this approach is that the theorem of Karlsson and Margulis is not quantitative: it does not guarantee that there are many good times (and, in fact, their proof gives a very small set of good times). Typically, there is no integer which is good both for ε_0 and ε_1, ruining the diagonal argument!

If we want to use this kind of approach, we need large sets of good times, when ε is fixed. This is what we will do to prove Theorem 4.3. The notion of largeness we will use is the (lower) asymptotic density

$$\underline{\mathrm{Dens}}\, B = \liminf_{N\to\infty} \frac{\mathrm{Card}(B \cap \{1,\dots,N\})}{N}.$$

The main steps in the proof are the following. Going to the natural extension if necessary, we can assume that U is invertible. Then we define a new subadditive cocycle $\tilde u(n,\omega') = u(n, U^{-n}\omega')$ (it is subadditive for U^{-1}). Its interest is that, writing $\omega' = U^n\omega$, then

$$u(n,\omega) - u(n-\ell, U^\ell\omega) = \tilde u(n,\omega') - \tilde u(n-\ell,\omega').$$

In the right hand side term, the same point ω' appears in both instances of $\tilde u$. This will make it possible to use some combinatorial arguments that do not work directly for u. The price to pay is that good times for $\tilde u$ are not good times for u: there is an additional change of variables, which spoils the argument if the information on the set of good times is only qualitative, but which works if we have quantitative estimates in terms of asymptotic density for the set of good times.

Then, we show that $\tilde{u}$ has many good times, with the following lemmas:

Lemma 4.4. *Let $\delta > 0$. Then there exists $C > 0$ such that, for almost every ω,*

$$\underline{\mathrm{Dens}}\{n \in \mathbb{N} \ :\ \forall \ell \in [1,n], \tilde{u}(n,\omega) - \tilde{u}(n-\ell,\omega) \geqslant (A-C)\ell\} \geqslant 1-\delta.$$

Lemma 4.5. *Let $\delta > 0$ and $\varepsilon > 0$. Then there exists an integer k such that, for almost every ω,*

$$\underline{\mathrm{Dens}}\{n \in \mathbb{N} \ :\ \forall \ell \in [k,n], \tilde{u}(n,\omega) - \tilde{u}(n-\ell,\omega) \geqslant (A-\varepsilon)\ell\} \geqslant 1-\delta.$$

The second lemma is essentially a more precise variant of the first one. Their proofs are essentially combinatorial, and borrow some ideas to the proof by Steele of Kingman's Theorem that we have described above.

As the intersection of two sets with asymptotic density close to 1 still has an asymptotic density close to 1, we will then be able to intersect the sets of good times produced by these lemmas (and, in the case of Lemma 4.5, for different values of ε), while keeping sets with large density. This makes it possible to implement the diagonal argument alluded to earlier. After a final change of variables to go back to u, we finally obtain Theorem 4.3. The details are rather delicate and technical, the interested reader is referred to Gouëzel and Karlsson [2015] for a full proof and to Gouëzel, Sébastien [2016] for a computer-checked formalization of the proof.

References

Martin R. Bridson and André Haefliger (1999). *Metric spaces of non-positive curvature*. Vol. 319. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer-Verlag, pp. xxii+643. MR: MR1744486. mathscinet: 1744486 (cit. on pp. 1957, 1958).

Aleksander Całka (1984). "On conditions under which isometries have bounded orbits". *Colloq. Math.* 48.2, pp. 219–227. MR: 758530 (cit. on p. 1957).

Arnaud Denjoy (1926). "Sur l'itération des fonctions analytiques." *C. R. Acad. Sci., Paris* 182, pp. 255–257. Zbl: 52.0309.04 (cit. on p. 1955).

Gouëzel, Sébastien (2016). "Ergodic Theory". *Archive of Formal Proofs*. Formal proof development (cit. on pp. 1951, 1965).

– (2017). "Cocycles sous-additifs et horofonctions". *Séminaires et congrès* 31, pp. 19–38 (cit. on p. 1951).

Sébastien Gouëzel and Anders Karlsson (2015). "Subadditive and multiplicative ergodic theorems". preprint (cit. on pp. 1951, 1960, 1963, 1965).

Anders Karlsson (2001). "Non-expanding maps and Busemann functions". *Ergodic Theory Dynam. Systems* 21, pp. 1447–1457. MR: MR1855841. mathscinet: 1855841 (cit. on pp. 1952, 1956).

Anders Karlsson and François Ledrappier (2006). "On laws of large numbers for random walks". *Ann. Probab.* 34.5, pp. 1693–1706. MR: 2271477 (cit. on p. 1960).

Anders Karlsson and Gregory A. Margulis (1999). "A multiplicative ergodic theorem and nonpositively curved spaces". *Comm. Math. Phys.* 208.1, pp. 107–123. MR: 1729880 (cit. on pp. 1960, 1964).

John F. C. Kingman (1968). "The ergodic theory of subadditive stochastic processes". *J. Roy. Statist. Soc. Ser. B* 30, pp. 499–510. MR: 0254907 (cit. on p. 1962).

Elon Kohlberg and Abraham Neyman (1981). "Asymptotic behavior of nonexpansive mappings in normed linear spaces". *Israel J. Math.* 38, pp. 269–275. MR: MR617673. mathscinet: 617673 (cit. on pp. 1952, 1954).

J. Michael Steele (1989). "Kingman's subadditive ergodic theorem". *Ann. Inst. H. Poincaré Probab. Statist.* 25, pp. 93–98. MR: MR995293. mathscinet: 995293 (cit. on p. 1962).

Julius Wolff (1926). "Sur l'itération des fonctions bornées." *C. R. Acad. Sci., Paris* 182, pp. 200–201. Zbl: 52.0309.03 (cit. on p. 1955).

Received 2017-11-29.

SÉBASTIEN GOUËZEL
LABORATOIRE JEAN LERAY, CNRS UMR 6629
UNIVERSITÉ DE NANTES
2 RUE DE LA HOUSSINIÈRE
44322 NANTES
FRANCE
sebastien.gouezel@univ-nantes.fr

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1967–1990)

DIMENSION THEORY OF SELF-SIMILAR SETS AND MEASURES

Michael Hochman

Abstract

We report on recent results about the dimension and smoothness properties of self-similar sets and measures. Closely related to these are results on the linear projections of such sets, and dually, their intersections with affine subspaces. We also discuss recent progress on the the Bernoulli convolutions problem.

1 Introduction

Consider a random walk $(X_n)_{n=0}^{\infty}$ on $\mathbb{R}^d$, started from a point, and with the transitions given by $X_{n+1} = \xi_n X_n$, where (ξ_n) is an independent sequence of similarity maps, chosen according to a common fixed distribution p.

The long term behavior of (X_n) depends on the scaling properties of the ξ_n. If they are expanding, there is no interesting limit. But the other cases are quite interesting. When the ξ_n are isometries, and act in some sense irreducibly on $\mathbb{R}^d$, both a central limit theorem and local limit theorem hold, i.e. $(X_n - \mathbb{E}X_n)/n^{d/2}$ converges in law to a Gaussian, and $X_n - \mathbb{E}X_n$ converges to Lebesgue measure on bounded open sets Tutubalin [1967], Gorostiza [1973], Roynette [1974], and P. Varjú [2016]. Thus, the limiting behavior is universal, and X_n spreads out "as much as possible".

The remaining case, namely, when the ξ_n contract, is our focus here. Then X_n converges in law (without any normalization) to a measure μ, which does not depend on the starting point X_0, but very strongly depends on the step distribution p. An important case is when p is finitely supported, in which case μ is called a *self-similar measure*, and its topological support, which is the set of accumulation points of any orbit of the semigroup generated by supp p, is called a *self-similar set*.

Research supported by NSF grant 0901534, ISF grant 1409/11, and ERC grant 306494. Parts of the work was carried out during a visit to Microsoft Research, Redmond, WA.

MSC2010: primary 28A80; secondary 37C45, 37F35, 11P70.

Many mathematical problems surround self-similar sets and measures, and in this paper we survey some of the recent progress on them. Perhaps the most natural problem is to determine the dimension, and, if applicable, smoothness, of μ. Although there is no universal limiting distribution as in the CLT, a weaker universal principle is believed to apply: namely, that μ should be "as spread out as possible" given the constraints imposed by the amount of contraction in the system, and given possible algebraic constraints, such as being trapped in a lower-dimensional subspace. This principle – that in algebraic settings, dynamical processes tend to spread out as much as the algebraic constraints allow – has many counterparts, such as rigidity theorems for unipotent flows and higher rank diagonal actions on homogeneous spaces (e.g. Ratner [1991b,a], Margulis and Tomanov [1994], and Ratner [1995], see also Lindenstrauss [2010])), and stiffness of random walks on homogeneous spaces (e.g. Bourgain, Furman, Lindenstrauss, and Mozes [2007] and Benoist and Quint [2009]).

Self-similar sets and measures are also natural examples of "fractal" sets; they possesses a rich set of symmetries, and a natural hierarchical structure. This has motivated a number of longstanding conjectures about the geometry of such sets and measures, and specifically, about the dimension of their linear images and their intersections with affine subspaces. Many of these conjectures are now confirmed, as we shall describe below.

Finally, we devote some time to the special case of Bernoulli convolutions, which is a problem with strong number-theoretic connections. This problem also has seen dramatic progress in the past few years.

Due to space constraints, we have omitted many topics, and kept to a minimum the discussion of classical results, except where directly relevant. A more complete picture can be found in the references.

The plan of the paper is as follows. We discuss the dimension problem for self-similar sets and measures in Section 2;. The Bernoulli convolutions problem in Section 3; and projections and slices in Section 4.

2 Self-similar sets and measures

Self-similar sets and measures are the prototypical fractals; the simplest example is the middle-$1/3$ Cantor set and the Cantor-Lebesgue measure, which arise from the system of contractions $\varphi_0(x) = \frac{1}{3}x$ and $\varphi_1(x) = \frac{1}{3}x + \frac{2}{3}$, taken with equal probabilities. In general, self-similar sets and measures are made up of copies of themselves, just as the Cantor set does. This is most evident using Hutchinson's construction Hutchinson [1981], which we specialize to our setting.

An *iterated function system* will mean a finite family $\Phi = \{\varphi_i\}_{i \in \Lambda}$ of contracting similarity maps of $\mathbb{R}^d$. A *self similar set* is the *attractor* of Φ, that is, the unique non-empty compact set $X = X_\Phi$ satisfying

$$X = \bigcup_{i\in\Lambda} \varphi_i(X) \tag{1}$$

The *self-similar measure* determined by $\Phi = \{\varphi_i\}_{i\in\Lambda}$ and a probability vector $p = (p_i)_{i\in\Lambda}$ (which we think of as a measure on Φ) is the unique Borel probability measure $\mu = \mu_{\Phi,p}$ satisfying

$$\mu = \sum_{i\in\Lambda} p_i \cdot \varphi_i \mu \tag{2}$$

where $\varphi\mu = \mu \circ \varphi^{-1}$ is the push-forward measure. When p is strictly positive, the topological support $\operatorname{supp}\mu$ of μ is X.

A similarity has the form $\varphi(x) = rUx + b$ where $r > 0$, U is an orthogonal matrix, and $b \in \mathbb{R}^d$. We call r the *contraction*, U the *orthogonal part*, and b the *translation part* of φ , respectively. We say that Φ, X or μ are *self-homothetic* if the φ_i are homotheties, i.e. the U_i are trivial; *uniformly contracting* if φ_i all have the same contraction ratio; and have *uniform rotations* if the φ_i all have the same orthogonal parts. Also, Φ is algebraic if all coefficients defining φ_i are algebraic.

By definition, X and μ are made up of smaller copies of themselves, and by iterating the identities (1) and (2) one gets such a representation at arbitrarily small scales. For a sequence $\mathbf{i} = i_1 \dots i_n \in \Lambda^n$ it is convenient to denote $\varphi_{\mathbf{i}} = \varphi_{i_1} \circ \dots \circ \varphi_{i_n}$ and $p_{\mathbf{i}} = p_{i_1} \cdot p_{i_2} \cdot \dots \cdot p_{i_n}$. Note that the contraction tends to 0 exponentially as $n \to \infty$. For $\mathbf{i} \in \Lambda^n$, we call $\varphi_{\mathbf{i}} X$ and $\varphi_{\mathbf{i}}\mu$ *generation-n cylinders*. These are the small-scale copies alluded to above, with the corresponding representation $X = \bigcup_{\mathbf{i}\in\Lambda^n} \varphi_{\mathbf{i}} X$ and $\mu = \sum_{\mathbf{i}\in\Lambda^n} p_{\mathbf{i}} \cdot \varphi_{\mathbf{i}}\mu$. The last identity shows that the definition above coincides with the earlier description using random walks: For $\mathbf{i} \in \Lambda^n$, the diameter of the support of $\varphi_{\mathbf{i}}\mu$ converges uniformly to zero as $n \to \infty$, so the identity $\mu = \sum_{\mathbf{i}\in\Lambda^n} p_{\mathbf{i}} \cdot \varphi_{\mathbf{i}}\mu$ implies that μ is the limit distribution of $\sum_{\mathbf{i}\in\Lambda^n} p_{\mathbf{i}} \cdot \delta_{\varphi_{\mathbf{i}} x}$ for every $x \in \operatorname{supp}\mu$. The last measure is just the distribution of the random walk from the introduction, started from x.

2.1 Preliminaries on dimension. We denote by $\dim(\cdot)$ the Hausdorff dimension for sets, and the lower Hausdorff dimension of Borel probability measures, defined by

$$\dim \mu = \inf\{\dim E \ : \ \mu(E) > 0\}$$

$\mathcal{H}^s$ denotes s-dimensional Hausdorff measure. Absolute continuity (a.c.) is with respect to Lebesgue measure.

There are many other notions of dimension which in general disagree, but for self-similar sets and measures most of them coincide. Specifically, Falconer [1989] proved that self-similar sets have equal Hausdorff and box dimensions. For self-similar measures, Feng and Hu [2009] proved that a self-similar measure is *exact dimensional*, meaning that for $\alpha = \dim \mu$, as $r \to 0$ we have

$$\mu(B_r(x)) = r^{\alpha(1+o(1))} \qquad \text{for } \mu\text{-a.e. } x \tag{3}$$

(non-uniformly in x).

With regard to smoothness of self-similar measures, in the special case of infinite convolutions it is a classical result of Jessen and Wintner [1935] that the measure is of pure type, i.e. is either singular with respect to Lebesgue, or absolutely continuous with respect to it. This is true for all self-similar measures, and is a consequence of Kolmogorov's zero-one law.

2.2 Similarity and Lyapunov dimension. In order to estimate the dimension of a set or measure, one must construct efficient covers of it, or estimate the mass of sets of small diameter. Self-similar sets and measures come equipped with the natural covers given by the cylinder sets of a given generation, or of a given approximate diameter. Counting cylinder sets, one arrives at the following estimates for the dimension:

- The *similarity dimension* is the unique $s = s(\Phi) \geq 0$ satisfying $\sum r_i^s = 1$, where r_i is the contraction constant of φ_i.
- The *Lyapunov dimension* of Φ and a probability vector $p = (p_i)_{i\in\Lambda}$ is $s(\Phi, p) = H(p)/\lambda(p)$, where $H(p) = -\sum p_i \log p_i$ is the Shannon entropy of p, and $\lambda(p) = -\sum p_i \log r_i$ is the asymptotic contraction, i.e. Lyapunov exponent, of the associate random product.

Note that $s(\Phi, p)$ is maximal when $p = (r_i^s)_{i\in\Lambda}$ (with $s = s(\Phi)$), and then the similarity and Lyapunov dimensions coincide: $s(\Phi, p) = s(\Phi)$.

These estimates ignore the possibility of coincidences between cylinders. When cylinders of the same generation intersect we say that the system has *overlaps*; it has *exact overlaps* if there exist finite sequences $\mathbf{i}, \mathbf{j} \in \Lambda^*$ such that $\varphi_{\mathbf{i}} = \varphi_{\mathbf{j}}$, or in other words, if the semigroup generated by Φ is not freely generated by it. If this happens, then without loss of generality we can assume that $\mathbf{i}, \mathbf{j}$ have the same length n (otherwise replace them with $\mathbf{ij}$ and $\mathbf{ji}$), and if such pairs exist for some n, then they exist for all large enough n.

In this generality, Hutchinson [1981] was the first to show that $\dim X \leq s(\Phi)$ and $\dim \mu \leq s(\Phi, p)$. Furthermore, these are equalities if we assume that there are only mild overlaps. Specifically, Φ is said to satisfy the *open set condition* (OSC) if there exists an open non-empty set U such that $\varphi_i U \subseteq U$ for all $i \in \Lambda$ and $\varphi_i U \cap \varphi_j U = \varnothing$ for all $i \neq j$. A special case of this is when the first generation cylinders are disjoint, which is called the *strong separation condition* (SSC). The OSC allows overlaps, but it implies that the overlaps have bounded multiplicity.

Theorem 2.1 (Hutchinson [ibid.]). *Suppose* $\Phi = \{\varphi_i\}_{i \in \Lambda}$ *is an IFS in* $\mathbb{R}^d$ *satisfying the OSC. Then* $\dim X_\Phi = s(\Phi)$ *and* $\dim \mu_{\Phi,p} = s(\Phi, p)$ *for every* p*. Furthermore, writing* $s = s(\Phi)$*, we have* $0 < \mathcal{H}^s(X) < \infty$*, and* $\mathcal{H}^s|_{X_\Phi}$ *is equivalent to the self-similar measure defined by* $p = (r_i^s)_{i \in \Lambda}$*.*

In fact, for $s = s(\Phi)$, Falconer [1989] showed that $\mathcal{H}^s(X) < \infty$ always holds. In general $\mathcal{H}^s$ can vanish on X; Schief [1994] (following some special cases Kenyon [1997] and Bandt and Graf [1992]) showed that $\mathcal{H}^s(X) > 0$ for $s = s(\Phi)$ is exactly equivalent to the OSC.

It must be emphasized that the OSC allows only "minor" overlaps between cylinders, and since $\dim X \leq d$ for $X \subseteq \mathbb{R}^d$, by Theorem 2.1, the OSC implies $s(\Phi) \leq d$. There exist IFSs with $s(\Phi) > d$, e.g. one can take an IFS $\Phi_n = \{\varphi_i\}_{i=1}^n$ on $\mathbb{R}$ with $\varphi_i(x) = \frac{1}{2}x + i$. The attractor is an interval, but $s(\Phi_n) = \log n / \log 2$. In any case, the dimension d of the ambient space $\mathbb{R}^d$ is also an upper bound on dimension, so whether or not the OSC holds, we have

$$\dim X \leq \min\{d, s(\Phi)\} \tag{4}$$

$$\dim \mu \leq \min\{d, s_p(\Phi)\} \tag{5}$$

We say that X or μ exhibits *dimension drop* if the corresponding inequality above is strict. The principle way dimension drop occurs is if there are exact overlaps. Indeed, suppose $\dim X < d$ and $\mathbf{i}, \mathbf{j} \in \Lambda^n$ are distinct with $\varphi_{\mathbf{i}} = \varphi_{\mathbf{j}}$. Let $\Phi^n = \{\varphi : \mathbf{u} \in \Lambda^n\}$. Then a short calculation shows that $s(\Phi^n) < s(\Phi)$. Since X is also the attractor of Φ^n, we get $\dim X \leq \min\{d, s(\Phi^n)\} < \min\{d, s(\Phi)\}$, and we have dimension drop.

2.3 The overlaps conjecture, and what we know about it.

In this section we specialize to $\mathbb{R}$, where exact overlaps are the only known mechanism that leads to dimension drop. The next conjecture is partly folklore. It seems to have first appeared in general form in Simon [1996].

Conjecture 2.2. *In* $\mathbb{R}$*, dimension drop occurs only in the presence of exact overlaps.*

Thus, non-exact overlaps should not lead to dimension drop. By Theorem 2.1, we know that minor overlaps can indeed be tolerated. Other examples come from parametric families such as the $\{0, 1, 3\}$-problem, which concerns the attractor of the IFS $\Phi_\lambda = \{x \mapsto \lambda x\,,\ x \mapsto \lambda x + 1\,,\ x \mapsto \lambda x + 3\}$. There are only countably many parameters λ with exact overlaps, and Pollicott and Simon showed that for a.e. $\lambda \in [\frac{1}{3}, \frac{1}{2}]$, there is no dimension drop, see also Keane, Smorodinsky, and Solomyak [1995].

To go further we must quantify the amount of overlap. Define the distance $d(\cdot,\cdot)$ between similarities $\varphi(x) = ax + b$ and $\varphi'(x) = a'x + b'$ by

$$d(\varphi,\varphi') = |b - b'| + |\log a - \log a'| \tag{6}$$

Alternatively, one can take any left- or right-invariant Riemannian metric on the group of similarities, or the operator norm on the standard embedding of the group into $GL_2(\mathbb{R})$. These metrics are not equivalent, but are mutually bounded up to a power distortion, which makes them equivalent for the purpose of what follows.

Given an IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ of similarities, let

$$\Delta_n = \min\{d(\varphi_{\mathbf{i}},\varphi_{\mathbf{j}}) \,:\, \mathbf{i},\mathbf{j} \in \Lambda^n \,,\; \mathbf{i} \neq \mathbf{j}\}$$

There are exact overlaps if and only if $\Delta_n = 0$ for some n (and hence all large enough n), and contraction implies that $0 \le \Delta_n \le r^n$ for some $0 < r < 1$. However, the decay of Δ_n generally need not be faster than exponential. We say that Φ is *exponentially separated* if there is a constant $c > 0$ such that $\Delta_n \ge c^n$ for all n.

Theorem 2.3 (Hochman [2014]). *Let Φ be an IFS in $\mathbb{R}$, let $\mu = \mu_{\Phi,p}$ be the self-similar measure, and write $s = s(\Phi, p)$. Then either* $\dim \mu = \min\{1, s\}$*, or else $\Delta_n \to 0$ super-exponentially. The same statement holds for sets.*[1]

Thus, exponential separation implies no dimension drop. We do not know of any IFS without exact overlaps for which $\frac{1}{n}\log\Delta_n \to \infty$, and it is conceivable that they simply do not exist, which would prove the conjecture.

Corollary 2.4. *Within the class of algebraic IFSs on $\mathbb{R}$, Conjecture 2.2 is true.*

Indeed, one can choose the metric so that $d(\varphi,\psi)$ is a polynomial in the coefficients of φ,ψ, and then $d(\varphi_{i_1}\dots\varphi_{i_n},\varphi_{j_1}\dots\psi_{j_n})$ is a polynomial of degree $O(n)$ in the coefficients of the φ_i. If these are algebraic, such an expression either vanishes, or is bounded below by an exponential c^n for some $c > 0$ (see Garsia's Lemma 3.4 below).

When exact overlaps exist, one can get a better bound than the Lyapunov dimension by taking the number of exact overlaps into account. Given Φ and p, let $(\xi_n)_{n=1}^{\infty}$ be i.i.d. elements of Φ with distribution p, and let $\sigma_n = \xi_n\xi_{n-1}\dots\xi_1$ be the associated random walk on the similarity group. The *random walk entropy* of p is defined by

$$h_{RW}(p) = \lim_{n\to\infty}\frac{1}{n}H(\sigma_n) \tag{7}$$

[1] The statement for sets follows from the measure case applied to $\mu_{\Phi,p}$, with p chosen so that $s(\Phi,p) = s(\Phi)$. The same remark holds for many theorems in the sequel.

where $H(\sigma_n)$ is the Shannon entropy of the discrete random variable σ_n. The limit exists by sub-additivity, and if Φ^* is freely generated by Φ, then $h_{RW}(p) = H(p)$. Corresponding to (5) we have the bound

$$\dim \mu \leq \min\{1, \frac{h_{RW}(p)}{\lambda(p)}\}$$

The following is a reasonable extension of Conjecture 2.2:

Conjecture 2.5. *If $\mu = \mu_{\Phi,p}$ is a self-similar measure on $\mathbb{R}$ then*

$$\dim \mu = \min\{1, h_{RW}(p)/\lambda(p)\}.$$

The next theorem is proved by the same argument as Theorem 2.3. The statement first appeared in P. Varjú [2016].

Theorem 2.6. *Let $\Phi = \{\varphi_i\}$ be an IFS of similarities in $\mathbb{R}$, and suppose that there is a $c > 0$ such that for every $\mathbf{i}, \mathbf{j} \in \Lambda^n$ either $\varphi_{\mathbf{i}} = \varphi_{\mathbf{j}}$ or $d(\varphi_{\mathbf{i}}, \varphi_{\mathbf{j}}) \geq c^n$. Let $\mu = \mu_{\Phi,p}$ be the self-similar measure for Φ. Then $\dim \mu = \min\{1, h_{RW}(p)/\lambda(p)\}$.*

An important strengthening of these results is obtained by replacing Hausdorff dimension with L^q-dimension. To define it, let $\mathfrak{D}_n$ denote the dyadic partition of $\mathbb{R}$ into intervals $[k/2^n, (k+1)/2^n)$, $k \in \mathbb{Z}$, and for $q > 1$ set

$$D(\mu, q) = \lim_{n\to\infty} -\frac{\log \sum_{I\in\mathfrak{D}_n} \mu(I)^q}{(q-1)n}$$

The limit is known to exist for self-similar measures Peres and Solomyak [2000], and for such measures, $\dim \mu = \lim_{q\searrow 1} D_\mu(q)$. The function $q \mapsto D(\mu, q)$ is non-increasing in q, and has the following property: for every $\alpha < D(\mu, q)$, there is a constant C such that $\mu(B_r(x)) \leq C \cdot r^{(1-1/q)\alpha}$, for *every* $x \in \mathbb{R}$. This is in stronger than (3), which holds only for μ-a.e. x, and non-uniformly.

The L^q-analog of the Lyapunov dimension for a self-similar measure $\mu = \mu_{\Phi,p}$, is the solution $s = s^q(\Phi, p)$ of the equation $\sum p_i^q |r_i|^{(q-1)s} = 1$, where r_i are the contraction ratios of the maps in Φ. Note that if $p = (r_i^{s(\Phi)})_{i\in\Lambda}$, then $s^q(\Phi, p) = s(\Phi, p) = s(\Phi)$, is independently of q. We always have $D_\mu(q) \leq s^q(\Phi, p)$.

Theorem 2.7 (Shmerkin [2016])**.** *Let $\Phi = \{\varphi_i\}$ be an IFS on $\mathbb{R}$ and $p = (r_i^{s(\Phi)})$. Let $\mu = \mu_{\Phi,p}$ and $s = s(\Phi, p) = s(\Phi)$. Then either $D_\mu(q) = s$ for all $q > 1$, or else $\Delta_n \to 0$ super-exponentially.*

In particular, if Φ is exponentially separated, then for every $t < \min\{1, s(\Phi)\}$, there is a $C = C(t) > 0$ such that $\mu(B_r(x)) \leq Cr^t$ for all $x \in X$ and all $r > 0$.

2.4 Some ideas from the proofs. The main idea is, very roughly, as follows. A self-similar measure μ on $\mathbb{R}$ can be written, locally, as a convolution of a scaled copy of itself with another measure ν whose "dimension" (in some finitary sense) is proportional to the difference $s(\Phi, p)-\dim\mu$. But convolution is a smoothing operation, and $\mu * \nu$ has larger dimension than μ if the dimension of ν is positive. Hence, if there were dimension drop, at small scales μ would be smoother than itself, which is impossible.

In order to give even a slightly more comprehensive sketch, some preparation is needed. First, by "smoothing", we mean that convolving measures generally results in more "spread out" measures than we started with. The discussion below is very much in the spirit of additive combinatorics, in which one asks when the sum $A+B$ of two finite sets $A, B \subseteq \mathbb{Z}$ is substantially larger than A. "Larger" is often interpreted as $|A+B| > C|A|$ where $C > 0$ is fixed and the sets are large, as in Freiman's theorem (e.g. Tao and Vu [2006]); but in our setting we mean $|A+B| \geq |A|^{1+\delta}$. Such a growth condition is closely related to the work of several authors on the sum-product phenomenon, notably Bourgain [2003] and Bourgain [2010]. The version used in Theorem 2.10 and presented below is from Hochman [2014]. See also the remark at the end of the section.

We measure how "spread out" a measure is using entropy at a finite scale. Recall that $\mathfrak{D}_n$ is the level-n dyadic partition of $\mathbb{R}$, whose atoms are the intervals $[k/2^n, (2+1)/2^n)$. The *scale-n entropy* of a probability measure ν is the Shannon entropy $H(\nu, \mathfrak{D}_n) = -\sum_{I\in\mathfrak{D}_n} \nu(I)\log_2 \nu(I)$ of ν with respect to $\mathfrak{D}_n$. We refer to Cover and Thomas [2006] for a more thorough introduction to entropy.

Scale-n entropy is a discretized substitute for dimension, and as a first approximation, one can think of it as the logarithm of the number of atoms of $\mathfrak{D}_n$ with non-trivial ν mass. In particular if ν is exact dimensional, then

$$\frac{1}{n}H(\nu, \mathfrak{D}_n) \to \dim\nu \tag{8}$$

For $m > n$ let $H(\nu, \mathfrak{D}_m|\mathfrak{D}_n) = H(\nu, \mathfrak{D}_m) - H(\nu, \mathfrak{D}_n)$ be the conditional entropy, i.e. the entropy increase from scale 2^{-n} to 2^{-m}. Assuming (8), we have

$$\frac{1}{m-n}H(\nu, \mathfrak{D}_m|\mathfrak{D}_n) = \dim\nu + o(1) \qquad \text{as } n\to\infty \text{ and } m-n\to\infty$$

In general, scale-n does not decrease under convolution,[2] and for generic measures it increases as much as possible, i.e. $H(\theta * \nu, \mathfrak{D}_n) \approx H(\nu, \mathfrak{D}_n) + H(\theta, \mathfrak{D}_n)$ assuming that the right hand side does not exceed one. There certainly are exceptions, even cases in which $H(\theta * \nu) \approx H(\nu)$. But for self-similar measures, some substantial entropy increase must occur.

[2] Actually it could decrease but only by an additive constant, which is negligible with our $1/n$-normalization.

Theorem 2.8 (Hochman [2014]). *For every $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. Let μ be a self-similar measure on $\mathbb{R}$ with $\dim \mu < 1 - \varepsilon$, and let θ be a probability measure. Then for n large enough (depending on ε and μ, but not θ),*

$$H(\theta, \mathfrak{D}_n) > \varepsilon n \qquad \Longrightarrow \qquad H(\theta * \mu, \mathfrak{D}_n) > H(\mu, \mathfrak{D}_n) + \delta n$$

This is a consequence of a more general result[3] describing the structure of pairs of measures θ, ν for which $H(\theta * \nu, \mathfrak{D}_n) \approx H(\nu, \mathfrak{D}_n)$. It says, roughly, that in this case each scale 2^{-i}, $1 \le i \le n$, is of one of two types: either ν looks approximately uniform (like Lebesgue measure) on 2^{-i}-balls centered at ν-typical points; or θ looks approximately like an atomic measure on 2^{-i}-balls centered at θ-typical points. The theorem above follows because self similar measures of dimension < 1 are highly homogeneous and don't look uniform on essentially any ball, while if $H(\theta, \mathfrak{D}_n) \ge \varepsilon n$ then there is a positive propostion of balls on which θ does not look atomic. For the full statement see Hochman [ibid.].

We return to the proof of Theorem 2.3; we assume for contradiction that there is both exponential separation and dimension drop. For simplicity, assume that all the maps in $\Phi = \{\varphi_i\}$ have the same contraction ratio r, and given n write

$$n' = \lfloor n \log(1/r) \rfloor$$

so that $\mathfrak{D}_{n'}$ contains atoms of diameter roughly r^n. For $\mathbf{i} \in \Lambda^n$ the map $\varphi_{\mathbf{i}}$ contracts by r^n, and all the generation-n cylinders appearing in the representation $\mu = \sum_{\mathbf{i} \in \Lambda^n} p_{\mathbf{i}} \cdot \varphi_{\mathbf{i}} \mu$ are translates of each other, so this identity can be re-written as a convolution

$$\mu = \mu^{(n)} * S_{r^n} \mu \tag{9}$$

where $\mu^{(n)} = \sum_{i \in \Lambda^n} p_i \cdot \delta_{\varphi_i(0)}$, and $S_t x = tx$ is the scaling operator. Because $S_{r^n} \mu$ is supported on a set of diameter $O(r^n) = O(2^{-n'})$, it contributes to scale-n' entropy only $O(1)$, so

$$H(\mu^{(n)}, \mathfrak{D}_{n'}) = H(\mu, \mathfrak{D}_{n'}) + O(1) = n' \dim \mu + o(n) \tag{10}$$

Next, chop the measure $\mu^{(n)}$ into a convex combination

$$\mu^{(n)} = \sum_{I \in \mathfrak{D}_{n'}} w_I \cdot (\mu_I^{(n)})$$

where $\mu_I^{(n)}$ is $\mu^{(n)}$ conditioned on I. Inserting this in (9) we get

$$\mu = \sum_{I \in \mathfrak{D}_{n'}} w_I \cdot \mu_I^{(n)} * S_{r^n} \mu \tag{11}$$

[3] For related work on entropy of convolutions, assuming less growth, see Tao Hochman [2014] and Madiman [2008]. More closely related is Bourgain's work on sumsets, Bourgain [2003] and Bourgain [2010].

Since Φ is exponentially separated there is a constant a such that every pair of atoms of $\mu^{(n)}$ is r^{an}-separated, and so lie in different atoms of $\mathfrak{D}_{an'}$. A direct calculation shows that

$$H(\mu^{(n)}, \mathfrak{D}_{an'}) = -\sum_{\mathbf{i}\in\Lambda^n} p_{\mathbf{i}} \log p_{\mathbf{i}} = nH(p) = n' \dim \mu + \varepsilon n$$

where $\varepsilon = H(p) - \dim \mu \cdot \log(1/r) > 0$ because of dimension drop. Combined with (10) this gives

$$H(\mu^{(n)}, \mathfrak{D}_{an'} | \mathfrak{D}_{n'}) = \varepsilon n + o(n) \tag{12}$$

By classical identities, this entropy is the average of the entropies of $\mu_I^{(n)}$, so for a μ-large proportion of $I \in \mathfrak{D}_{n'}$,

$$H(\mu_I^{(n)}, \mathfrak{D}_{an'} | \mathfrak{D}_{n'}) \geq \frac{\varepsilon}{2} n + o(n)$$

Thus,

$$\begin{aligned}
(a-1)n' \cdot \dim \mu &= H(\mu, \mathfrak{D}_{an'} | \mathfrak{D}_{n'}) \\
&\geq \sum_{I \in \mathfrak{D}_{n'}} w_I \cdot H(\mu_I^{(n)} * S_{r^n}\mu, \mathfrak{D}_{an'} | \mathfrak{D}_{n'}) \\
&= \sum_{I \in \mathfrak{D}_{n'}} w_I \cdot H(\mu_I^{(n)} * S_{r^n}\mu, \mathfrak{D}_{an'}) - o(n) \\
&\geq (a-1)n' \cdot \dim \mu + \delta n - o(n)
\end{aligned}$$

where in the first inequality we plugged in the identity (11) and used concavity of the entropy function, in the next line we eliminated the conditioning because $\mu_I^{(n)} * S_{r^n}\mu$ is supported on $O(1)$ atoms of $\mathfrak{D}_{n'}$, and in the last line, we applied Theorem 2.8 together with (12). This is the desired contradiction.

Theorem 2.7 is proved using a very similar philosophy, but with the L^q-dimension of finite-scale approximations replacing entropy as the measure of smoothness. We refer the reader to the original paper for details.

2.5 Higher dimensions. In higher dimensions, Conjecture 2.2 is false in its stated form. To see this start with two IFSs, Φ_1 and Φ_2, on $\mathbb{R}$ with the same contraction ratios. Assume that $s(\Phi_1) > 1$ and the attractor X_1 of Φ_1 is an interval, and the attractor X_2 of Φ_2 satisfies $\dim X_2 = s(\Phi_2)$. We can also assume neither Φ_1 nor Φ_2 have exact overlaps. Let $\Phi = \Phi_1 \times \Phi_2$ be the IFS consisting of all maps of the form $x \mapsto (\varphi(x), \psi(x))$ with $\varphi \in \Phi_1$ and $\psi \in \Phi_2$. Then Φ has attractor $X_1 \times X_2$ of dimension $\dim X_1 + \dim X_2 < s(\Phi_1) + s(\Phi_2) = s(\Phi)$, and Φ has no exact overlaps.

In the example, there are horizontal lines intersecting X in an interval (a copy of X_1) and the family of such lines is preserved by Φ. It seems likely that this is the only new phenomenon possible in higher dimensions. More precisely, let us say that a set X has *full slices* on a linear subspace $V \leq \mathbb{R}^d$, if $\dim X \cap (V + a) = \dim V$ for some $a \in \mathbb{R}^d$. Similarly we say that a measure μ has full slices on V if the system $\{\mu_x^V\}_{x \in \mathbb{R}^d}$ of conditional measure on parallel translates of V satisfies $\dim \mu_x^V = \dim V$ for μ-a.e. x. Finally, we say that V is *linearly invariant under* $\Phi = \{\varphi_i\}$ if $U_i V = V$ for all i, where U_i is the linear part of φ_i. We say that V is non-trivial if $0 < \dim V < d$.

Conjecture 2.9. *Let $X = X_\Phi$ and $\mu = \mu_{\Phi,p}$ be a self-similar measure in $\mathbb{R}^d$. Then dimension drop for X implies that either there are exact overlaps, or there is a non-trivial linearly invariant subspace $V \leq \mathbb{R}^d$ on which X has full slices, and the analogous statement holds for μ.*

Define a metric d on the similarity group Aff($\mathbb{R}^d$) of $\mathbb{R}^d$ using any of the metrics described after Equation (6).

Theorem 2.10 (Hochman [2015]). *Let $\mu = \mu_{\Phi,p}$ be a self-similar measure in $\mathbb{R}^d$ with Lyapunov dimension s. Then at least one of the following holds:*

- $\dim \mu = \min\{d, s(\Phi)\}$.
- $\Delta_n \to 0$ *super-exponentially.*
- *There is a non-trivial linearly invariant subspace $V \leq \mathbb{R}^d$ on which μ has full slices.*

In particular, if the linear parts of the maps φ_i act irreducibly on $\mathbb{R}^d$ then dimension drop implies $\Delta_n \to 0$ super-exponentially. If additionally $\{U_i\}$ generate a free group and have algebraic entries, then there is no dimension drop.

The same holds for the attractor.

The analogous statements for L^q dimension are at present not established, but we anticipate that some version of them holds, at least in the case where the linear parts of the contractions are homotheties.

In dimension $d \geq 3$ the orthogonal group is non-abelian, and the random walk associated to the matrices A_i may have a spectral gap. Then a much stronger conclusion is possible:

Theorem 2.11 (Lindenstrauss and Varjú [2016]). *Let $U_1, \ldots, U_k \in SO(d)$ and $p = (p_1, \ldots, p_k)$ a probability vector. Suppose that the operator $f \mapsto \sum_{i=1}^k p_i f \circ U_i$ on $L^2(SO(d))$ has a spectral gap. Then there is a number $\widetilde{r} < 1$ such that for every choice*

$\widetilde{r} < r_1, \ldots, r_k < 1$, and for every $a_1, \ldots, a_k \in \mathbb{R}^d$, the self similar measure with weights p for the IFS $\{r_i U_i + a_i\}_{i=1}^k$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$.

Contrasting this statement with the previous corollary, in the former we get $\dim \mu = d$ as soon as $s(\Phi) > d$, whereas in the latter we get absolute continuity when the contraction is close enough to 1, but miss part of the potential parameter range. It is not known if this additional assumption is necessary for the conclusion.

2.6 Parametric families. Many classical problems in geometric measure theory involve parametric families of IFSs, e.g. the Bernoulli convolutions and projection problems discussed below, and the $\{0, 1, 3\}$-problem mentioned earlier. In these problems one wants to show that dimension drop is rare in the parameter space.

To set notation, suppose that $\{\Phi_t\}_{t \in I}$ is a parametric family of IFSs on $\mathbb{R}$, so $\varphi_i^t(x) = r_i(t)(x - a_i(t))$ where $r_i : I \to (-1, 1) \setminus \{0\}$ and $a_i : I \to \mathbb{R}$ are given functions. For infinite sequences $\mathbf{i}, \mathbf{j} \in \Lambda^{\mathbb{N}}$ set

$$\Delta_{\mathbf{i},\mathbf{j}}(t) = \lim_{n \to \infty} \varphi^t_{i_1 \ldots i_n}(0) - \varphi^t_{j_1 \ldots j_n}(0)$$

It is clear that if Φ_t has exact overlaps then there exist $\mathbf{i}, \mathbf{j} \in \Lambda^{\mathbb{N}}$ with $\Delta_{\mathbf{i},\mathbf{j}}(t) = 0$, but $\Delta_{\mathbf{i},\mathbf{j}}$ may certainly vanish also when there are non-exact overlaps. However, under an analyticity and non-degeneracy assumption, the zeros of $\Delta_{\mathbf{i},\mathbf{j}}(\cdot)$ will be isolated, and the function will grow polynomially away from its zeros.[4] Furthermore, by a compactness argument, the rate is uniform. Also, assuming analyticity, lower bounds on $\Delta_{\mathbf{i},\mathbf{j}}$ can be translated to lower bounds for $|\varphi^t_{i_1 \ldots i_n}(0) - \varphi^t_{j_1 \ldots j_n}(0)|$, and hence for $d(\varphi^t_{i_1 \ldots i_n}, \varphi^t_{j_1 \ldots j_n})$. From these ingredients one obtains efficient covers of the set of parameters for which Φ_t is not exponentially separated. The end result of this analysis is the following.

Theorem 2.12 (Hochman [2014]). *Let $I \subseteq \mathbb{R}$ be a compact interval, let $r : I \to (-1, 1) \setminus \{0\}$ and $a_i : I \to \mathbb{R}$ be real analytic, and let $\Phi_t = \{\varphi_{i,t}\}_{i \in \Lambda}$ be the associated parametric family of IFSs, as above. Suppose that*

$$\forall \mathbf{i}, \mathbf{j} \in \Lambda^{\mathbb{N}} \quad (\, \Delta_{\mathbf{i},\mathbf{j}} \equiv 0 \text{ on } I \quad \Longleftrightarrow \quad \mathbf{i} = \mathbf{j} \,).$$

Then the set of $t \in I$ for which there is dimension drop has Hausdorff and packing dimension 0.

[4] In contrast, the classical transversality method for parametric families depends on showing that $\Delta_{\mathbf{i},\mathbf{j}}$ grows linearly away from its zeros, i.e., it requires one to show that all zeros are simple. When this holds one often gets stronger conclusions, e.g. absolute continuity of the measures outside a small (though generally not zero-dimensional) set of parameters. But it is much harder to establish that the zeros are simple, and not always true. For more information we refer to Peres, Schlag, and Solomyak [2000] and Peres and Schlag [2000]. A major benefit of the method above is that polynomial growth is automatic.

Analogous statements hold in $\mathbb{R}^d$, giving, under some assumptions, that the exceptional parameters have dimension $\le d-1$. For details see Hochman [2015].

2.7 Further developments. The same questions can be asked about attractors of non-linear IFSs. The only such case where a version of Theorem 2.3 is known is for linear fractional transformations Hochman and Solomyak [2017]. Little is known beyond this case.

Another natural problem is to extend the results to self-affine sets and measures, defined in the same way but using affine maps rather than similarities. This area is developing rapidly, and it seems likely that analogous results will be established there in the near future.

Finally, we mention a result of Fraser, Henderson, Olson, and Robinson [2015], showing that if a self-similar set in $\mathbb{R}$ does not have exact overlaps then its Assouad dimension is one. This is a very weak notion of dimension, equal to the maximal dimension of any set which is a Hausdorff limit of magnifications of X. It says nothing about $\dim X$ itself, but it lends moral support to the idea that without exact overlaps, X is "as large as possible".

3 Bernoulli convolutions

In the "supercritical" case $s(\Phi) > 1$, Conjecture 2.2 has a stronger variant:

Conjecture 3.1. *Let $\mu = \mu_{\Phi,p}$ be a self-similar measure on $\mathbb{R}$. If there are no exact overlaps and*[5] *$s(\Phi, p) > 1$, then μ is absolutely continuous with respect to Lebesgue measure.*

The main evidence supporting the conjecture comes from the study of parametric families, the primary example of which are Bernoulli convolutions. For $0 < \lambda < 1$ the *Bernoulli convolution with parameter λ* is the distribution ν_λ of the real random variable $\sum_{n=0}^{\infty} \pm\lambda^n$, where the signs are chosen i.i.d. with $\mathbb{P}(+) = \mathbb{P}(-) = \frac{1}{2}$. The name derives from the fact that ν_λ can be written as the infinite convolution of the measures $\frac{1}{2}(\delta_{-\lambda^n} + \delta_{\lambda^n})$, $n = 0, 1, 2, \ldots$, but it is also a self-similar measure for the IFS $\Phi_\lambda = \{\varphi_{\pm 1}\}$, defined by assigning equal probabilities to each of the maps

$$\varphi_{\pm 1}(x) = \lambda x \pm 1. \tag{13}$$

Let $\Lambda = \{\pm 1\}$. For $\mathbf{i}, \mathbf{j} \in \Lambda^n$, the maps $\varphi_{\mathbf{i}}, \varphi_{\mathbf{j}}$ contract by λ^n, so

$$d(\varphi_{\mathbf{i}}, \varphi_{\mathbf{j}}) = |\varphi_{\mathbf{i}}(0) - \varphi_{\mathbf{j}}(0)| = |\sum_{k=0}^{n} (i_k - j_k)\lambda^k|$$

[5] By a theorem of Schief [1994], absolute continuity can fail in the "critical" case $s(\Phi, p) = 1$.

Hence, since $i_k - j_k \in \{-2, 0, 2\}$, exact overlaps occur if and only if λ is the root of a polynomial in with coefficients $-1, 0, 1$. Write $\nu_\lambda^{(n)}$ for the distribution measure of the finite sum $\sum_{k=0}^{n} \pm\lambda^k$.

The case $\lambda < \frac{1}{2}$ is simple from the point of view of dimension: Φ_λ satisfies the SSC and $\dim \nu_\lambda = \log \lambda / \log(1/2)$. Also, $\nu_{1/2}$ is uniform on $[-2, 2]$.

Things are more interesting for $\lambda > \frac{1}{2}$. Then the Lyapunov dimension is > 1, the attractor is an interval, and Φ_λ has overlaps. From Conjectures 2.2 and 3.1 one would expect that ν_λ is absolutely continuous (and $\dim \nu_\lambda = 1$) unless there are exact overlaps. Thus, we shall say that λ is *a.c.-exceptional* or *dim-exceptional*, if ν_λ is singular or $\dim \nu_\lambda < 1$, respectively. We denote the sets of these parameters by E_{ac} and $E_{\dim}$.

It was Erdős who found the first, and so far only, exceptional parameters: if λ^{-1} is a Pisot number[6] then λ is a.c.-exceptional, and Garsia later showed that such λ are also dim-exceptional. Perhaps these are the only ones; some support for this is Salem's theorem that $|\widehat{\nu}_\lambda(t)| \to 0$ as $t \to \infty$ if and only if λ^{-1} is not Pisot Salem [1944].

3.1 Bounds on the size of the exceptional parameters. Much of the work on Bernoulli convolutions has focused on bounding the size of the set of exceptions. The work of Erdős and Kahane implies that $\dim((a, 1) \cap E_{\mathrm{ac}}) \to 0$ as $a \nearrow 1$, and Erdős proved that $\dim \nu_\lambda \to 1$ as $\lambda \to 1$ (see also Peres and Schlag [2000]).

A major step forward was the 1995 proof by Solomyak [1995] that ν_λ is a.c. with L^2 density, for almost every $\lambda \in (\frac{1}{2}, 1)$. This was one of the early successes of the transversality method. Some improvements, including some bounds on the dimension of exceptions, were later obtained by Peres and Schlag [2000].

Theorem 2.12 leads to further improvements:

Theorem 3.2 (Hochman [2014]). $\dim \nu_\lambda = 1$ *outside a set of* λ *of Hausdorff and packing dimension* 0.

Currently, these techniques don't recover Solomyak's theorem directly, but combined with Fourier-theoretic information, Shmerkin managed to prove

Theorem 3.3 (Shmerkin [2014, 2016]). *Outside a set of* λ *of Hausdorff dimension* 0*, the measure* ν_λ *is absolutely continuous with density in* L^p *for all* $1 \le p < \infty$.

Here is the idea of the proof. Fix an integer k, and split the random sum as

$$\sum_{n=0}^{\infty} \pm\lambda^n = \sum_{n=0 \bmod k} \pm\lambda^{kn} + \sum_{n \neq 0 \bmod k} \pm\lambda^n \tag{14}$$

[6] A Pisot number is an algebraic integer greater than one, all of whose conjugates lie in the interior of the unit disk.

The first term on the right has distribution ν_{λ^k}. Write τ_{λ^k} for the distribution of the second term. Since the two series are mutually independent, we get $\nu_\lambda = \nu_{\lambda^k} * \tau_{\lambda^k}$. Next, using an energy-theoretic argument, it is shown that the convolution will be absolutely continuous provided that the Fourier transform $\widehat{\nu}_{\lambda^k}$ has power decay (i.e. $|\widehat{\nu}_{\lambda^k}(t)| \leq t^{-c}$ for some $c = c(\lambda) > 0$) and $\dim \tau_{\lambda^k} = 1$. Now, a classical result of Erdős and Kahane (see Peres, Schlag, and Solomyak [2000]) says that outside a zero-dimensional set E' of parameters, the Fourier transform $\widehat{\nu}_{\lambda^k}$ indeed has power decay; on the other hand, τ_{λ^k} is itself a parametric family of self-similar measures, and Theorem 2.12 implies that there is dimension drop for a set E'' of parameters of dimension zero. For $\lambda \notin E''$ we will have $\dim \tau_{\lambda^k} = 1$ whenever the Lyapunov dimension is ≥ 1, which by a short calculation happens when $\lambda \in ((1/2)^{1-1/k}, 1)$. Thus, ν_λ is absolutely continuous for $\lambda \in [(1/2)^{1-1/k}, 1] \setminus (E' \cup E'')$. Taking the union over k gives the claim.

In order to obtain densities in L^p, an analogous argument is carried out using Theorem 2.7 instead of Theorem 2.3.

There remains a difference between Theorems 3.2 and 3.3. In the former, a parameter is "good" if Φ_λ is exponentially separated, which gives new explicit examples, e.g. all rational parameters. The set E'' in the above is similarly explicit. But E' is completely ineffective. Consequently, to get new examples of absolutely continuous ν_λ requires other methods, see Section 3.3.

3.2 Mahler measure. The *Mahler measure* of an algebraic number λ is $M_\lambda = |a| \cdot \prod_{|\xi|>1} |\xi|$ where the product is over Galois conjugates ξ of λ, and a is the leading coefficient of its minimal polynomial. This is a standard measure of the size or complexity of an algebraic number. It first appeared in connection with Bernoulli convolution in the following lemma of Garsia:

Lemma 3.4 (Garsia [1962]). *Let $\lambda > 1$ be algebraic with conjugates $\lambda_1, \ldots, \lambda_s \neq \lambda$, of which σ lie on the unit circle. Let $p(x) = \sum_{i=0}^n a_i x^i$ be an integer polynomial and $A = \max\{|a_i|\}$. Then either $p(\lambda) = 0$, or*

$$|p(\lambda)| \geq \frac{\prod_{|\lambda_i| \neq 1} ||\lambda_i| - 1|}{A^s (n+1)^\sigma \left(\prod_{|\lambda_i|>1} |\lambda_i|\right)^n} \geq \frac{C_\lambda}{A^s (n+1)^\sigma M_\lambda^n}$$

In particular, any distinct atoms of the n-th approximation $\nu_\lambda^{(n)}$ of ν_λ are separated by at least $C_\lambda((n+1)^\sigma M_\lambda^n)^{-1}$.

Recently, Mahler measure has been related to the random walk entropy h_λ associated to ν_λ in Section 2.3:

Theorem 3.5 (Breuillard and P. P. Varjú [2015]). *There exists a constant $c > 0$ such that for any algebraic number $\lambda \in (\frac{1}{2}, 1)$,*

$$c \min\{1, \log M_\lambda\} \leq h_\lambda \leq \min\{1, \log M_\lambda\}$$

In particular, if $\lambda > \min\{2, M_\lambda\}^{-c}$ (for c a as in the theorem), then $\dim \nu_\lambda = 1$. We do not discuss the proof of the theorem here, the reader may consult Breuillard and P. P. Varjú [ibid.].

3.3 Absolute continuity for algebraic parameters. By the identity $\nu_\lambda = \nu_{\lambda^k} * \tau_{\lambda^k}$, which follows from (14), $\nu_{1/2^{1/k}} = \nu_{1/2} * \tau_{1/2}$, hence, since $\nu_{1/2}$ is the uniform measure on an interval, $\nu_{1/2^{1/k}}$ is absolutely continuous. Garsia identified a less trivial class of examples: those λ such that λ^{-1} is an algebraic integer, $\sigma = 0$ and $M_\lambda = 2$. Such numbers are not roots of $0, \pm 1$-polynomials, so $\sum_{i=0}^{n} \pm\lambda^n$ takes 2^n equally likely values, and by Lemma 3.4 the values are $c \cdot 2^{-n}$-separated (for some c). This implies that $\nu_\lambda = \lim \nu_\lambda^{(n)}$ is absolutely continuous. Until recently these were the only examples. We now have the following, which gives many more. For example, it applies to every rational number close enough to one in a manner depending on their denominator P. Varjú [2016, Section 1.3.1].

Theorem 3.6 (P. Varjú [ibid.]). *For every $\varepsilon > 0$ there is a $c > 0$ such such if $\lambda \in (\frac{1}{2}, 1)$ is algebraic and satisfies*

$$\lambda > 1 - c \min\{\log M_\lambda, (\log M_\lambda)^{-(1-\varepsilon)}\}$$

Then ν_λ is absolutely continuous with density in $L \log L$.

The proof relies on the following, which goes back to Garsia:

Theorem 3.7 (Garsia [1963]). *ν_λ is absolutely continuous with density in $L \log L$ if and only if $H(\nu_\lambda^{(n)}) = n - O(1)$.*

The argument in Theorem 2.3 gives $H(\nu_\lambda^{(n)}) = n - o(n)$; in order to get the $O(1)$ error required by Garsia's theorem, Varjú proves two quantitative variants of the general entropy-growth result underlying Theorem 2.8. Roughly speaking,[7] the first shows that if α is small enough, and measures θ_1, θ_2 satisfy $H(\theta_i, \mathfrak{D}_{n+1}|\mathfrak{D}_n) > 1 - \alpha$, then $H(\theta_1 * \theta_2, \mathfrak{D}_{n+1}|\mathfrak{D}_n) > 1 - \alpha^2$. The second is analogous to Theorem 2.8 but with $\delta = c\varepsilon$. Now, fixing N, split the series $\sum_{n=0}^{N} \pm\lambda^n$ into $k = [\log(N^2)]$ finite sums $\sum_{n \in I_i} \pm\lambda^n$ of distribution $\nu_\lambda^{(N,i)}$ respectively, so that $\nu_\lambda^{(N)} = \nu_\lambda^{(N,1)} * \ldots * \nu_\lambda^{(N,k)}$. If we can choose

[7] We have omitted many assumptions, logarithmic factors, and even then the entropy inequalities are false using Shannon entropy; one must use spatially averaged entropy, see P. Varjú [2016].

$I_1, \ldots, I_k$ so that $H(\nu_\lambda^{(N,i)}, \mathfrak{D}_N|\mathfrak{D}_{N-1}) > 1-\alpha_0$, with α_0 small, we can apply the first entropy growth result iteratively, and get $H(\nu_\lambda^{(N)}, \mathfrak{D}_N|\mathfrak{D}_{N-1}) > 1-\alpha_0^{2^{\log k}} = O(1/N^2)$; summing over $0 \le N \le M$ gives $H(\nu_\lambda^{(M)}) = M - O(1)$, as desired. In order to find $I_1, \ldots, I_k$ as above, one uses a similar argument, relying on the second entropy-growth theorem to amplify the random walk entropy provided by Theorem 3.5. For more details, we refer to the original paper.

3.4 Dimension results for other parameters. Let $\mathcal{P}$ be the set of polynomials with coefficients $0, \pm 1$ and set $\mathcal{P}_n = \{f \in \mathcal{P} : \deg f \le n\}$. Suppose that $\dim \nu_\lambda < 1$. Then by Theorem 2.3, $\Delta_n \to 0$ super-exponentially, i.e. there exist $p_n \in \mathcal{P}_n$ such that $p_n(\lambda) \to 0$ super-exponentially. This does not force λ to be algebraic, but using transversality arguments or Jensen's formula, one can find roots λ_n of p_n such that $|\lambda_n - \lambda| \to 0$ super-exponentially, so λ_n, λ_{n+1} are super-exponentially close. If the roots of elements of $\mathcal{P}_n$ were sufficiently (i.e. exponentially) separated, this would force the sequence λ_n to stabilize, and λ would be algebraic, in fact a root of some $p_{n_0} \in \mathcal{P}$. Thus, an affirmative answer to the following problem would imply $\dim \nu_\lambda = 1$ for all λ without exact overlaps:

Question 3.8. *Does there exist a constant $c > 0$ such that if $\alpha \neq \beta$ are roots of (possibly different) polynomials in $\mathcal{P}_n$, then $|\alpha - \beta| > c^n$?*

The best current bound, due to Mahler [1964], is of the form $|\alpha - \beta| > n^{-cn}$ for some constant $c > 0$. In order for this to be useful, one needs to get a similar rate for the decay of Δ_n in Theorem 2.3. This is essentially the content of the following:

Theorem 3.9 (Breuillard and P. P. Varjú [2016]). *Suppose that $\dim \nu_\lambda < 1$ for some $\lambda \in (\frac{1}{2}, 1)$. Then there exist arbitrarily large n for which there is an algebraic number ξ that is a root of a polynomial in $\mathcal{P}_n$, such that*

$$|\lambda - \xi| < \exp(-n^{\log\log\log n})$$

and

$$\dim \nu_\xi < 1.$$

Compared to the argument at the start of the section, notice that the numbers ξ are only guaranteed to exist for infinitely many n, not all large enough n. Therefore, even though the rate is better than Mahler's bound, one cannot conclude that the approximants stabilize. But, most importantly, the algebraic number ξ are guaranteed to be themselves exceptional for dimension. The latter has a dramatic implication:

Theorem 3.10 (Breuillard and P. P. Varjú [ibid.]). $E_{\dim} = \overline{E_{\dim} \cap \overline{\mathbb{Q}}}$, *where $\overline{\mathbb{Q}}$ denotes the algebraic closure of $\mathbb{Q}$.*

This reduces the question of the dimension of Bernoulli convolutions to the algebraic case. In particular, it is known that the Pisot numbers form a closed set, so if these were shown to be the only algebraic parameters with dimension drop, these would be the only exceptions altogether. Also, recall that Lehmer's famous problem asks whether $\inf\{M_\lambda : M_\lambda \neq 1\} > 1$. If this were true, then Theorem 3.5 would imply that there is an $\varepsilon > 0$ such that $\dim \nu_\lambda = 1$ for every algebraic $\lambda \in (1-\varepsilon, 1)$; combined with Theorem 3.10, this gives:

Theorem 3.11 (Breuillard and P. P. Varjú [2016]). *If the answer to Lehmer's problem is affirmative, then there is an $\varepsilon > 0$ such that $\dim \nu_\lambda = 1$ for every $\lambda \in (1-\varepsilon, 1)$.*

There is no known converse, but for a related result see Peres, Schlag, and Solomyak [2000, Proposition 5.1].

Finally, using the information gained about algebraic approximants of exceptional parameters, Breuillard and Varjú have managed to find the first explicit transcendental parameters for which ν_λ has full dimension; e.g. $e, 1/\ln 2$, and other natural constants. For details see Breuillard and P. P. Varjú [2016].

4 Projection and slice theorems

A basic principle in (fractal) geometry is that projections of a set typically should be "as large as possible", and slices should be correspondingly small. By a projection of a set we mean its image under an orthogonal projection π_V to a linear subspace V, and by a slice we mean its intersection with an affine subspace. The trivial bound for projections is $\dim \pi_V X \leq \min\{\dim X, \dim V\}$ (because π_V is Lipschitz and has range V). The following theorem shows that this is generally the right bound, and that slices behave dually. Let $G(d,k)$ denote the manifold of k-dimensional linear subspaces of $\mathbb{R}^d$.

Theorem 4.1 (Marstrand [1954], Mattila [1975]). *Let $E \subseteq \mathbb{R}^d$ be Borel and $1 \leq k < d$. Then for a.e. $V \in G(d,k)$,*

$$\dim \pi_V E = \min\{k, \dim E\}$$

If in addition $\dim E > k$ then[8] *$\pi_V(E)$ has positive k-dimensional volume for a.e. $V \in G(d,k)$, and for a.e. $y \in V$ with respect to the volume,*

$$\dim E \cap \pi_V^{-1}(y) \leq \max\{0, \dim E - k\}$$

[8] Although unrelated to our discussion, it is worth mentioning that if $\dim E = k$, then $\pi_V E$ will have positive k-dimensional volume depending on whether it is rectifiable or purely unrectifiable. See Mattila [1995].

Kaufman, Falconer and Mattila (see e.g. Mattila [1995]) also bounded the dimension of the set of exceptional $V \in G(d,k)$, e.g. for $d = 2$ and $k = 1$, if $\dim E < k$ the exceptions have dimension $\leq \dim E$, and if $\dim E > k$ it is at most $2 - \dim E$.

What these general results for generic directions fail to give is any information at all about particular directions. For "natural", well-structured sets, one would expect to be able to be more precise. What one expects in such cases is that the projections will be as large as possible unless there is some combinatorial or algebraic obstruction; and that slices be correspondingly small.

In the coming discussion we restrict attention to self-similar sets in $\mathbb{R}^2$, where results are more complete. We mention measures and multi-dimensional analogues only occasionally, since these require more assumptions. We also do not discuss results on randomly generated fractals, for this see K. Rams M. S. [2015], Peres and M. Rams [2016], and Shmerkin and Suomala [2017].

4.1 Dimension conservation. Heuristically, projections to V and slices in direction $V^\perp$ are complementary, in the sense that having a large image forces most slices to be small, and vice versa. This is exactly true for finite-scale entropy, and combinatorial versions can also be formulated. For dimension, this duality does not always hold. The following "dimension conservation" result , a relative of the Ledrappier-Young formula, marked the start of the current phase of research.

Theorem 4.2 (Furstenberg [2008]). *Let $X \subseteq \mathbb{R}^d$ be a self-homothetic set. Then for every $V \in G(d,k)$, we have*

$$\dim \pi_V X + \sup_{y \in V} \dim(X \cap \pi_V^{-1}(y)) \geq \dim X$$

For self-homothetic measures μ there is in fact equality for $\pi_V \mu$-a.e. y, and until recently it was not known whether a similar phenomenon might hold also outside of homotheties. It turns out that it does not:

Theorem 4.3 (Rapaport [2016]). *There exists a self-similar measure μ on $\mathbb{R}^2$ with* $\dim \mu > 1$, *uniform contractions and uniform dense rotations, such that for a dense G_δ set of directions V the conditional measures on translates of V are a.s. atomic, hence*

$$\dim \pi_V \mu + \operatorname*{esssup}_{y \sim \pi_V \mu} \dim \mu_{\pi^{-1}(y)} < \dim \mu$$

Problem 4.4. For a self-similar set in $\mathbb{R}^2$ with dense rotations and dimension greater than one, do all projections have positive length?

4.2 Projections of self-homothetic sets. Self-homothetic self-similar sets X have the special property that each projection $\pi_V X$ is also self-homothetic, being the attractor of and IFS $\Phi_V = \{\varphi_{i,V}(x) = r_i x + \pi_V a_i\}$. Assuming that Φ has strong separation, and fixing a self-similar measure on X, Theorem 2.12 applies to the parametric family $\{\Phi_V\}_{V \in G(2,1)}$. In the case $\dim X > 1$ an argument similar to that in Theorem 3.3 also applies. All in all, we get

Theorem 4.5 (Hochman [2014], Shmerkin [2015])**.** *If $X \subseteq \mathbb{R}^2$ is self-homothetic then $\dim \pi_V X = \min\{1, \dim X\}$ for all but a zero dimensional set of $V \in G(2,1)$. If also $\dim X > 1$, the projection will have positive length outside a set of V of dimension zero.*

The reason no separation condition is needed is that a self-similar set X always contains smaller self-similar sets of dimension arbitrarily close to $\dim X$, and satisfying strong separation and uniform contraction (Peres and Shmerkin [2009]), and it is enough to show that these subsets have large projections.

More is true when the maps φ_i are algebraic, by which we mean that the parameters r_i, a_i are algebraic. The conjecture that this is the case was raised by Furstenberg for the so-called 1-dimensional Sierpinski gasket, first appearing in the work of Kenyon [1997]. The next theorem follows from Corollary 2.4 and an argument (originally due to Solomyak) showing that if the projected IFS Φ_V does not have exact overlaps, then it is exponentially separated.

Theorem 4.6 (Hochman [2014])**.** *If X is the attractor of an algebraic IFS consisting of homotheties, then $\dim \pi_V X = \min\{1, \dim X\}$ for all except at most countably many $V \in G(2,1)$, which are among the V which collapse cylinders, i.e. $\pi_V \varphi_{\mathbf{i}} = \pi_V \varphi_{\mathbf{j}}$ for some n and distinct $\mathbf{i}, \mathbf{j} \in \Lambda^n$.*

There certainly can exist exceptional directions, but they have not been entirely characterized (a special case was analyzed by Kenyon [1997]). Currently, no analogous result exists for the Lebesgue measure of the projection in the regime $\dim X > 1$. Finally, note that the result for measures seems to require strong separation, since for measures, there is no analog of the trick of passing to a sub-self-similar set.

All statements above hold if instead of self-homothetic sets we allow IFSs whose orthogonal parts generate a finite group of rotations.

4.3 Projections of sets and measures with rich symmetries . Let $\Phi = \{\varphi_i\}$ be an IFS in $\mathbb{R}^2$ with $\varphi_i(x) = r_i O_i x + a_i$, with $0 < r_1 < 1$, O_i an orthogonal matrix, and $a_i \in \mathbb{R}^2$. We say that Φ has *irrational rotations* if at least one O_i is an irrational rotation (has infinite order). Notice that $\pi_V \varphi_i X \subseteq \pi_V X$, and $\pi_V \varphi_i$, up to change of coordinates, is projection to $O_i^{-1} V$. Iterating this and using the fact that $\{O_{i_1}^{-1} \dots O_{i_n}^{-1} V\}_{\mathbf{i} \in \Lambda^*}$ is dense, we see that

$\pi_V X$ contains subsets approximating every other projection $\pi_W X$. This can be used to prove that there are no exceptional directions:

Theorem 4.7 (Peres and Shmerkin [2009], Nazarov, Peres, and Shmerkin [2012], Hochman and Shmerkin [2011]). *Let $X \subseteq \mathbb{R}^2$ be a self-similar set with dense rotations. Then* $\dim \pi_V X = \min\{1, \dim X\}$ *for every V. The same holds for self-similar measures assuming the open set condition.*

This was first proved for sets by Peres and Shmerkin [2009] for sets. Nazarov, Peres, and Shmerkin [2012] proved an analog for measures assuming uniform rotations. Hochman and Shmerkin [2011] proved the general version. See also Farkas [2016].

We briefly sketch the proof of Hochman-Shmerkin. A central ingredient is the method of local entropy averages. Suppose that ν is any probability measure on $\mathbb{R}^d$. Let $\mathfrak{D}_n(x)$ denote the unique atom of $\mathfrak{D}_n$ containing x and let

$$\nu_{x,n} = \frac{1}{\nu(\mathfrak{D}_n(x))} \nu|_{\mathfrak{D}_n(x)}$$

Thus the sequence $(\nu_{x,n})_{n=1}^{\infty}$ is what you see when you "zoom in" to x along dyadic cells.

Theorem 4.8 (Hochman and Shmerkin [2011]). *Let ν be a Borel probability measure on $\mathbb{R}^d$, and $V \in G(d,k)$. If for some $\alpha \geq 0$ and $m \in \mathbb{N}$*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} H(\nu_{x,n}, \mathfrak{D}_{n+m}) \geq \alpha \tag{15}$$

then $\dim \pi_V \nu \geq \alpha - O_{d,k}(1/m)$.

If μ is self-similar with OSC, then $\mu_{x,n}$ is a piece (or a combination of boundedly many pieces) of a cylinder measure of diameter approximately 2^{-n}, and if there are dense rotations, for μ-a.e. x the rotations observed along the sequence $(\mu_{x,n})_{n=1}^{\infty}$ equidistribute in the circle. Now, by Marstrand's theorem, $\dim \pi_V \mu = \min\{1, \dim \mu\}$ for a.e. line V, so by (8), for large enough m, with high probability over V we have $\frac{1}{m} H(\pi_V \mu, \mathfrak{D}_m) > \min\{1, \dim \mu\} - \varepsilon$. This condition on V is essentially open, so $(\mu_{x,n})_{n=1}^{\infty}$ consists predominantly of measures that, after re-scaling by 2^n, satisfy this inequality. Theorem 4.7 now follows from Theorem 4.8.

This method also gives the following, which was conjectured by Furstenberg:

Theorem 4.9 (Hochman and Shmerkin [ibid.]). *Let $Y_a, Y_b \subseteq [0,1]$ be closed and invariant under $x \mapsto ax \bmod 1$, $x \mapsto bx \bmod 1$, and assume $\log a / \log b \notin \mathbb{Q}$. Then* $\dim \pi_V(Y_a \times Y_b) = \min\{1, \dim Y_a + \dim Y_b\}$ *for all V except the horizontal and vertical directions.*

The methods also apply in non-linear cases. This has some overlap with the work of Moreira on nonlinearly generated fractals, which predates all the results above, but does not apply in the linear case. See Moreira [1998].

Problem 4.10. If Y_a, Y_b are as in the theorem, and $\dim Y_a \times Y_b > 1$, does $\pi_V(Y_a \times Y_b)$ have positive Lebesgue measure for all V not parallel to the axes?

Surprisingly the analogous problem for products of self-similar measures has a negative answer, see Nazarov, Peres, and Shmerkin [2012].

4.4 Slices . Dual to the projections problem is that of slices. When projections achieve their maximal value, one might expect slices not to exceed their typical value. Furstenberg [1970] conjectured this for non-vertical/horizontal slices of products as in Theorem 4.9. Very recently two independent proofs of this emerged:

Theorem 4.11 (Shmerkin [2016], Wu [2016])**.** *Let $X = Y_a \times Y_b$ be as in Theorem 4.9. Then $\dim(X \cap \ell) \leq \max\{0, \dim X - 1\}$ for all lines $\ell \subseteq \mathbb{R}^2$ not parallel to the axes. Similarly if X is self-similar set with uniform contraction and uniform dense rotations, the bounds holds for all ℓ.*

The case $\dim X < 1/2$ was proved by Furstenberg. He showed that if $\dim X \cap \ell = \alpha$, then there exists a stationary ergodic process $Z = (x_n, \theta_n)_{n=1}^{\infty}$ with $x_n \in X$ and $\theta_n \in [0,1]$, such that a.s. the line $\ell(x_n, \theta_n)$ of slope θ_n through x_n satisfies $\dim(X \cap \ell(x_n, \theta_n)) = \alpha$, and the process (θ_n) has pure point spectrum. Thus θ_1 is uniform on $[0,1]$, so in a.e. direction there is a pair $x', x'' \in X$ with $x' - x''$ pointing in that direction. This implies $\dim X \geq 1/2$.

Wu's proof is ergodic-theoretic and begins with the same construction. Now, if there were a point $\xi \in X$ such that the distribution of θ_1 is uniform given $x_1 = \xi$, this would give a "bouquet" of α-dimensional slices passing through ξ and pointing in a 1-dimensional set of directions, and imply $\dim X \geq 1 + \alpha$, the desired bound on α. To find such ξ, first apply a classical theorem of Sinai to get a Bernoulli factor W of the process $Z = (x_n, \theta_n)$ exhausting the entropy, i.e. $h(Z|W) = 0$. Let P_w denote the disintegration of the distribution of Z over $w \in W$ and let $Q_w = \mathbb{E}(x_1|w)$ be the image of P_w in X. Next, self-similarity gives an expanding conformal dynamics on X, and the process Z can be constructed such that $\dim Q_w$ is proportional to $h(Z|W)$, hence $\dim Q_w = 0$. Finally, $\Theta = (\theta_n)$ is a rotation and W is Bernoulli, so by the disjointness theory of Furstenberg [1967], Θ is independent of W, hence θ_1 is distributed uniformly conditioned on w. Thus, there is a family of slices of dimension α, with uniformly distributed directions, passing through the points of the zero-dimensional measure Q_w. From this one can construct (approximations of) the desired bouquets.

Shmerkin's proof is entirely different. Consider the self-similar case with the natural self-similar measure μ on X. It is a basic fact that if $\inf_{q>1} D(\pi_V \mu, q) = \alpha$, then $\dim(X \cap \pi_V^{-1}(y)) \leq \dim X - \alpha$ for all y. Thus the goal is to show that the L^q dimension is maximal in all directions. Now, $\pi_V \mu$ are not self-similar but it has a convolutions structure, because all cylinders of a given generation in μ differ only by translations. By an argument similar to Theorems 2.7 and 2.12, we conclude that $\inf_{q>1} D(\pi_V \mu, q) = \min\{1, \dim X\}$ for a large set of $V \in G(2,1)$. To extend this for all directions, one uses (among many other things) unique ergodicity of a certain cocycle arising from the rotational symmetry of X.

Finally, Theorem 2.7 implies an L^q version of Theorem 4.6, which gives a dual result for the slices, confirming another old conjecture of Furstenberg:

Theorem 4.12 (Shmerkin [2016]). *Let $X \subseteq \mathbb{R}^2$ be a self-homothetic algebraic self-similar measure. Then outside a countable set of directions, there are no exceptional slices.*

References

Christoph Bandt and Siegfried Graf (1992). "Self-Similar Sets 7. A Characterization of Self-Similar Fractals with Positive Hausdorff Measure". *Proceedings of the American Mathematical Society* 114.4, pp. 995–1001 (cit. on p. 1971).

Yves Benoist and Jean-François Quint (2009). "Mesures stationnaires et fermés invariants des espaces homogènes". *C. R. Math. Acad. Sci. Paris* 347.1-2, pp. 9–13. MR: 2536741 (cit. on p. 1968).

J. Bourgain (2003). "On the Erdős-Volkmann and Katz-Tao ring conjectures". *Geom. Funct. Anal.* 13.2, pp. 334–365. MR: 1982147 (cit. on pp. 1974, 1975).

Jean Bourgain (2010). "The discretized sum-product and projection theorems". *J. Anal. Math.* 112, pp. 193–236. MR: 2763000 (cit. on pp. 1974, 1975).

Jean Bourgain, Alex Furman, Elon Lindenstrauss, and Shahar Mozes (2007). "Invariant measures and stiffness for non-abelian groups of toral automorphisms". *C. R. Math. Acad. Sci. Paris* 344.12, pp. 737–742. MR: 2340439 (cit. on p. 1968).

Emmanuel Breuillard and Péter P Varjú (2015). "Entropy of Bernoulli convolutions and uniform exponential growth for linear groups". arXiv: 1510.04043 (cit. on p. 1982).

– (2016). "On the dimension of Bernoulli convolutions". arXiv: 1610.09154 (cit. on pp. 1983, 1984).

Thomas M. Cover and Joy A. Thomas (2006). *Elements of information theory*. Second. Hoboken, NJ: Wiley-Interscience [John Wiley & Sons], pp. xxiv+748. MR: 2239987 (cit. on p. 1974).

K. J. Falconer (1989). "Dimensions and measures of quasi self-similar sets". *Proc. Amer. Math. Soc.* 106.2, pp. 543–554. MR: 969315 (cit. on pp. 1969, 1971).

Ábel Farkas (2016). "Projections of self-similar sets with no separation condition". *Israel Journal of Mathematics* 214, pp. 389–404 (cit. on p. 1987).

De-Jun Feng and Huyi Hu (2009). "Dimension theory of iterated function systems". *Comm. Pure Appl. Math.* 62.11, pp. 1435–1500. MR: 2560042 (cit. on p. 1970).

J.M. Fraser, A.M. Henderson, E.J. Olson, and J.C. Robinson (2015). "On the Assouad dimension of self-similar sets with overlaps". *Advances in Mathematics* 273.Supplement C, pp. 188–214 (cit. on p. 1979).

Harry Furstenberg (1967). "Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation". *Math. Systems Theory* 1, pp. 1–49. MR: 0213508 (cit. on p. 1988).

– (1970). "Intersections of Cantor sets and transversality of semigroups". In: *Problems in analysis (Sympos. Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*. Princeton, N.J.: Princeton Univ. Press, pp. 41–59. MR: 0354562 (cit. on p. 1988).

Hillel Furstenberg (2008). "Ergodic fractal measures and dimension conservation". *Ergodic Theory Dynam. Systems* 28.2, pp. 405–422. MR: 2408385 (cit. on p. 1985).

Adriano M. Garsia (1962). "Arithmetic properties of Bernoulli convolutions". *Trans. Amer. Math. Soc.* 102, pp. 409–432. MR: 0137961 (cit. on p. 1981).

– (1963). "Entropy and singularity of infinite convolutions". *Pacific Journal of Math.* 13, pp. 1159–1169. MR: 0156945 (cit. on p. 1982).

Luis G. Gorostiza (1973). "The central limit theorem for random motions of d-dimensional Euclidean space". *Ann. Probability* 1, pp. 603–612. MR: 0353408 (cit. on p. 1967).

Michael Hochman (2014). "On self-similar sets with overlaps and inverse theorems for entropy". *Ann. of Math. (2)* 180.2, pp. 773–822. MR: 3224722 (cit. on pp. 1972, 1974, 1975, 1978, 1980, 1986).

– (2015). "On self-similar sets with overlaps and inverse theorems for entropy in $\mathbb{R}^d$". To appear in Memoirs of the Amer. Math. Soc. arXiv: 1503.09043 (cit. on pp. 1977, 1979).

Michael Hochman and Pablo Shmerkin (2011). "Equidistribution from fractals" (cit. on p. 1987).

Michael Hochman and Boris Solomyak (Dec. 2017). "On the dimension of Furstenberg measure for $SL_2(\mathbb{R})$ random matrix products". *Inventiones mathematicae* 210.3, pp. 815–875 (cit. on p. 1979).

John E. Hutchinson (1981). "Fractals and self-similarity". *Indiana Univ. Math. J.* 30.5, pp. 713–747. MR: 625600 (cit. on pp. 1968, 1970, 1971).

Børge Jessen and Aurel Wintner (1935). "Distribution functions and the Riemann zeta function". *Trans. Amer. Math. Soc.* 38.1, pp. 48–88. MR: 1501802 (cit. on p. 1970).

Mike Keane, Meir Smorodinsky, and Boris Solomyak (1995). "On the morphology of γ-expansions with deleted digits". *Trans. Amer. Math. Soc.* 347.3, pp. 955–966. MR: 1290723 (cit. on p. 1971).

Richard Kenyon (1997). "Projecting the one-dimensional Sierpinski gasket". *Israel J. Math.* 97, pp. 221–238. MR: 1441250 (cit. on pp. 1971, 1986).

Elon Lindenstrauss (2010). "Equidistribution in homogeneous spaces and number theory". In: *Proceedings of the International Congress of Mathematicians. Volume I*. Hindustan Book Agency, New Delhi, pp. 531–557. MR: 2827904 (cit. on p. 1968).

Elon Lindenstrauss and Péter P. Varjú (2016). "Random walks in the group of Euclidean isometries and self-similar measures". *Duke Math. J.* 165.6, pp. 1061–1127. MR: 3486415 (cit. on p. 1977).

M. Madiman (2008). "On the entropy of sums". In: *Information Theory Workshop, 2008. ITW '08. IEEE*, pp. 303–307 (cit. on p. 1975).

K. Mahler (1964). "An inequality for the discriminant of a polynomial". *Michigan Math. J.* 11, pp. 257–262. MR: 0166188 (cit. on p. 1983).

G. A. Margulis and G. M. Tomanov (1994). "Invariant measures for actions of unipotent groups over local fields on homogeneous spaces". *Invent. Math.* 116.1-3, pp. 347–392. MR: 1253197 (cit. on p. 1968).

J. M. Marstrand (1954). "Some fundamental geometrical properties of plane sets of fractional dimensions". *Proc. London Math. Soc. (3)* 4, pp. 257–302. MR: 0063439 (cit. on p. 1984).

Pertti Mattila (1975). "Hausdorff dimension, orthogonal projections and intersections with planes". *Ann. Acad. Sci. Fenn. Ser. A I Math.* 1.2, pp. 227–244. MR: 0409774 (cit. on p. 1984).

– (1995). *Geometry of sets and measures in Euclidean spaces*. Vol. 44. Cambridge Studies in Advanced Mathematics. Fractals and rectifiability. Cambridge: Cambridge University Press, pp. xii+343. MR: 28006 (cit. on pp. 1984, 1985).

Carlos Gustavo T. de A. Moreira (1998). "Sums of regular Cantor sets, dynamics and applications to number theory". *Period. Math. Hungar.* 37.1-3. International Conference on Dimension and Dynamics (Miskolc, 1998), pp. 55–63. MR: 1719468 (cit. on p. 1988).

Fedor Nazarov, Yuval Peres, and Pablo Shmerkin (Jan. 2012). "Convolutions of cantor measures without resonance". *Israel Journal of Mathematics* 187.1, pp. 93–116 (cit. on pp. 1987, 1988).

Yuval Peres and Michał Rams (2016). "Projections of the natural measure for percolation fractals". *Israel J. Math.* 214.2, pp. 539–552. MR: 3544692 (cit. on p. 1985).

Yuval Peres and Wilhelm Schlag (2000). "Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions". *Duke Math. J.* 102.2, pp. 193–251. MR: 1749437 (cit. on pp. 1978, 1980).

Yuval Peres, Wilhelm Schlag, and Boris Solomyak (2000). "Sixty years of Bernoulli convolutions". In: *Fractal geometry and stochastics, II (Greifswald/Koserow, 1998)*. Vol. 46.

Progr. Probab. Basel: Birkhäuser, pp. 39–65. MR: 1785620 (cit. on pp. 1978, 1981, 1984).

Yuval Peres and Pablo Shmerkin (2009). "Resonance between Cantor sets". *Ergodic Theory Dynam. Systems* 29.1, pp. 201–221. MR: 2470633 (cit. on pp. 1986, 1987).

Yuval Peres and Boris Solomyak (2000). "Existence of L^q dimensions and entropy dimension for self-conformal measures". *Indiana Univ. Math. J.* 49.4, pp. 1603–1621. MR: 2002f:28012 (cit. on p. 1973).

Károly Rams Michałand Simon (2015). "Projections of fractal percolations". *Ergodic Theory Dynam. Systems* 35.2, pp. 530–545. MR: 3316924 (cit. on p. 1985).

Ariel Rapaport (2016). "On self-affine measures with equal Hausdorff and Lyapunov dimensions". To appear in Trans. Amer. Math. Soc. arXiv: 1607.07864 (cit. on p. 1985).

Marina Ratner (1991a). "On Raghunathan's measure conjecture". *Ann. of Math. (2)* 134.3, pp. 545–607. MR: 1135878 (cit. on p. 1968).

– (1991b). "Raghunathan's topological conjecture and distributions of unipotent flows". *Duke Math. J.* 63.1, pp. 235–280. MR: 1106945 (cit. on p. 1968).

– (1995). "Raghunathan's conjectures for Cartesian products of real and p-adic Lie groups". *Duke Math. J.* 77.2, pp. 275–382. MR: 1321062 (cit. on p. 1968).

Bernard Roynette (1974). "Théorème central-limite pour le groupe des déplacements de $\mathbf{R}^d$". *Ann. Inst. H. Poincaré Sect. B (N.S.)* 10, 391–398 (1975). MR: 0375422 (cit. on p. 1967).

R. Salem (1944). "A remarkable class of algebraic integers. Proof of a conjecture of Vijayaraghavan". *Duke Math. J.* 11.1, pp. 103–108 (cit. on p. 1980).

Andreas Schief (1994). "Separation properties for self-similar sets". *Proc. Amer. Math. Soc.* 122.1, pp. 111–115. MR: 1191872 (cit. on pp. 1971, 1979).

Pablo Shmerkin (2014). "On the exceptional set for absolute continuity of Bernoulli convolutions". *Geom. Funct. Anal.* 24.3, pp. 946–958. MR: 3213835 (cit. on p. 1980).

– (2015). "Projections of Self-Similar and Related Fractals: A Survey of Recent Developments". In: *Fractal Geometry and Stochastics V*. Ed. by Christoph Bandt, Kenneth Falconer, and Martina Zähle. Cham: Springer International Publishing, pp. 53–74 (cit. on p. 1986).

– (2016). "On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions". arXiv: 1609.07802 (cit. on pp. 1973, 1980, 1988, 1989).

Pablo Shmerkin and Ville Suomala (2017). *Spatially independent martingales, intersections and applications*. Memoirs of the Amer. Math. Soc. (Cit. on p. 1985).

Károly Simon (1996). "Overlapping cylinders: the size of a dynamically defined Cantor-set". In: *Ergodic theory of $\mathbf{Z}^d$ actions (Warwick, 1993–1994)*. Vol. 228. London Math. Soc. Lecture Note Ser. Cambridge: Cambridge Univ. Press, pp. 259–272. MR: 1411222 (cit. on p. 1971).

Boris Solomyak (1995). "On the random series $\sum \pm\lambda^n$ (an Erdős problem)". *Ann. of Math. (2)* 142.3, pp. 611–625. MR: 1356783 (cit. on p. 1980).

Terence Tao and Van Vu (2006). *Additive combinatorics*. Vol. 105. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, pp. xviii+512. MR: 2289012 (cit. on p. 1974).

V. N. Tutubalin (1967). "The central limit theorem for random motions of Euclidean space". *Vestnik Moskov. Univ. Ser. I Mat. Meh.* 22.6, pp. 100–108. MR: 0242233 (cit. on p. 1967).

Peter Varjú (2016). "Absolute continuity of Bernoulli convolutions for algebraic parameters". arXiv: 1602.00261 (cit. on pp. 1967, 1973, 1982).

Meng Wu (2016). "A proof of Furstenberg's conjecture on the intersections of $\times$p and $\times$q-invariant sets". arXiv: 1609.08053 (cit. on p. 1988).

Received 2017-11-30.

Michael Hochman
Einstein Institute of Mathematics, Edmond J. Safra Campus, The Hebrew University of Jerusalem, Givat Ram. Jerusalem, 9190401, Israel
mhochman@math.huji.ac.il

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (1991–2012)

RENORMALIZATION AND RIGIDITY

Konstantin Khanin

Abstract

The ideas of renormalization was introduced into dynamics around 40 years ago. By now renormalization is one of the most powerful tools in the asymptotic analysis of dynamical systems. In this article we discuss the main conceptual features of the renormalization approach, and present a selection of recent results. We also discuss open problems and formulate related conjectures.

1 Renormalization Group in Statistical Mechanics and Critical Phenomena

The ideas of renormalization and universality emerged in statistical mechanics in the 1960s in the works of L.Kadanoff, M.Fisher, A.Patashinski, V.Pokrovski, B.Widom, and K.Wilson in connection with a problem of phase transitions and critical phenomena. As parameters of physical system change the system may go through a dramatic change of its behaviour. This phenomenon is called phase transition. The simplest examples are provided by lattice spin systems. When temperature T decreases the system changes from a state of almost independent (statistically) spins to a highly correlated state. This transition happens at a particular value of temperature $T = T_{cr}$, called the critical temperature. It was discovered by physicists that at the critical temperature one can observe highly non-trivial scaling invariant structures. Moreover, this scaling invariant structures have strong universality properties. It means that their statistical properties and scalings exponents do not depend on the details of the interaction, but only on global characteristics, such as dimension, symmetries etc. Physicists also developed technical tools to calculate critical exponents and other parameters of the asymptotic statistical objects. These tools were based on the concept of renormalization group, or renormalization. The main idea is to look at large and increasing chunks of a system, but also rescale them in order to deal with objects of order 1. The transformation describing transition from one block size to

MSC2010: primary 37E20; secondary 37F25, 37E05, 37E10, 82B27, 82B28.

the next one, typically constant time larger, is called renormalization group transformation. The core of the renormalization group theory is a concept of the renormalization fixed point. It was realized that under the above procedure the system will converge to a statistical state which is a fixed point of the renormalization transformation. If the temperature $T > T_{cr}$ the renormalization converges to a trivial, or Gaussian, fixed point, which corresponds to the Central Limit Theorem (CLT) for statistically weakly depended spins. The system should be renormalized by a factor $N^{1/2}$ where N is a number of spins in a block. Thus the scaling exponent in this case is a trivial CLT exponent $\gamma = 2 \times 1/2 = 1$. Interesting behaviour can be observed only at a critical temperature $T = T_{cr}$ and at the temperatures infinitesimally close to it. It turns out that at the critical point the system under renormalization converges to a non-trivial statistical state which is also a fixed point for renormalization group transformation but with non-trivial critical scaling exponent γ. In the case of 2D Ising model $\gamma = 7/4$ which corresponds to a scaling $N^{7/8}$ instead of Gaussian scaling $N^{1/2}$.

Next step of a conceptual picture is a discussion of a stability of the above fixed points. One can linearize renormalization transformation near a fixed point and ask about the number of unstable directions. It turns out that there exists a single non-trivial unstable direction. According to the hyperbolic theory, it means that in the infinite-dimensional space of systems there exists a co-dimension 1 stable manifold of systems converging to a fixed point under renormalization. A one-parameter family of systems, parametrized by the temperature, will typically intersect this stable manifold. Such an intersection point corresponds to a critical temperature $T = T_{cr}$ at which the system converges to the critical renormalization fixed point. For temperatures $T > T_{cr}$ the system will converge to a trivial Gaussian fixed point which can be shown to be stable. For temperatures $T > T_{cr}$ and close to T_{cr} the system first comes close to a critical fixed point, and then goes away following a one-dimensional unstable manifold corresponding to the unique unstable direction.

This is a very brief sketch of the physical theory. I started with it to provide a historical background, and also to emphasize similarities with dynamical renormalization which I discuss below. I finish this physical detour with a few remarks. Firstly, the wording "renormalization group" is used less frequently in our days. In fact, one can speak of a semi-group of iterates of the renormalization transformation. In what follows I use only the word renormalization. Secondly, when I write above "it was discovered", "it was realized", "it turns out" etc, I mean statements established on the physical level. Mathematically rigorous results in this direction are very few, and they are incredibly difficult to prove. However, the conceptual picture described above is extremely simple and very attractive. It provided a completely different point of view, and played a crucial role in shaping of modern understanding of critical phenomena. Whenever physicists see a universal behaviour and universal numbers they immediately think about renormalization.

Thirdly, physical renormalization theory is much more sophisticated and elaborated than the sketch above. It also provides tools to calculate things. That is why the 1982 Nobel Prize in Physics was awarded to K.Wilson for his work on critical phenomena.

2 Dynamical Renormalization

In the context of the theory of dynamical systems renormalization was introduced by M.Feigenbaum and P.Coullet, Ch.Tresser in the middle of the 70s (Feigenbaum [1978] and Tresser and Coullet [1978]). Feigenbaum was playing around with two families of unimodal maps, the logistic family and the sine family. He looked at the sequences of the period-doubling bifurcations trying to find numerically the bifurcations parameter values. Feigenbaum observed that parameter values seems to converge exponentially fast to a limiting value. To help with numerics he estimated the rate of the exponential convergence. Astonishingly the rate looked the same for both families, $\mu_{n+1} - \mu_n \sim C\delta^{-n}$, where $\delta = 4.669\ldots$ is the famous Feigenbaum constant. As I pointed out above, physicists would look for a renormalization explanation when they come across universal numbers. Feigenbaum developed renormalization theory which explained the universality phenomenon. He defined a renormalization transformation and showed numerically that it has a non-trivial hyperbolic fixed point with essentially unique unstable direction. This unique unstable eigenvalue is exactly equal to δ. A large block of spins is replaced by an exponentially increasing sequence of iterates of a map. In the period-doubling case the n-th step of the renormalization procedure, or, in other words, n-th iterate of a renormalization transformation, corresponds to 2^n iterate of an original map. It turns out that interesting dynamics for 2^n iterate happens near the critical point, in fact, exponentially close to it. Thus it makes sense to rescale a space variable so that the effective dynamics is described in terms of order 1 maps. Next $(n+1)$-step requires iterating twice the rescaled n-th step map, as well as an additional rescaling. What is described in words is exactly period doubling transformation. Namely, consider the space of unimodal maps $f(x)$ with non-degenerate point of maximum at $x = 0$, and normalized in such a way that $f(0) = 1$. Then the renormalization transformation can be defined in the following way:

$$Rf(x) = -\alpha f(f(-\alpha^{-1}x)), \ \alpha^{-1} = -f(1). \tag{1}$$

The concept of metrical universality was so revolutionary and novel for mathematicians at the time that it was initially even difficult to make them believe in it. Shortly, however, mathematicians realized the importance of the discovery. O.Lanford (Lanford [1982]) gave the first rigorous computer-assisted proof of the existence of the Feigenbaum fixed point with one unstable direction. Later E.Vul, Ya.Sinai and K.Khanin developed the thermodynamic formalism for the Feigenbaum attractor (Vul, Sinai, and Khanin [1984]), important Epstein classes were introduced by H.Epstein(Epstein [1989]). However, a real mathematical theory, so-called Sullivan-McMullen-Lyubich theory was developed only in the 90s (Sullivan [1992], McMullen [1994], and Lyubich [1999]). In fact, the theory covers a much more general case. The accumulation points of period-doubling bifurcations corresponds to a particular combinatorics of the so-called infinitely renormalizable maps. For other combinatorial types the renormalization transformation can also be defined. On the other hand, in this more general situation it does not make sense to speak about fixed points of renormalization since the renormalization transformation itself changes at every step. However, it still make sense to speak about convergence of renormalization. Namely, consider two different analytic infinitely renormalizable unimodal maps T_1 and T_2 with the same order of critical points, and the same combinatorics. It means that at every step the same renormalization transformation is applied to both of them. Denote by $f_n^{(1)}$ and $f_n^{(2)}$ the n-th step renormalization for T_1 and T_2 respectively. It follows from the Sullivan-McMullen-Lyubich theory that $\| f_n^{(1)} - f_n^{(2)} \| \to \infty$ as $n \to \infty$ exponentially fast. At the same time, changing of the combinatorial type results in an exponential instability. One can show that there are no other unstable directions. In other words, the hyperbolicity of renormalization with one unstable direction is valid in full generality. The set of maps with a given combinatorial type form a stable manifold of co-dimension one. The fact that this set of maps has a smooth manifold structure follows from general fact of the hyperbolicity theory. This is a kind of dream result for the renormalization ideology. Indeed, it provides a full justification of the renormalization picture. The only drawback is the requirement that the order of critical points is given by even integer numbers. At the same time all the experts agree that similar result must hold for all orders greater than 1. While the analyticity assumption can be significantly relaxed, currently there are almost no results on convergence of renormalization for maps with critical points of non-integer order. In our opinion this is one of the central open problems in the theory of renormalization. We shall discuss this problem in more details later.

Our brief discussion of renormalization for unimodal maps would be incomplete without mentioning the paper by S. Davie on period-doubling for $C^{2+\epsilon}$ maps (Davie [1996]), another paper by M. Martens providing a construction of a large class of periodic point for renormalization transformation (Martens [1998]), and a recent important paper by A. Avila and M. Lyubich where a new approach to convergence of renormalization for unimodal maps is developed (Avila and Lyubich [2011]).

Many aspects of the renormalization theory can be presented in a cleaner and easier way in the case of circle homeomorphisms. This will be the main object in what we discuss below. The problem for interval maps is very similar. However circle maps have several advantages. Firstly, the combinatorial types are completely determined by the rotation numbers. Infinitely renormalizable are simply maps with irrational rotation number. But most importantly, exponential instability in changing of a combinatorial type is obvious in the circle case and is highly non-trivial fact for interval maps.

3 Renormalization and Rigidity for Circle Homeomorphisms

Let T be a homeomorphism of the unit circle $\mathbb{S}^1$ with irrational rotation number ρ. Although renormalization can be defined for any homeomorphism, a meaningful theory requires certain regularity. We shall either consider the case when T is a smooth diffeomorphism, or assume that T is smooth outside of a finite set of singular points. In this section we discuss the case of one singularity. It can be either a critical inflection point, for so-called critical circle maps, or a break point, that is a point where the first derivative has a jump discontinuity.

To implement a general renormalization scheme one has to determine a sequence of times such that an iterate of a initial point comes close to itself, and then rescale the space coordinate. Let the continued fraction expansion for ρ be given by $\rho = [k_1, k_2, \ldots, k_n, \ldots]$, where $k_i \in \mathbb{N}$, $i \in \mathbb{N}$ is a sequence of partial quotients. It is well known that denominators q_n of convergents $p_n/q_n = [k_1, k_2, \ldots, k_n]$ correspond to a sequence of times of closest returns for a linear rotation $T_\rho : x \mapsto x + \rho \pmod 1$. Moreover, $T_\rho^{q_{2n}} x_0$ converges to x_0 from the right, and $T_\rho^{q_{2n-1}} x_0$ from the left. To define a sequence of renormalization one has to fix a point x_0, called a marked point, about which the renormalization will be defined, then for n-level renormalization consider a closed interval I_n containing x_0 with the end points $x_{q_{n-1}} = T^{q_{n-1}} x_0$ and $x_{q_n} = T^{q_n} x_0$. For simplicity we assume that n is even, so $I_n = [x_{q_{n-1}}, x_{q_n}]$. The first return map from I_n into itself has two branches. The first one is given by $T^{q_n} : [x_{q_{n-1}}, x_0] \to [x_{q_n+q_{n-1}}, x_{q_n}]$, the second branch corresponds to $T^{q_{n-1}} : [x_0, x_{q_n}] \to [x_{q_{n-1}}, x_{q_n+q_{n-1}}]$. Here and below the trajectory of x_0 is denoted by $x_i = T^i x_0$, $i \in \mathbb{Z}$. Next we choose the n-th level renormalized coordinate, denoted by z, such that the length of I_n in the coordinate z will be order one. Usually z is defined by an affine change of variables such that $z(x_{q_{n-1}}) = -1, z(x_0) = 0$. Hence, $z(x) = (x - x_0)/(x_0 - x_{q_{n-1}})$. To simplify notations we do not indicate dependence of the coordinate z on the renormalization level n. Now we can define $R^n(T)$ as a pair of first return maps $T^{q_{n-1}}$, T^{q_n} expressed in the renormalized coordinate z. Denote $a_n = z(x_{q_n}), -b_n = z(x_{q_n+q_{n-1}})$. Then $R^n(T) = (f_n(z), g_n(z))$, $f_n(z) : [-1, 0] \to [-b_n, a_n]$, $g_n(z) : [0, a_n] \to [-1, -b_n]$, where

(2)
$$f_n(z) = \frac{T^{q_n}(x_0 + z(x_0 - x_{q_{n-1}})) - x_0}{x_0 - x_{q_{n-1}}}, \quad g_n(z) = \frac{T^{q_{n-1}}(x_0 + z(x_0 - x_{q_{n-1}})) - x_0}{x_0 - x_{q_{n-1}}}.$$

A simple but important observation is that in order to find maps (f_{n+1}, g_{n+1}) one does not need to know the original map T. Indeed, maps (f_{n+1}, g_{n+1}) are completely determined by the maps (f_n, g_n). Since $q_{n+1} = k_{n+1}q_n + q_{n-1}$, we have $f_{n+1} = A_{n+1}^{-1} \circ f_n^{k_{n+1}} \circ g_n \circ A_{n+1}$, $g_{n+1} = A_{n+1}^{-1} \circ f_n \circ A_{n+1}$, where A_{n+1} is a linear rescaling $A_{n+1}(z) = -a_n z$, and $a_n = f_n(0)$. An integer number k_{n+1} is also determined by f_n. It is just a number of iterates of f_n such that a trajectory of point (-1) reaches positive semi-axis. Namely, $f_n^i(-1) < 0, i \le k_{n+1}$ and $f_n^{k_{n+1}+1}(-1) > 0$. The transformation from a pair (f_n, g_n) to a pair (f_{n+1}, g_{n+1}) is called the *renormalization transformation*, and is denoted by R. Note that maps g_{n+1} are obtained from f_n by linearly rescaling the coordinates. Hence, in most cases, it is enough to keep track only of a sequence of maps f_n.

The main goal of the renormalization theory is to show that the renormalization transformation is hyperbolic, and circle maps with the same irrational rotations numbers and with the same local structure of their singular points belong to the same stable manifold for R. In other words, renormalizations $f_n^{(1)}$, $f_n^{(2)}$ constructed from two such maps T_1, T_2 converge to each other with exponential rate. Moreover, the rate of convergence is universal. It means that it does not depend on the rotation number and the maps T_1, T_2, but only on local characteristics of singular points. Circle maps which we consider below satisfy the Denjoy property. Namely, they all topologically conjugate to the linear rotation with the same rotation number. This allows us to formulate renormalization conjecture in a more precise way. We shall consider two type of singularities: critical points of finite order x_{cr} and break points x_{br}. At a critical point the derivative $T'(x_{cr}) = 0$, and a map T locally behaves like $T(x) - T(x_{cr}) \sim A(x - x_{cr})|x - x_{cr}|^{\alpha-1}$, where $A > 0$ and $\alpha > 1$ is the order of the critical point. At a break point x_{br} the first derivative of T has a jump discontinuity. Namely, both one sided derivatives exist and both are positive, but $T'(x_{br}-) \ne T'(x_{br}+)$. Parameter $c = \sqrt{T'(x_{br}-)/T'(x_{br}+)}$ is called the size of a break. We take square root in the formula for c to simplify some formulas below. When we say above "local characteristics of singular points" we mean precisely parameters α and c. Notice that both α and c are smooth invariants. In other words, they are preserved by smooth changes of variables.

We shall say that two circle maps T_1 and T_2 with the same irrational rotation number are *singularity equivalent* if there exists a topological conjugacy ϕ which is also a bijection between the singular points of T_1 and T_2. Moreover, conjugated singular points are of the same type (critical or break), and with the same value of parameters α or c.

Renormalization Conjecture. Suppose circle homeomorphisms T_1 and T_2 have a finite number of singular points and are singularity equivalent. Assume also that T_1 and T_2 are $C^{2+\alpha}$-smooth outside of a set of singularity points. Then the renormalizations $f_n^{(1)}$, $f_n^{(2)}$

constructed from the conjugated marked points converge to each other with a universal exponential rate.

Although convergence of renormalizations is interesting in its own right, it is also directly related to the rigidity theory. In many cases convergence of renormalization implies that two maps which a priori are only topologically conjugate are, in fact, smoothly conjugate to each other. Such an upgrade from topological equivalence to a smooth one is called the *rigidity*. Below we discuss rigidity results in parallel with results on renormalization convergence. In this section we consider only the unimodal setting, that is maps with a single singularity which is either a critical point or a break point. In this case it is natural to take the singularity point as a marked point for the renormalization construction. However, we start with the case of $C^{2+\alpha}$-smooth diffeomorphisms where the renormalization picture is much simpler.

Linearization of circle diffeomorphisms. Since in the smooth setting linear rotations form a distinguished class, rigidity in this case is usually discussed in terms of the linearization problem. The main question here is when a smooth circle diffeomorhism with irrational rotations number ρ is smoothly conjugate to a linear rotation T_ρ. This problem was first addressed in a classical paper by V.Arnold (Arnold [1961]). Arnold proved that smooth (in fact, analytic) linearization holds for analytic circle diffeomorphisms close to linear rotations, provided their rotation numbers are typical, i.e. satisfy certain conditions of a Diophantine type. He also showed that smooth linearization cannot be extended to all irrational rotation numbers, and conjectured that local condition, that is closeness to linear rotations, can be removed. Global result was proved by M.Herman (Herman [1979]) and extended to a larger class of rotation numbers by J-C.Yoccoz (Yoccoz [1984a]). Although renormalization was not explicitly used by Herman and Yoccoz, the linearization problem is closely related to the Renormalization Conjecture above. It is easy to see that for linear rotations T_ρ the renormalization $f_n(z) = a_n + z$, where $a_n = [k_{n+1}, k_{n+2}, \dots]$. Hence, convergence of renormalization reduces to showing that in a general case $\max_{z \in [-1,0]} |f_n'(z) - 1| \to 0$ as $n \to \infty$, and the convergence is fast enough. Since $f_n'(z) = (T^{q_n})'(x)$, the task is to prove that $\epsilon_n = \max_{x \in \mathbb{S}^1} |\log (T^{q_n})'(x)| \to 0$ as $n \to \infty$. This case is simpler than the case of maps with singularities since renormalization converge to a "trivial fixed-point" given by a family of linear maps $f_a(z) = z + a$. What is much harder is to get sharp estimates for ϵ_n which are needed for rigidity results. Indeed, even if we know that f_n is ϵ_n-close to a linear family, after many iterates one can lose control, and iterated maps may be not close to linear ones anymore. This may and will happen when k_{n+1} are very large. Convergence rate for ϵ_n is related to a growth rate of the denominators q_n. The following estimate was proved in Khanin and Teplinsky [2009].

Lemma 3.1. *Let $T \in C^{2+\alpha}(\mathbb{S}^1)$. Denote by $l_n = \max_{x \in \mathbb{S}^1} |T^{q_n} x - x|$. Then there exists a constant $C > 0$ such the following estimate holds:*

$$\epsilon_n \leq C \left[l_{n-1}^{\alpha} + \frac{l_n}{l_{n-1}} l_{n-2}^{\alpha} + \frac{l_n}{l_{n-2}} l_{n-3}^{\alpha} + \cdots + \frac{l_n}{l_0} \right]$$

It is easy to show that the sequence l_n decays exponentially fast, and, uniformly in n, the ratio l_n/l_{n-k} is exponentially small in k. It follows immediately that $\epsilon_n \to 0$ exponentially fast as $n \to \infty$. But, in fact, the lemma above gives much more. Its proof is quite elementary. It is based on renormalization ideology and cross-ratio distortion relations. Cross-ratio distortion methods were first used by Yoccoz (Yoccoz [1984b]) in his study of analytical critical circle maps. As we will see below it is a powerful tool in renormalization theory. The lemma above allows to prove the following sharp rigidity result. Denote by $D_\delta = \{\rho : |q\rho - p| \geq C(\rho) q^{-1-\delta},\ \forall p \in \mathbb{Z}, q \in \mathbb{N}\}$. It is well known that for any $\delta > 0$ the set D_δ has full Lebesgue measure.

Theorem 3.2. *(Sinai and Khanin [1989] and Khanin and Teplinsky [2009])Let $T \in C^{2+\alpha}(\mathbb{S}^1)$, $\rho(T) \in D_\delta$, and $\alpha > \delta$. Then T is $C^{1+\alpha-\delta}$-smoothly conjugate to T_ρ.*

It is well known that C^1-rigidity cannot hold for typical rotation numbers when the map T is only C^2-smooth. Hence $C^{2+\alpha}$ is a natural setting here. Note also that the smoothness result above is sharp. For maps T of higher smoothness the best result is due to Y. Katznelson and D.Ornstein (Katznelson and Ornstein [1989]).

Theorem 3.3. *Let $T \in C^{k+\alpha}(\mathbb{S}^1, k \geq 2$, $\rho(T) \in D_\delta$, and $k - 1 + \alpha - \delta > 1$. Then, for any arbitrary small ϵ, the map T is $C^{k-1+\alpha-\delta-\epsilon}$-smoothly conjugate to T_ρ.*

Note extra ϵ which has to be subtracted from the smoothness exponent which is not necessary in the case $k = 2$. I was recently informed by D.Ornstein that he can now improve the last result and remove ϵ from the estimate above (Ornstein [2017]). Finally, notice that C^1 linearization implies regularity of invariant measure for T. It is well known that any circle homeomorphism with irrational rotation number is uniquely ergodic. In other words, it has a unique invariant probability measure. This measure is absolutely continuous with respect to the Lebesgue measure with a positive continuous density if and only if the map T is C^1 linearizable.

I should also mention a closely related problem of simultaneous linearization of commuting circle diffeomorphisms. In this case one can speak about a higher rank action by the group $\mathbb{Z}^d$ where d is the number of diffeomorphisms. A conjecture by J. Moser stated that Herman theory can be extended to this setting in the following sense. Suppose a vector $\vec{\rho} = (\rho_1, \dots \rho_d)$, where $\rho_i, 1 \le i \le d$ are the rotation numbers of commuting C^∞ circle diffeomorphisms $T_1, \dots T_d$, is a Diophantine vector in the sense of simultaneous rational approximations. Then the diffeomorphisms $T_1, \dots T_d$ can be simultaneously linearized, and the corresponding conjugacy is C^∞ smooth. Moser proved a local version of the above conjecture. The global result was proved in Fayad and Khanin [2009].

Summarizing, we see that in the case of smooth circle diffeomorphisms the convergence of renormalization holds for all irrational rotation numbers. At the same time, the rigidity results require conditions of a Diophantine type.

Critical circle maps. Convergence of renormalization for critical circle maps was studied by E. de Faria, W. de Melo (de Faria and de Melo [1999, 2000] and M. Yampolsky (Yampolsky [2002] and Yampolsky [2001]). Although they used different approach, in both cases analysis was based on a combination of real-analytic methods and methods from the holomorphic dynamics. This required an assumption that the order of critical points is given by odd integer numbers: $\alpha = 3, 5, 7, \dots$. In de Faria and de Melo [1999] convergence of renormalization was proved in the C^∞ setting for rotation numbers of bounded type. Yampolsky developed a new approach based on parabolic renormalization which allowed him to prove exponential convergence of renormalizations in the analytic case C^ω for all irrational rotation numbers. Below I formulate the result in the analytic setting. Note that in the critical case we always choose a critical point as a marked point.

Consider a double-infinite sequence of natural numbers $\mathbf{k} = \{k_i, k_i \in \mathbb{N}, i \in \mathbb{Z}\}$, and form two irrational numbers $\rho_+ = [k_1, k_2, \dots, k_n \dots]$, $\rho_- = [k_0, k_{-1}, \dots, k_{-n} \dots]$. Denote by $\hat{G}$ a natural extension of a Gauss map $G : \rho \mapsto \rho^{-1} \pmod 1$, $\rho \in [0, 1]$. The transformation $\hat{G}$ acts on pairs (ρ_-, ρ_+) : $\hat{G}(\rho_-, \rho_+) = (([\rho_+{}^{-1}] + \rho_-)^{-1}, G\rho_+)$, where $[\cdot]$ denotes the integer part of a number. It is easy to see that $\hat{G}$ corresponds to the unit shift of a sequence $\mathbf{k}$. Denote by $\mathfrak{F}_\alpha$ the space of pairs of commuting analytic functions $(f(z), g(z)$ with a critical point of the order α at the origin. The following statements describing a hyperbolic horseshoe attractor for the renormalization transformation R follow from the results in Yampolsky [2002].

• Fix $\alpha = 2k + 1, k \in \mathbb{N}$. Then for every irrational $\rho_+ \in \mathbb{S}^1$ there exists a smooth co-dimension 1 manifold $\Gamma_s(\rho_+)$ which consists of pairs $(f(z), g(z)) \in \mathfrak{F}_\alpha$ with a "forward" rotation number ρ_+. Also, for every irrational $\rho_- \in \mathbb{S}^1$ there exists a smooth one-dimensional manifold $\Gamma_u(\rho_-)$ which consists of pairs $(f(z), g(z)) \in \mathfrak{F}_\alpha$ with a "backward" rotation number ρ_-. For every pair $(\rho_+, \rho_-) \in \mathbb{S}^1 \times \mathbb{S}^1$ the manifolds $\Gamma_s(\rho_+)$ and $\Gamma_u(\rho_-)$ intersect transversally at a unique point $(f_{\rho_-,\rho_+}(z), g_{\rho_-,\rho_+}(z))$.

• A set $A \subset \mathfrak{F}_\alpha$ which consists of all intersection points $(f_{\rho_-,\rho_+}(z), g_{\rho_-,\rho_+}(z))$ for different pairs (ρ_-, ρ_+) is a compact subset of $\mathfrak{F}_\alpha$ invariant for the renormalization transformation R :

$$R(f_{\rho_-,\rho_+}(z), g_{\rho_-,\rho_+}(z)) = (f_{\hat{G}(\rho_-,\rho_+)}(z), g_{\hat{G}(\rho_-,\rho_+)}(z)).$$

• The set A is a Cantor-type hyperbolic attractor for the renormalization transformation R. At every point $(f_{\rho_-,\rho_+}, g_{\rho_-,\rho_+}) \in A$ the global stable and unstable manifolds are precisely $\Gamma_s(\rho_+)$ and $\Gamma_u(\rho_-)$.

• Let T be an analytic critical circle map with a critical point of the order α and with an irrational rotation number ρ. Denote $\rho_+^{(n)} = G^n\rho = [k_{n+1}, k_{n+2}, \dots]$, $\rho_-^{(n)} = [k_n, k_{n-1}, \dots, k_1, 1, 1, \dots]$. Then $\|(f_n, g_n) - (f_{\rho_-^{(n)},\rho^{(n)}{}_+}, g_{\rho^{(n)}{}_-,\rho^{(n)}{}_+})\| \to 0$ exponentially fast as $n \to \infty$.

Note that filling in the tail of $\rho_-^{(n)}$ with 1s is not essential since the family $(f_{\rho_-,\rho_+}, g_{\rho_-,\rho_+})$ depends exponentially weakly on the tails of a sequence $\mathbf{k}$. One can say that all critical circle maps with the same irrational rotation number and the same order α belong to the same stable manifold for renormalization transformation R. The analyticity assumptions can and was substantially weakened. In Guarino and de Melo [2017] convergence of renormalizations has been proved in the C^4 setting. However the condition of odd integer orders of critical points is used in all existing results (apart from some perturbative results for α close to odd integers). At the same time there are no doubts that the renormalization conjecture must be true for all orders $\alpha > 1$. This is still an open problem of central importance.

We shall now discuss rigidity results for critical circle maps. We always consider a particular conjugacy which maps a critical point into another critical point. It turns out that critical maps are more rigid than diffeomorphisms. So-called *robust rigidity* result was proved in Khanin and Teplinsky [2007].

Theorem 3.4. *Let T_1 and T_2 be two analytical critical circle maps with the same order of critical points, and with the same irrational rotation number $\rho = \rho(T_1) = \rho(T_2)$. Then T_1 and T_2 are C^1-smoothly conjugate to each other.*

Note that C^1 rigidity holds for all irrational rotation numbers. Also note that the above result is sharp. A.Avila (Avila [2013]) proved that it cannot be extended even on a level of a modulus of continuity of the conjugacy between T_1 and T_2. In fact, the result in Khanin and Teplinsky [2007] is more general. It says that as long as convergence of renormalization is established, robust rigidity holds. In particular, in view of Guarino and de Melo [2017], it can be applied to C^4 critical circle maps with the same odd order of critical points.

To explain the mechanism of the robust rigidity we introduce a sequence of dynamical partitions which are closely related to renormalization, and discuss geometry of these partitions. Let x_0 be a marked point for renormalization. A partition $\xi_n(x_0)$ of the n-th level is a partition on $q_{n-1}+q_n$ intervals whose endpoints are given by a finite trajectory of a point $x_0 : \{x_i = T^i x_0, 0 \le i < q_{n-1}+q_n\}$. Denote $\Delta_0^{(n)}$ a closed interval with endpoints x_0 and x_{q_n}. It is easy to see that all the elements of the partition $\xi_n(x_0)$ belong to two sequences of closed intervals: $\Delta_i^{(n-1)} = T^i \Delta_0^{(n-1)}, 0 \le i < q_n$ and $\Delta_j^{(n)} = T^j \Delta_0^{(n)}, 0 \le j < q_{n-1}$. These two sequences do not intersect except at their endpoints, and cover the whole unit circle. It turn out that in the case of critical circle maps the geometry of partitions is uniformly bounded. Swiatek (Swiatek [1988]) showed that for any such T there exists a constant $C > 1$ such that for any two neighbouring intervals Δ_1, Δ_2 of a partition $\xi_n(x_0)$ the ratio of their lengths is bounded by C: $C^{-1} \le |\Delta_1|/|\Delta_2| \le C$. Moreover, a constant C is asymptotically universal. Namely, there exists a constant $C > 1$ which depends only on the order of a critical point such that the above estimate holds for all T if n is large enough. Note that in the diffeomorphism case the geometry is unbounded. Obviously $|\Delta_0^{(n)}|/|\Delta_0^{(n-1)}| \sim k_{n+1}^{-1}$ and can be arbitrary small for large k_{n+1}. Due to the bounded geometry, one can show that in the critical case when k_{n+1} is large renormalization $f_n(z)$ has a unique point of almost parabolic tangency with the diagonal. Iterations near such points of almost tangency are very regular. This fact together with convergence of renormalization is responsible for the robust rigidity.

For typical rotation numbers the conjugacy between two maps is $C^{1+\beta}$, $0 < \beta < 1$ smooth. It was first proved in de Faria and de Melo [2000] for rotation numbers of bounded type and later extended to the Lebesgue typical case. Although I don't know rigorous results in this direction, it is expected that in general the smoothness of conjugacy cannot be improved to C^2 or above even for rotation numbers of bounded type. This feature makes the critical case substantially different from the case of smooth diffeomorphism.

Concluding, one can prove the renormalization conjecture in full generality under a rather annoying condition on the order of critical points. The "fixed point family" given by the attractor $A = \{(f_{\rho_-,\rho_+}(z), g_{\rho_-,\rho_+}(z))\}$ is highly non-trivial. Smooth C^1 rigidity holds for all irrational rotation numbers.

Circle maps with breaks. Circle maps with breaks were introduced in Khanin and Vul [1991]. One possible motivation for their study can be explained if we adopt a point of view of generalized interval exchange transformations. It is well known that linear circle rotations can be viewed as interval exchange transformations of two intervals. Imagine now that the maps transforming two intervals are still smooth and monotone but are non-linear. The endpoints will be still matched. However a condition of matching of the derivatives of two branches of non-linear maps at the endpoints is rather unnatural. If the derivatives are not matched then we get two break points with break sizes c_1 and c_2. In fact, since both break points belong to a single trajectory, the renormalization for such maps is equivalent to renormalization for maps with one break point with a size of a break $c = c_1 c_2$.

The renormalization theory for maps with a single break point can be considered as a one-parameter extension of the Herman theory where parameter is a size of a break c. The value $c = 1$ corresponds to the diffeomorphisms case where, as we have seen above, renormalization converge to a trivial fixed-family consisting of linear maps with slope 1. In the break case renormalization converge to a space of Möbius transformations. The renormalization transformation has a non-trivial Cantor-type hyperbolic attractor similar to the critical case. Appearance of Möbius transformations is a conceptual fact. In the holomorphic dynamics this fact is related to the Köbe principle. Here we give a real-analytic explanation. Suppose we have a sequence of C^3-smooth one-dimensional maps $T_i, 0 \le i \le n-1$ acting on intervals $\Delta_i = [a_i, b_i], 0 \le i \le n$, $T_i : \Delta_i \to \Delta_{i+1}$. For each interval Δ_i define a relative coordinate $z_i(x) = (x - a_i)/(b_i - a_i)$. Assume that all intervals Δ_i are smaller than ϵ and $\sum_{i=0}^{n-1} |\Delta_i| = O(1)$. Assume also that all derivatives of T_i are uniformly bounded by a constant $C > 1$ and $T_i'(x) \ge C^{-1}$. Then the function $z_n(z_0)$ is order ϵ close to a fractional-linear function. The proof is based on a simple application of cross-ratio distortion estimates. If the maps T_i are only $C^{2+\alpha}$-smooth than closeness above will be of the order ϵ^α. In terms of renormalization this property implies that the renormalization $(f_n(z), g_n(z))$ are getting exponentially close to the space of pairs of fractional-linear functions as $n \to \infty$. It is also easy to see that asymptotically as $n \to \infty$ the limiting pairs $(f(z), g(z))$ must satisfy the commutativity relation $f \circ g(c^2 x) = g \circ f(x)$. Here there is a little twist. Since on every step of renormalization the orientation changes, c is also changing to c^{-1}, and back to c on the next step. Thus for odd n the commutativity relation is replaced by $f \circ g(c^{-2} x) = g \circ f(x)$. It is convenient to define $c_n = c^{(-1)^n}$. Taking into account the commutativity relations one can show that $(f_n(z), g_n(z))$ converges to invariant two-parameter family of pairs of fractional-linear functions which can be written explicitly. It is possible to write explicit formulas using geometrically define parameters a_n and b_n. However it is more convenient to replace b_n by another parameter $v_n = (c - a_n)/b_n - 1$ which characterizes the nonlinearity of the map $f_n(z)$. Let us define a family

$$F_{a,v,c}(z) = \frac{a + cz}{1 - vz}, \; G_{a,v,c}(z) = \frac{a(z - c)}{ac + z(1 + v - c)} \tag{3}$$

Then $(f_n(z), g_n(z))$ gets exponentially close to $F_{a_n,v_n,c_n}(z) = G_{a_n,v_n,c_n}(z)$ as $n \to \infty$. The limiting family (3) is invariant for the renormalization transformation R. The action of R on the parameter c is trivial: $Rc = 1/c$. The hyperbolic properties of R acting on the two-dimensional plane of parameters (a, v) were studied in Khanin and Khmelev [2003] and Khanin and Teplinsky [2013]. An important role is played by the special time-reversible symmetry. Define an involution $I(a, v, c) = ((c - 1 - v)/av, -v/c, 1/c)$.

Then, $R^{-1} = I \circ R \circ I$. Since R is essentially a two-dimensional transformation, the time-reversible symmetry provides duality between stable and unstable directions. The results proved in Khanin and Teplinsky [2013] are very similar to the convergence of renormalization results for critical circle maps. Again, there exist smooth stable and unstable manifolds $\Gamma_s(\rho_+)$ and $\Gamma_u(\rho_-)$, in this case they both are one-dimensional, parametrized by forward and backward rotation numbers ρ_+ and ρ_- respectively, and a hyperbolic attractor A which consists of all points of intersection between $\Gamma_s(\rho_+)$ and $\Gamma_u(\rho_-)$. As in the critical case, for any pair (ρ_+, ρ_-) the manifolds $\Gamma_s(\rho_+)$ and $\Gamma_u(\rho_)$ intersect transversally in a single point. All Lyapunov exponents are uniformly bounded away from zero by a positive constant which depends only on c. The main difference with the critical case is that the whole picture is now two-dimensional. Another, more conceptual, proof of the above statement can be found in Khanin and Yampolsky [2015]. It is based on the following idea. There is always one unstable direction for the renormalization transformation. This direction is related to a change of a rotation number. The other direction must be stable by the time-reversible symmetry. Together these two statements imply the hyperbolicity.

Although renormalization converge exponentially to the two-dimensional invariant family, and renormalization dynamics restricted to this family is hyperbolic and well understood, it is highly non-trivial to combine this two facts into the statement of global convergence of renormalization. The main difficulty is a strongly unbounded geometry when k_{n+1} are getting large. Assume that $c > 1$. Then for even n renormalization $f_n(z)$ will be convex, and for odd n it will concave. Thus, when k_{n+1} is large and n is even the function $f_n(z)$ has an almost parabolic tangency with the diagonal, like in the case of critical circle maps. This is a good case of bounded geometry. However, when n is odd, function $f_n(z)$ come close to the diagonal at a point $z = 0$ corresponding to the break point. One can show that in this case $f'(-1)$ is close to c, and $f'(0)$ is close to $1/c$. As a result, geometry is strongly unbounded. Namely, $|\Delta_0^{(n)}|/|\Delta_0^{(n-1)}| \sim c^{-k_{n+1}/2}$ which decays extremely fast as $k_{n+1} \to \infty$. Recall, that in the diffeomorphisms case $|\Delta_0^{(n)}|/|\Delta_0^{(n-1)}| \sim k_{n+1}^{-1}$. In the case $c < 1$ the situation is dual: the geometry is strongly unbounded for even n, and bounded for odd n. Despite the above difficulty the following global convergence result was proved in Khanin and Kocić [2014]. Consider two circle maps T_1 and T_2 with a break of size c. Assume that they are $C^{2+\alpha}$ smooth outside of break points. Denote $f_n^{(1)}(z)$, $f_n^{(2)}(z)$ renormalization for T_1 and T_2 respectively.

Theorem 3.5. *For any c there exist a constant $\lambda(c) \in (0,1)$ such that for all T_1 and T_2 as above and all n large enough we have $\| f_n^{(1)}(z) - f_n^{(2)}(z)\|_{C^2[-1,0]} \leq \lambda(c)^n$, provided $\rho(T_1) = \rho(T_2)$.*

Theorem 3.5 allows to prove the rigidity result for maps with breaks. Due to unbounded geometry robust rigidity does not hold in this case (Khanin and Kocić [2013]). One has to impose certain restrictions for a growth rate of k_{n+1} for odd n in case $c > 1$, and even n in the case $c < 1$. Assume that in both cases the corresponding sequence of k_{n+1} is bounded by λ_1^{-n} for large enough n of the right parity. Here $\lambda_1 < 1$ is an arbitrary constant greater than $\lambda(c)$. Then the following result holds (Khanin, Kocić, and Mazzeo [2017]).

Theorem 3.6. *Suppose $\rho = \rho(T_1) = \rho(T_2)$ satisfies the above condition. Then T_1 and T_2 are C^1-smoothly conjugate to each other.*

In this theorem we again assumed that $c(T_1) = c(T_2)$, and both T_1 and T_2 are $C^{2+\alpha}$ smooth outside of their break points. We have seen above that in the case of critical circle maps $C^{1+\beta}$ rigidity holds for Lebesgue almost all rotation numbers. S.Kocic showed that this is not the case for circle maps with breaks (Kocić [2016]). Of course $C^{1+\beta}$ rigidity still holds for rotation numbers of bounded type, or if k_{n+1} grow slowly enough.

4 Critical behaviour and parameter dependence

The renormalization for maps with singularities is very different from the diffeomorphism case. Using the statistical mechanics analogy we can say that it demonstrate the critical behaviour. Non-trivial scalings and fractal, or, more precisely, multifractal, structure of the renormalization attractor are just two of many manifestation of criticality.

Another feature is singularity of invariant measure. The singularity follows from Graczyk and Swiatek [1993] in the case of critical circle maps and from Dzhalilov and Khanin [1998] in the case of maps with a break point. Properties of invariant measure for maps with several breaks were studied in Dzhalilov, Liousse, and Mayer [2009].

It was known that if breaks belong to the same trajectory and the product of their sizes is equal to 1, then they can effectively compensate each other, and the invariant measure can be absolutely continuous. For a while it was conjectured that in all other cases the invariant measure is singular. Recently A. Teplinsky (Teplinsky [2018]) constructed interesting example of a piecewise linear circle maps with four breaks where the invariant measure is absolutely continuous, although the breaks do not compensate each other completely.

Another interesting property of critical behaviour is related to a parameter dependence. For any circle homeomorphism it is natural to include it in a one-parameter family which allows to change a rotation number. In can be done in different ways. Below we discuss the simplest construction when a one parameter family is given by $T_\omega(x) = T(x) + \omega \pmod 1$, $\omega \in [0, 1]$. An object of interest here is a function $\rho(\omega)$. Obviously, $\rho(\omega) = \omega$ in the linear case $T(x) = x$. In a nonlinear case $\rho(\omega)$ is a monotone, but not a strictly monotone, function. It has flat pieces $\omega \in I(p/q) = [a(p/q), b(p/q)]$ for all rational $0 \le \rho = p/q < 1$. These closed intervals $I(p/q)$ sometimes are called mode-locking intervals. On the contrary for any irrational $0 < \rho < 1$ there exists a unique value $\omega = \omega(\rho)$ corresponding to it. Denote by I_{ir} a Cantor-type of "irrational" parameter values, that is parameter values corresponding to irrational rotation numbers. It follows from the KAM theory that the set I_{ir} has a positive Lebesgue measure provided T is smooth enough. On the contrary, in the case of critical circle maps (Swiatek [1988]) and in the case of circle maps with a break point (Khanin and Vul [1991]) $\mathfrak{L}(I_{ir}) = 0$, where $\mathfrak{L}$ is the Lebesgue measure.

We next discuss how to define the notion of typical irrational rotation number in a one-parameter family. Since the parameter space is equipped with the Lebesgue measure it is natural to consider the conditional distribution on $\omega \in [0, 1]$ under condition that the rotation number is irrational. It is easy to define such a conditional distribution in the diffeomorphism case when $\mathfrak{L}(I_{ir}) > 0$. In the other two cases it is much less straightforward. A natural approach here is to consider a decreasing sequence of sets $I_n, n \in \mathbb{N}$ such that for $\omega \in I_n$ the continued fraction expansion for the rotation number $\rho(\omega)$ is longer than n. Obviously $\cap_n I_n = I_{ir}$. Since $\mathfrak{L}(I_n) > 0$ one can define the conditional distribution μ_n under the condition $\omega \in I_n$. Then the conditional distribution under the condition $\omega \in I_{ir}$ can be defined as $\mu(T) = \lim_{n\to\infty} \mu_n$. Using the hyperbolicity of renormalization one can proof that such a limit exists (Dolgopyat, Fayad, Khanin, and Kocic [2018]). Next one can consider the measure $\nu(T)$ which is a pushforward of $\mu(T)$ by the map $\omega \mapsto \rho(\omega)$. This measure provides a natural probability distribution on the irrational rotation numbers in a one-parameter family T_ω. Although measure $\nu(T)$ depends on a map T its asymptotic properties are universal. One way to see it is to consider pushforward of $\nu(T)$ by iterates of the Gauss map G. It turns out that in the case of critical circle maps $G^m\nu(T) \to \nu_\alpha$ as $m \to \infty$, where the limiting measure ν_α is already universal and depends only on the order of a critical point α (Dolgopyat, Fayad, Khanin, and Kocic [ibid.]). In the symbolic representation corresponding to the continued fraction expansion the measure ν_α has a Gibbs structure with a nice Hölder continuous potential. Note that in the diffeomorphism case the corresponding measure ν is absolutely continuous with the density $\frac{1}{\ln 2}\frac{1}{1+\rho}$. The statistical properties of measure ν are very different from ν_α. One can show that a probability with respect to ν_α of ρ such that an entry k_n in a continued fraction expansion for ρ takes a large value $k_n = k$ decays as k^{-3} while in the diffeomorphism case it decays as k^{-2}.

Similar statements can be proved for circle maps with a break point. The only difference is that in this case the sequence $G^m\nu(T)$ converges to a periodic orbit of the period two: $G^{2n}\nu(T) \to \nu_c^e$, $G^{2n+1}\nu(T) \to \nu_c^o$ as $n \to \infty$. This periodic orbit $G\nu_c^e = \nu_c^o$, $G\nu_c^o = \nu_c^e$ is also universal and depends only on a break size c. Obviously, $\nu_c^e = \nu_{1/c}^o$, $\nu_c^o = \nu_{1/c}^e$. In the case $c > 1$ a probability of large values of $k_n = k$ decays again as k^{-3} for odd n. For even n the probability decays exponentially with k. In the case $c < 1$ the parity is opposite. It looks plausible that for two families $T_{1,\omega}$, $T_{2,\omega}$ for typical "irrational" parameter values

ω_1, ω_2 the conjugacy between T_{1,ω_1} and T_{2,ω_2} will be $C^{1+\beta}$ smooth provided $\rho(T_{1,\omega_1}) = \rho(T_{2,\omega_2})$.

5 Beyond dimension one

All the results which we discussed above were one-dimensional. The combinatorics of trajectories which is present in the one-dimensional setting either in terms of rotation numbers, or in terms of kneading sequences is a crucially important feature in the universality phenomenon. This is especially true in the case of critical behaviour.

The multidimensional results are either related to the linear KAM regime (Mackay [1982], Koch [1999], and Khanin, Lopes Dias, and Marklof [2007, 2006]), or to the cases where one-dimensional structure is effectively present. By this I mean situations where dynamics is essentially one-dimensional with dissipation in other directions. This approach goes back to the paper by Collet, Eckmann, and Koch [1981]. Important results related to the geometry of embeddings of quasi one-dimensional attractors into two-dimensional space were studied by De Carvalho, Lyubich, and Martens [2005] and by Gaidashev and Yampolsky [2016].

Very important renormalization problems were studied in connection with critical invariant curves for area-preserving twist maps of a cylinder. An interesting and rich critical behaviour was discovered there by R.MacKay who also developed a renormalization scheme for such maps (Mackay [1982]). Unfortunately most of the results in this direction are still not rigorous. Recently Koch, using computer-assisted methods, was able to construct a corresponding fixed point for critical invariant curves with the golden mean rotation number $\rho = \frac{\sqrt{5}-1}{2}$ (Koch [2008]). Although dynamics in this case is really two-dimensional without dissipation, a footprint of the 1D case is still present through the invariant curve.

In the KAM regime renormalization schemes are based on the multidimensional continued fraction algorithm (Khanin, Lopes Dias, and Marklof [2007, 2006]). One can consider either a problem of constructing smooth invariant tori, or study the linearization problem for smooth diffeomorphisms of the torus $\mathbb{T}^d$. In both cases, in addition to the coordinate rescaling, on every step of the renormalization procedure one also have to implement certain nonlinear coordinate changes. These coordinate changes are eliminating some non-essential unstable direction. There are not essential precisely because they are related to coordinate changes. In fact, in the KAM regime all the eigen-directions, even stable ones, can be removed by means of proper coordinate changes. This fact can be seen as a renormalization explanation of a fast convergence of the linearization approximations in the KAM setting.

It is well known that KAM type of results on linearization in the multidimensional case require closeness of the corresponding maps to the linear ones. At the same time linearization results in the one-dimensional case are global. One of the main difficulties is related to Denjoy theory. In the 1D case we know when a diffeomorphism is topologically conjugated to a linear one, while in the multidimensional situation it is not the case. In certain sense this is the only obstacle to global results. The following conjecture by R.Krikorian states that the global rigidity results hold in any dimension.

Conjecture (R.Krikorian). Let T be a C^∞ diffeomorphism of $\mathbb{T}^d$. Assume that T is topologically conjugate to a linear translation $T_{\vec{\omega}} : \vec{x} \mapsto \vec{x} + \vec{\omega} \pmod 1)\, \vec{x} \in \mathbb{T}^d$ with a Diophantine rotation vector $\vec{\omega} = (\omega_1, \ldots, \omega_d)$. Then the conjugacy between T and $T_{\vec{\omega}}$ is C^∞ smooth.

It is obvious that in the presence of periodic orbits smooth linearization is virtually impossible. In the above conjecture there are no periodic orbits for T since it is conjugate to an irrational translation. Diophantine condition guarantees that there are also no orbits which are too close to periodic ones. The conjecture basically says that there are no other obstructions to smooth linearization.

It looks natural to propose the following generalization of the above conjecture. Let M be a compact Riemannian manifold. Let T_1 and T_2 be C^∞ diffeomorphisms of M. Assume also that T_1 and T_2 satisfy the Diophantine property. Namely, there exists $\tau, C > 0$ such that for $x \in M$ and all $n \in N$ we have: $dist(T_1^n x, x) \geq Cn^{-\tau}$, $dist(T_2^n x, x) \geq Cn^{-\tau}$.

Rigidity Conjecture. Suppose T_1 and T_2 are topologically conjugate. Then the conjugacy is C^∞ smooth.

Probably it is enough to require the Diophantine condition for only one of the maps T_1 and T_2.

6 Concluding remarks

We have seen above how the renormalization ideology can be implemented in several dynamical problems. Although the conceptual picture is very simple and appealing the proofs are quite difficult. A variety of different techniques is required in various settings. At the same time the phenomenon seems to be extremely general. In all the cases studied we see the hyperbolicity of renormalization, and convergence of renormalization for maps which are equivalent topologically and have the same structure of critical points.

The renormalization behaviour is extremely rigid. The geometrical properties of trajectories are completely determined by the backward and the forward rotation numbers and parameters (like α and c) characterizing the local behaviour near singular points.

Above we formulated a number of conjectures and discussed several open problem. As we have emphasized repeatedly, the central open problem in renormalization theory is the problem of convergence of renormalization for critical maps with an arbitrary order of critical points. Below we formulate few more open problems.

Interesting problems on convergence of renormalization arise in the multimodal setting when the map has several singular points. Consider for example the case of a map T_1 which has several break points u_i with break sizes c_i, $1 \leq i \leq k$. We shall present below an heuristic argument behind the hyperbolicity of renormalization. According to the Renormalization Conjecture, in order to have convergence of renormalization the second map T_2 should have the same rotation number and a matching structure of its break points. Denote v_i, $1 \leq i \leq k$ the break points for T_2. Matching means that their break sizes are also c_i, $1 \leq i \leq k$, and $\mu_1([u_i, u_{i+1}] = \mu_2([v_i, v_{i+1}]$, $1 \leq i < k$, where μ_1, μ_2 are probability invariant measures for T_1 and T_2 respectively. Changing the values of ρ and $m_i = \mu_1([u_i, u_{i+1}]$, $1 \leq i < k$ correspond to a k-dimensional unstable manifold. The task is to show that all other directions are stable. If there are no other unstable directions then by fixing the values of ρ and m_i, $1 \leq i < k$ we put T_1 and T_2 on the same stable manifold for renormalization transformation. In the case of k break points one has $(k+1)$ smooth branches of renormalization. Like in the case of one break point, all branches after rescaling will converge to the space of fractional-linear maps. Using k commutativity conditions one can reduce the number of parameters to $2k$. Renormalization transformation corresponds to some inducing scheme. It is possible to use the Rauzy induction. Indeed, at every step of renormalization the map can be viewed as a nonlinear interval exchange transformation of $(k+1)$ intervals. However one can develop a different inducing scheme which is more suitable for our purposes (Khanin and Teplinsky [2018]). This scheme corresponds to interchanges between k disjoint intervals, moreover, each of them is subdivided onto two subintervals. More precisely, each break point is surrounded by two intervals, one to the right and one to the left of it. These intervals are mapped by proper iterates of T. The union of their images covers the initial collection of intervals. On the next step of renormalization intervals get smaller and number of iterates get larger. An inductive scheme of the above type can be constructed in such a way that the renormalization transformation R will again have an explicit time-reversible symmetry provided by an involution I (Khanin and Teplinsky [ibid.]). We have seen above that such an involution exists in the case of one break. Of course the involution I is more complicated in the multiple breaks case. It acts on the subset of $\mathbb{R}^{2k}$ corresponding to parameters of fractional-linear functions. It also acts on a finite set of combinatorial types. In the case $k = 1$ it was just two combinatorial types related to a change of orientation.

Interchanges of k subdivided intervals have a more complicated combinatorial structure. Because of the time-reversible symmetry the number of stable and unstable directions are the same. This means that as long as the existence of k unstable directions is established, all other directions must be stable.

A simpler case of maps with k breaks under additional condition $\prod_{i=1}^{k} c_i = 1$ was considered in Cunha and Smania [2014, 2013]. In this case the renormalization converge to the space of linear maps with different slopes. Cunha and Smania used the Rauzy induction and proved convergence of renormalization and smooth rigidity. One can say that inductive schemes of the Rauzy type have three different representations: the usual one in the space of linear maps with slop 1, another one in the space of linear maps with different slopes studied in Cunha and Smania [2014, 2013], and the third one in the space of fractional-linear functions. Moreover, all three representations are hyperbolic.

Another interesting open problem is related to critical circle maps with asymmetric critical points. Namely, the orders α_1 to the right of the critical point and α_2 to the left of it are different. Since the maps in this case are not quasisymmetric the results of Yoccoz (Yoccoz [1984b]) cannot be applied. We even do not know whether such maps are topologically conjugate to linear rotations, although this should be expected. We believe that the renomalization for such maps will behave in the following way. The geometry of the dynamical partition ξ_n will be strongly unbounded. The intervals $\Delta_0^{(n)}$ and $\Delta_0^{(n-1)}$ will be exponentially small with different exponents. However, their images $\Delta_1^{(n)} = T(\Delta_0^{(n)})$ and $\Delta_1^{(n-1)} = T(\Delta_0^{(n-1)})$ will be already of the same order. In other words, a meaningful renormalization theory can be developed near the critical value, rather than near the critical point. One can expect that renormalization will still converge within the universality class.

This paper provides a relatively brief introduction into the theory of dynamical renormalization. It was certainly impossible to comment on all important contributions made in this area in the last 40 years. Many important papers were not discussed above. The reason for their omission is a lack of space rather than lack of respect.

References

V. I. Arnold (1961). “Small denominators. I. Mapping the circle onto itself”. *Izv. Akad. Nauk SSSR Ser. Mat.* 25, pp. 21–86. MR: 0140699 (cit. on p. 1997).

A. Avila (2013). “On rigidity of critical circle maps”. *Bull. Braz. Math. Soc. (N.S.)* 44.4, pp. 611–619. MR: 3167125 (cit. on p. 2000).

A. Avila and M. Lyubich (2011). “The full renormalization horseshoe for unimodal maps of higher degree: exponential contraction along hybrid classes”. *Publ. Math. Inst. Hautes Études Sci.* 114, pp. 171–223. MR: 2854860 (cit. on p. 1994).

Artur Avila and Raphaël Krikorian (2006). "Reducibility or nonuniform hyperbolicity for quasiperiodic Schrödinger cocycles". *Ann. of Math. (2)* 164.3, pp. 911–940. MR: 2259248.

P. Collet, J.-P. Eckmann, and H. Koch (1981). "Period doubling bifurcations for families of maps on $\mathbf{R}^n$". *J. Statist. Phys.* 25.1, pp. 1–14. MR: 610688 (cit. on p. 2006).

K. Cunha and D. Smania (2013). "Renormalization for piecewise smooth homeomorphisms on the circle". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 30.3, pp. 441–462. MR: 3061431 (cit. on p. 2009).

– (2014). "Rigidity for piecewise smooth homeomorphisms on the circle". *Adv. Math.* 250, pp. 193–226. MR: 3122166 (cit. on p. 2009).

A. M. Davie (1996). "Period doubling for $C^{2+\epsilon}$ mappings". *Comm. Math. Phys.* 176.2, pp. 261–272. MR: 1374413 (cit. on p. 1994).

A. De Carvalho, M. Lyubich, and M. Martens (2005). "Renormalization in the Hénon family. I. Universality but non-rigidity". *J. Stat. Phys.* 121.5-6, pp. 611–669. MR: 2192529 (cit. on p. 2006).

D. Dolgopyat, B. Fayad, K. Khanin, and S. Kocic (2018). *On typical irrational rotation numbers in one-parameter families of circle maps with singularities*. In preparation (cit. on p. 2005).

A. Dzhalilov and K. Khanin (1998). "On the invariant measure for homeomorphisms of the circle with a break". *Funktsional. Anal. i Prilozhen.* 32.3, pp. 11–21, 95. MR: 1659647 (cit. on p. 2004).

A. Dzhalilov, I. Liousse, and D. Mayer (2009). "Singular measures of piecewise smooth circle homeomorphisms with two break points". *Discrete Contin. Dyn. Syst.* 24.2, pp. 381–403. MR: 2486582 (cit. on p. 2004).

H. Epstein (1989). "Fixed points of composition operators. II". *Nonlinearity* 2.2, pp. 305–310. MR: 994095 (cit. on p. 1994).

E. de Faria and W. de Melo (1999). "Rigidity of critical circle mappings. I". *J. Eur. Math. Soc. (JEMS)* 1.4, pp. 339–392. MR: 1728375 (cit. on p. 1999).

– (2000). "Rigidity of critical circle mappings. II". *J. Amer. Math. Soc.* 13.2, pp. 343–370. MR: 1711394 (cit. on pp. 1999, 2001).

B. Fayad and K. Khanin (2009). "Smooth linearization of commuting circle diffeomorphisms". *Ann. of Math. (2)* 170.2, pp. 961–980. MR: 2552115 (cit. on p. 1999).

M. J. Feigenbaum (1978). "Quantitative universality for a class of nonlinear transformations". *J. Statist. Phys.* 19.1, pp. 25–52. MR: 0501179 (cit. on p. 1993).

D. Gaidashev and M. Yampolsky (Apr. 2016). "Renormalization of almost commuting pairs". arXiv: 1604.00719 (cit. on p. 2006).

J. Graczyk and G. Swiatek (1993). "Singular measures in circle dynamics". *Comm. Math. Phys.* 157.2, pp. 213–230. MR: 1244865 (cit. on p. 2004).

P. Guarino and W. de Melo (2017). "Rigidity of smooth critical circle maps". *J. Eur. Math. Soc. (JEMS)* 19.6, pp. 1729–1783. arXiv: 1511.02792. MR: 3646874 (cit. on p. 2000).

M.R. Herman (1979). "Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations". *Inst. Hautes Études Sci. Publ. Math.* 49, pp. 5–233. MR: 538680 (cit. on p. 1997).

Y. Katznelson and D. Ornstein (1989). "The differentiability of the conjugation of certain diffeomorphisms of the circle". *Ergodic Theory Dynam. Systems* 9.4, pp. 643–680. MR: 1036902 (cit. on p. 1998).

K. Khanin and D. Khmelev (2003). "Renormalizations and rigidity theory for circle homeomorphisms with singularities of the break type". *Comm. Math. Phys.* 235.1, pp. 69–124. MR: 1969721 (cit. on p. 2002).

K. Khanin and S. Kocić (2013). "Absence of robust rigidity for circle maps with breaks". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 30.3, pp. 385–399. MR: 3061428 (cit. on p. 2003).

– (2014). "Renormalization conjecture and rigidity theory for circle diffeomorphisms with breaks". *Geom. Funct. Anal.* 24.6, pp. 2002–2028. MR: 3283933 (cit. on p. 2003).

K. Khanin, S. Kocić, and E. Mazzeo (2017). "C^1-rigidity of circle maps with breaks for almost all rotation numbers". *Ann. Sci. Éc. Norm. Supér. (4)* 50.5, pp. 1163–1203. MR: 3720027 (cit. on p. 2004).

K. Khanin, J. Lopes Dias, and J. Marklof (2006). "Renormalization of multidimensional Hamiltonian flows". *Nonlinearity* 19.12, pp. 2727–2753. MR: 2273756 (cit. on p. 2006).

– (2007). "Multidimensional continued fractions, dynamical renormalization and KAM theory". *Comm. Math. Phys.* 270.1, pp. 197–231. MR: 2276445 (cit. on p. 2006).

K. Khanin and A. Teplinsky (2007). "Robust rigidity for circle diffeomorphisms with singularities". *Invent. Math.* 169.1, pp. 193–218. MR: 2308853 (cit. on p. 2000).

– (2009). "Herman's theory revisited". *Invent. Math.* 178.2, pp. 333–344. MR: 2545684 (cit. on p. 1998).

– (2013). "Renormalization horseshoe and rigidity for circle diffeomorphisms with breaks". *Comm. Math. Phys.* 320.2, pp. 347–377. MR: 3053764 (cit. on pp. 2002, 2003).

– (2018). *Renormalization and rigidity for circle maps with multiple breaks*. In preparation (cit. on p. 2008).

K. Khanin and E. Vul (1991). "Circle homeomorphisms with weak discontinuities". In: *Dynamical systems and statistical mechanics (Moscow, 1991)*. Vol. 3. Adv. Soviet Math. Amer. Math. Soc., Providence, RI, pp. 57–98. MR: 1118158 (cit. on pp. 2001, 2005).

K. Khanin and M. Yampolsky (2015). "Hyperbolicity of renormalization of circle maps with a break-type singularity". *Mosc. Math. J.* 15.1, pp. 107–121, 182. MR: 3427415 (cit. on p. 2003).

H. Koch (1999). "A renormalization group for Hamiltonians, with applications to KAM tori". *Ergodic Theory Dynam. Systems* 19.2, pp. 475–521. MR: 1685404 (cit. on p. 2006).

H. Koch (2008). "Existence of critical invariant tori". *Ergodic Theory Dynam. Systems* 28.6, pp. 1879–1894. MR: 2465604 (cit. on p. 2006).

S. Kocić (2016). "Generic rigidity for circle diffeomorphisms with breaks". *Comm. Math. Phys.* 344.2, pp. 427–445. MR: 3500244 (cit. on p. 2004).

O. E. Lanford III (1982). "A computer-assisted proof of the Feigenbaum conjectures". *Bull. Amer. Math. Soc. (N.S.)* 6.3, pp. 427–434. MR: 648529 (cit. on p. 1993).

M. Lyubich (1999). "Feigenbaum-Coullet-Tresser universality and Milnor's hairiness conjecture". *Ann. of Math. (2)* 149.2, pp. 319–420. MR: 1689333 (cit. on p. 1994).

R. S. Mackay (1982). *Renormalization in area preserving maps*. Thesis (Ph.D.)–Princeton University. ProQuest LLC, Ann Arbor, MI, p. 518. MR: 2632364 (cit. on p. 2006).

M. Martens (1998). "The periodic points of renormalization". *Ann. of Math. (2)* 147.3, pp. 543–584. MR: 1637651 (cit. on p. 1994).

C.T. McMullen (1994). *Complex dynamics and renormalization*. Vol. 135. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, pp. x+214. MR: 1312365 (cit. on p. 1994).

D. Ornstein (2017). Private communication (cit. on p. 1998).

Ya. Sinai and K. Khanin (1989). "Smoothness of conjugacies of diffeomorphisms of the circle with rotations". *Uspekhi Mat. Nauk* 44.1(265), pp. 57–82, 247. MR: 997684 (cit. on p. 1998).

D. Sullivan (1992). "Bounds, quadratic differentials, and renormalization conjectures". In: *American Mathematical Society centennial publications, Vol. II (Providence, RI, 1988)*. Amer. Math. Soc., Providence, RI, pp. 417–466. MR: 1184622 (cit. on p. 1994).

G. Swiatek (1988). "Rational rotation numbers for maps of the circle". *Comm. Math. Phys.* 119.1, pp. 109–128. MR: 968483 (cit. on pp. 2001, 2005).

A. Teplinsky (2018). "A circle diffeomorphism with breaks that is absolutely continuously linearizable". *Ergodic Theory Dynam. Systems* 38.1, pp. 371–383. MR: 3742551 (cit. on p. 2004).

Ch. Tresser and P. Coullet (1978). "Itérations d'endomorphismes et groupe de renormalisation". *C. R. Acad. Sci. Paris Sér. A-B* 287.7, A577–A580. MR: 512110 (cit. on p. 1993).

E. Vul, Ya. Sinai, and K. Khanin (1984). "Feigenbaum universality and thermodynamic formalism". *Uspekhi Mat. Nauk* 39.3(237), pp. 3–37. MR: 747790 (cit. on p. 1993).

M. Yampolsky (2002). "Hyperbolicity of renormalization of critical circle maps". *Publ. Math. Inst. Hautes Études Sci.* 96, 1–41 (2003). MR: 1985030 (cit. on p. 1999).

M.l Yampolsky (2001). "The attractor of renormalization and rigidity of towers of critical circle maps". *Comm. Math. Phys.* 218.3, pp. 537–568. MR: 1828852 (cit. on p. 1999).

J.-C. Yoccoz (1984a). "Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne". *Ann. Sci. École Norm. Sup. (4)* 17.3, pp. 333–359. MR: 777374 (cit. on p. 1997).

– (1984b). “Il n’y a pas de contre-exemple de Denjoy analytique”. *C. R. Acad. Sci. Paris Sér. I Math.* 298.7, pp. 141–144. MR: 741080 (cit. on pp. 1998, 2009).

Received 2017-12-18.

Konstantin Khanin
University of Toronto
khanin@math.toronto.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2013–2034)

BOUNDARY DYNAMICS FOR SURFACE HOMEOMORPHISMS

Andres Koropecki and Meysam Nassiri

Abstract

We discuss some aspects of the topological dynamics of surface homeomorphisms. In particular, we survey recent results about the dynamics on the boundary of invariant domains, its relationship with the induced dynamics in the prime ends compactification, and its applications in the area-preserving setting following our recent works with P. Le Calvez.

1 Introduction

The dynamics of homeomorphisms of the circle $\mathbb{S}^1$ has been completely understood and classified from the topological viewpoint since the early 20th century, with the classical Poincaré rotation number as a key concept. If $f : \mathbb{S}^1 \to \mathbb{S}^1$ is an orientation-preserving homeomorphism, the rotation number $\rho(f) \in \mathbb{R}/\mathbb{Z}$ measures the asymptotic average rotation of its orbits, and one has:

(i) If $\rho(f) = p/q \in \mathbb{Q}/\mathbb{Z}$ then every nonwandering point is a fixed point of f^q, and all periodic points have the same least period. In particular, the ω-limit and α-limit sets of any point consist of fixed points of f^q.

(ii) If $\rho(f) \notin \mathbb{Q}/\mathbb{Z}$ then there are no periodic points. Moreover, f is topologically semi-conjugate to the rigid rotation by $\rho(f)$, and it is uniquely ergodic. In particular, there is a unique minimal set. If f is sufficiently smooth, this semi-conjugation is a conjugation.

On the other hand, the jump from dimension 1 to dimension 2 introduces new behavior and rich dynamical phenomena which suggest that one cannot expect a general classification. For instance, one may have coexistence of periodic orbits of arbitrary periods, positive topological entropy, different coexistent types of rotational behavior, mixing, among several other phenomena that are not present in dimension 1.

A. K. acknowleges support by the German Research Council (Mercator fellowship, DFG-grant OE 538/9-1), as well as FAPERJ-Brasil and CNPq-Brasil.

MSC2010: primary 37E30; secondary 37B45, 37E45, 30D40.

Nevertheless, generalizations of the notion of rotation number have been useful in the study of two-dimensional dynamics. As in dimension one, the general idea is to measure the (asymptotic) average rotation of an orbit along a given homological direction of the surface, by means of a lift to the universal cover or an isotopy. It is observed that the coexistence of different types of rotational behavior often forces the existence of new types of orbits, periodic orbits, positive topological entropy, and other dynamical phenomena. The concept of rotation intervals for annulus homeomorphisms, motivated by the classical Poincaré-Birkhoff theorem, has been useful to obtain generalizations and refined versions of this result, in the context of twist maps Mather [1982a], area-preserving homeomorphisms, and even in the general setting Franks [1988] and Handel [1990]. The notion of rotation set of a toral homeomorphism had a similar success Franks [1989] and Llibre and MacKay [1991], and to a lesser extent its generalization to arbitrary surfaces of positive genus Pollicott [1992] and Koropecki and Tal [2017]. In late years, a great deal of progress has been made in understanding the dynamical consequences of rotation sets and intervals, e.g. Le Calvez and Tal [2017], Dávalos [2016], Jäger [2009], Koropecki and Tal [2014], Boyland, de Carvalho, and Hall [2016], Passeggi [2014], Addas-Zanata [2015], Kocsard [2016], and Koropecki, Passeggi, and Sambarino [2016].

A different approach consists in studying the dynamics of a surface homeomorphism at the boundary of an invariant domain. This is the focus of this article. To fix ideas, suppose U is a simply connected open set in an orientable surface S, invariant by an orientation-preserving homeomorphism f. Can one describe the dynamics of $f|_{\partial U}$ in simple terms, as in (i) and (ii) above? If ∂U is a simple curve, this is clearly the case. However, often ∂U is far from being a curve. An illustrative example is given by Wada-type continua, which are the common boundary of three or more disjoint topological disks. For instsnce the example in Figure 1 is nowhere locally connected and *indecomposable*, i.e. it is not the union of two proper nonempty subcontinua. This type of continuum appears frequently and robustly in smooth dynamics, for instance as a hyperbolic planar attractor Plykin [1974]. Another example, in a sense more drastic, is the *pseudo-circle* (see Figure 2), which is hereditarily indecomposable. This means every subcontinuum is indecomposable, and in particular it does not contain any arcs. The pseudo-circle may appear as an attractor Herman [1986] and Boroński and Oprocha [2015], as an invariant set of an area-preserving C^∞ diffeomorphism Handel [1982], or even as the boundary of a Siegel disk for a local holomorphic diffeomorphism Chéritat [2011]. Other relevant examples are the *hedgehogs* that appear in holomorphic dynamics Pérez-Marco [1997].

There is no hope for a classification such as (i)-(ii) in the general setting. Most of the rich dynamical dynamical phenomena that appear in dimension 2 can also appear simultaneously in the boundary of U. For instance, a connected hyperbolic attractor such as the Plykin attractor in the sphere includes dense periodic points of arbitrary periods, positive topological entropy, topological mixing, and yet it is the boundary of a simply connected

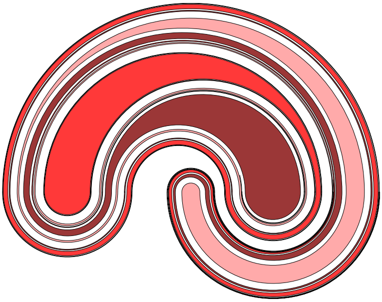


Figure 1: Wada-type continuum

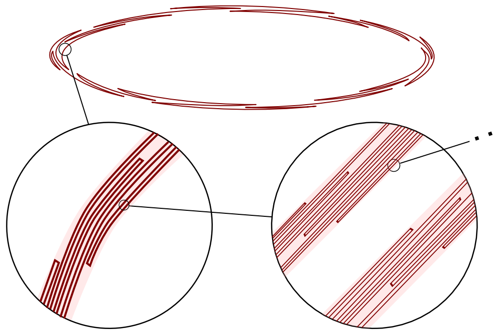


Figure 2: A pseudo-circle

domain. But it turns out that the presence of such rich dynamics in ∂U is related to the existence of attracting or repelling regions near ∂U, and if one introduces some mild form of recurrence near ∂U the situation changes completely. This is particularly the case in the area-preserving setting, which was the main motivation behind our joint works with P. Le Calvez described below.

The idea of studying the dynamics of an invariant continuum by means of its complementary components was already present in the work of Cartwright and Littlewood [1951] where they proved their celebrated fixed point theorem, and one of their key ideas was the use of Caratheodory's prime ends compactification and the prime ends rotation number. This was further explored by many authors (see for instance Mather [1981], Walker [1991], Barge and Gillette [1991], Alligood and Yorke [1992], Ortega and Ruiz del Portal [2011], and Hernández-Corbato, Ortega, and Ruiz del Portal [2012]). In the area-preserving setting, one of the motivations for this approach is the conjecture (dating back to Poincaré) that for a C^r-generic symplectic surface diffeomorphism the periodic points are dense. This is well known for $r = 1$ Pugh and Robinson [1983], but for $r > 1$ the usual local perturbation techniques do not work well, and a completely different approach is needed. Recent developments in symplectic topology led to a proof of the conjecture for any r in the case of Hamiltonian diffeomorphisms Asaoka and Irie [2016]. The general case remains open, but a relevant step is understanding the closures of invariant manifolds of hyperbolic periodic points. The following result, proved first in the sphere Franks and Le Calvez [2003] and later generalized to arbitrary surfaces Koropecki, Le Calvez, and Nassiri [2015] and Xia [2006] illustrates the usefulness of the topological study of boundaries of invariant domains (see Section 2.5):

Theorem 1.1. *For a C^r-generic area-preserving diffeomorphism of a closed surface, the set of all stable (or unstable) manifolds of hyperbolic periodic points is dense.*

Previous results by Mather [1981] showed that in the generic setting, the closures of any two stable or unstable branches of a hyperbolic periodic point coincide. This also relied on the use of prime ends.

If $U \subsetneq \mathbb{R}^2$ is an open simply connected set, the prime ends compactification $c_{\mathcal{E}}U$ is a way of compactifying U by adjoining a *circle of prime ends* $b_{\mathcal{E}}U \simeq \mathbb{S}^1$, in such a way that the resulting space is homeomorphic to the closed unit disk $\overline{\mathbb{D}}$ (see Section 2). The most direct approach to define it is to consider a conformal Riemann uniformization $\phi\colon U \to \mathbb{D}$, and then define the compactification as $U \sqcup \mathbb{S}^1$ with the topology generated by sets of the form $\phi^{-1}(V) \cup (V \cap \mathbb{S}^1)$ where V is (relatively) open in $\overline{\mathbb{D}}$. If $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a homeomorphism such that $f(U) = U$, then $f|_U$ always extends to a homeomorphism $f_e\colon c_{\mathcal{E}}U \to c_{\mathcal{E}}U$. The restriction of f_e to the circle of of prime ends is an orientation-preserving circle homeomorphism if f is orientation-preserving, and this allows us to define a *prime ends rotation number* $\rho(f, U)$ by considering the Poincaré rotation number of f_e on the circle of prime ends. One may then try to describe the dynamics in ∂U in terms of the this rotation number, and hope to recover results such as (i) and (ii) above. This transition from the prime ends to the boundary dynamics is a subtle problem, and there are examples showing that one cannot hope to do it without any additional hypotheses (see Figures 3 and 4).

However, when U has finite area and f is area-preserving, or more generally when there is some mild form of recurrence near ∂U, this approach has been more successful: an argument due to Cartwright and Littlewood [1951] shows that if $\rho(f, U) = p/q$ then there is a fixed point of f^q in ∂U (see Theorem 2.1 ahead). On the other hand, Mather showed that under certain generic conditions for an area-preserving diffeomorphism, $\rho(f, U)$ is always irrational Mather [1981] (although he did not obtain any direct consequences about the dynamics of $f|_{\partial U}$). Finally, in recent joint works Koropecki, Le Calvez, and Nassiri [2015, 2017], we obtained a more complete picture, which is very similar to the situation for homeomorphisms of the circle. The results apply under a more general local condition near ∂U which we will call the *boundary condition*. The precise definition is given in Section 2.2. We only mention here that this condition holds whenever $f|_U$ is nonwandering (in particular if f is area-preserving and U has finite area). Summarizing some of the results from Koropecki, Le Calvez, and Nassiri [2015, 2017] we may state the following:

Theorem 1.2. *Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is orientation-preserving and leaves invariant an open simply connected set $U \subsetneq \mathbb{R}^2$. Assume further that f has the boundary condition in U. Then:*

(1) *If $\rho(f, U) = p/q \in \mathbb{Q}$, then every nonwandering point of $f|_{\partial U}$ is a fixed point of f^q, and all periodic points have the same least period. In particular if U is bounded then the ω-limit and α-limit sets of any point in ∂U consist of fixed points of f^q.*

(2) *If $\rho(f, U) \notin \mathbb{Q}$ then there are no periodic points in ∂U. Moreover, if U is unbounded, there are no periodic points in the complement of U.*

The theorem is stated in $\mathbb{R}^2$ for simplicity, but similar results hold in more general surfaces. Note that part (1) is identical to what happens for circle homeomorphisms. Part (2) is only partially so, since examples such as Handel [1982] show that one cannot hope to obtain a semiconjugation to an irrational rotation, except in some special cases (see Section 3.4).

Note that this may be seen as a result about dynamics, without invoking prime ends. For instance, we may state the following:

Corollary 1.3. *Under the same hypotheses, if there is a fixed point of f in ∂U then the nonwandering points of $f|_{\partial U}$ are all fixed points.*

We remark that the claim about the nonwandering points of $f|_{\partial U}$ is very strong. If U is bounded, it implies that every orbit goes from the fixed point set to the fixed point set. This leads to the following Koropecki, Le Calvez, and Nassiri [2017]:

Theorem 1.4. *Suppose that $f : \mathbb{R}^2 \to \mathbb{R}^2$ is an area-preserving diffeomorphism, and U is a bounded f-invariant open topological disk with a fixed point in its boundary. Then:*

- *The derivative of f at every fixed point in ∂U has positive eigenvalues. In particular, there are no elliptic points in ∂U.*
- *If there is no fixed point in ∂U with an eigenvalue 1, then ∂U is the union of (finitely many) hyperbolic saddles together with saddle connections. In particular, ∂U is locally connected.*

The same thing holds on any surface if f is isotopic to the identity, with a possible exceptional case on the sphere (see Figure 5). A consequence of this fact is that under the explicit C^r-generic condition that f has no saddle connections and all periodic points are either hyperbolic or elliptic, there are no periodic points in ∂U. A similar result, with a completely different proof (still unpublished), was also announced by Fernando Oliveira several years ago.

As part of the proof of Theorem 1.2 we also obtain results about the topology of ∂U. Let us mention a particularly striking one. To simplify, assume that U is bounded, and ∂U is also the boundary of the unbounded connected component of $\mathbb{R}^2 \setminus \overline{U}$. In that case, if $\rho(f, U) = p/q$ and f has the boundary condition in U, we show that there are two possibilities: either $f^q|_{\partial U} = \mathrm{Id}$, or ∂U is *compactly generated* in the annulus $A = \mathbb{R}^2 \setminus \{z_0\}$ where $z_0 \in U$ is any point. This means that there is a compact connected set in the universal cover of A which projects onto ∂U, and it is a restrictive condition. For example, the pseudo-circle is not compactly generated, so we can state the following:

Theorem 1.5. *Suppose that f is area-preserving and K is an invariant pseudo-circle. If there is a fixed point in K, then $f|_K = \mathrm{Id}$.*

This is another instance of a general property that is seen in invariant boundaries in the area-preserving setting: *if the topology of the boundary is too wild in a certain way, then its dynamics is forced to be trivial.* The results on homotopical boundedness described in Section 2.7 also reflect this observation.

2 Dynamics under a boundary condition

Instead of defining the prime ends compactification by means of a uniformizing conformal map Milnor [2011] and Pommerenke [1992], one may use a purely topological definition, which is more convenient and flexible in many situations. We summarize it here; more details can be found, for instance, in Mather [1982b] or Koropecki, Le Calvez, and Nassiri [2015]. Most definitions in the literature assume that U is relatively compact, but following Koropecki, Le Calvez, and Nassiri [ibid.] we do not make this assumption.

Let S be a surface (which is always assumed to be boundaryless, connected, orientable, of finite genus and endowed with a metric) and $U \subset S$ an open topological disk such that $S \setminus U$ has more than one point. A *cross-cut* of U in S is the image of a simple arc $\gamma\colon (0,1) \to U$ that extends to an arc $\overline{\gamma}\colon [0,1] \to \overline{U}$ joining two points[1] of ∂U, and such that each of the two components of[2] $U \setminus \gamma$ has some boundary point in $\partial U \setminus \overline{\gamma}$. A *cross-section* of U in S is any connected component of $U \setminus \gamma$ for some cross-cut γ of U in S. Each cross-cut corresponds to exactly two cross-sections, which are topological disks.

A *chain* for U in S is a sequence $\mathcal{C} = (D_n)_{n\in\mathbb{N}}$ of cross-sections such that $D_i \subset D_j$ for all $i \geq j \geq 1$ and $\overline{\partial_U D_i} \cap \overline{\partial_U D_j} = \varnothing$ for all $i \neq j$. If D is any cross-section of U, we say that the chain $\mathcal{C}$ *divides* D if $D_i \subset D$ for some $i \in \mathbb{N}$. If $\mathcal{C}' = (D'_n)_{n\in\mathbb{N}}$ is another chain, we say that $\mathcal{C}$ *divides* $\mathcal{C}'$ if $\mathcal{C}$ divides D'_n for each $n \in \mathbb{N}$. We say that $\mathcal{C}$ and $\mathcal{C}'$ are equivalent if $\mathcal{C}$ divides $\mathcal{C}'$ and $\mathcal{C}'$ divides $\mathcal{C}$. A chain $\mathcal{C}$ is called a *prime chain* if $\mathcal{C}$ divides $\mathcal{C}'$ whenever $\mathcal{C}'$ is a chain that divides $\mathcal{C}$. An equivalence class of prime chains is called a *prime end* of U.

For a cross-section D of U, we say that the prime end $\mathfrak{p}$ *divides* D if some (hence any) chain representing p divides D. We denote by $\mathcal{E}_U D$ the set of all prime ends that divide D, and by $\mathrm{b}_{\mathcal{E}} U$ the set of all prime ends of U. The *prime ends compactification* of U is the set $\mathrm{c}_{\mathcal{E}} U = U \sqcup \mathrm{b}_{\mathcal{E}} U$, with the topology which has as a basis of open sets the family of all open subsets of U together with all sets of the form $D \cup \mathcal{E}_U D$ for some cross-section D of U. With this topology, $\mathrm{c}_{\mathcal{E}} U$ is homeomorphic to the closed unit disk $\overline{\mathbb{D}}$, and $\mathrm{b}_{\mathcal{E}} U$ endowed with the restricted topology is homeomorphic to the unit circle $\mathbb{S}^1$.

[1] Often it is assumed that the two points are different; we do not make this assumption.

[2] For simplicity we abuse notation and denote by γ both the parametrized arc and its image.

A useful observation is that when U is relatively compact in S, every prime end has a representative chain $(D_i)_{i\in\mathbb{N}}$ such that the diameters of the cross-cuts $\partial_U D_i$ tend to 0 as $i \to \infty$ (see Koropecki, Le Calvez, and Nassiri [ibid.]).

If $f : S \to S$ is an orientation preserving homeomorphism, then $f|_U$ extends to an orientation-preserving homeomorphism of the prime ends compactification, which we denote by $f_e : c_{\mathcal{E}}U \to c_{\mathcal{E}}U$. The *prime ends rotation number* of f in U, denoted $\rho(f, U) \in \mathbb{R}/\mathbb{Z}$, is defined as the Poincaré rotation number of the homeomorphism of the circle given by the restriction of f_e to $b_{\mathcal{E}}U \simeq \mathbb{S}^1$.

2.1 Prime ends vs. boundary dynamics. Suppose $U \subset S$ is an open topological disk invariant by the orientation-preserving homeomorphism f. What can be said about $f|_{\partial U}$ in terms of the prime ends rotation number? In particular, is it true that there are periodic points in ∂U if and only if $\rho(f, U)$ is rational? The answer in general is no, in both directions: The example in Figure 3 has no fixed points in ∂U, but the prime ends dynamics is a north pole-south pole map. On the other hand, the example in Figure 4 (see Walker [1991]) has fixed points in the boundary (every point in the outer circle is fixed), but the prime ends dynamics is a (Denjoy) homeomorphism with irrational rotation number.

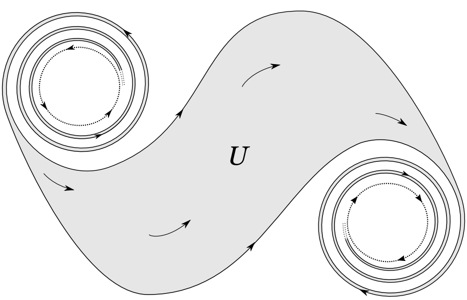


Figure 3: $\rho(f, U) = 0$, no fixed points

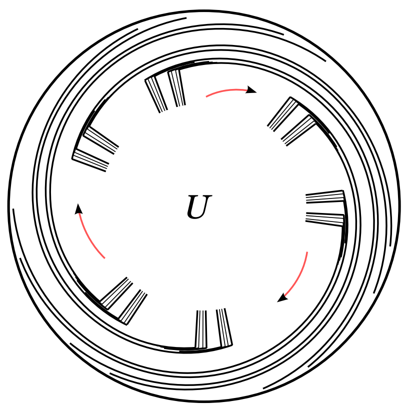


Figure 4: $\rho(f, U) \notin \mathbb{Q}/\mathbb{Z}$

Both examples have attracting or repelling regions near the boundary, and this is not a coincidence. If one excludes this behavior (for instance, if f is area-preserving and U has finite area), the situation changes. To illustrate this fact, we begin with a result proved by Cartwright and Littlewood. A cross-section D is *trapping* if it satisfies $f(D) \subset D$ and

$\overline{\partial_U D}$ is disjoint from its own image. Note that this also implies $\mathrm{cl}_U\, f(D) \subset D$. We also say that a cross-cut is trapping if one of its corresponding cross-sections is trapping.

Theorem 2.1. *If U is relatively compact and $\rho(f, U) = 0$ then either there is a fixed point in ∂U or there are trapping cross-cuts arbitrarily close to ∂U.*

Note that if f is area-preserving in U then one cannot have any trapping cross-cuts, so only the first case may hold.

Proof. Since $\rho(f, U) = 0$ there exists a prime end $\mathfrak{p}$ such that $f_e(\mathfrak{p}) = \mathfrak{p}$. Since U is relatively compact, one may choose a prime chain $(D_n)_{n\in\mathbb{N}}$ representing $\mathfrak{p}$ such that the diameters of the cross-cuts $\alpha_n = \partial_U D_n$ tend to 0 as $n \to \infty$. If infinitely many of these cross-cuts are trapping, then we have trapping cross-cuts arbitrarily close to $\partial_U U$. Thus we may assume that there exists i such that α_n is non-trapping for all $n \geq i$. The fact that $\mathfrak{p}$ is fixed means $(f(D_n))_{n\in\mathbb{N}}$ is also a prime chain representing $\mathfrak{p}$. Thus there is j_0 such that $D_j \subset f(D_i)$ for any $j > j_0$. We claim that if $j > j_1 := \min\{i, j_0\}$, then $f(\overline{\alpha}_j) \cap \overline{\alpha}_j \neq \varnothing$. Indeed, assume on the contrary that $f(\overline{\alpha}_j) \cap \overline{\alpha}_j = \varnothing$. Note that $D_j \subset f(D_i) \cap D_i$. By the previous argument using f^{-1} instead of f, there is $k > j$ such that $D_k \subset f^{-1}(D_j) \cap D_j$. In particular we have $D_k \cup f(D_k) \subset D_j \subset D_i \cap f(D_i)$. This means that both components of $U \setminus \alpha_j$ intersect their own image by f, and in particular letting D be the connected component of $U \setminus \alpha_j$ which does not contain $f(\alpha_j)$ one has $D \cap f(D) \neq \varnothing$. Since D is disjoint from $\partial_U f(D) = f(\alpha_j)$, this implies that $D \subset f(D)$. Note that $U \setminus D = \mathrm{cl}_U\, D'$, where D' is the cross-section determined by α_j which is not D; thus $f(\mathrm{cl}_U\, D') \subset \mathrm{cl}_U\, D'$, and since we assumed that $\overline{\alpha}_j$ is disjoint from $f(\overline{\alpha}_j)$, it follows that $f(\mathrm{cl}_U\, D') \subset D'$. This means that α_j is trapping, contradicting our assumption.

Thus $\overline{\alpha}_j \cap f(\overline{\alpha}_j) \neq \varnothing$ for all $j > j_1$, and we may choose $z_j \in \overline{\alpha}_j \cap f(\overline{\alpha}_j)$ for each $j \geq j_1$. Since the diameter of α_j tends to 0, we have $d(z_j, f^{-1}(z_j)) \to 0$ as $j \to \infty$, so any limit point of the sequence $(z_j)_{j\geq j_1}$ is a fixed point in ∂U. □

In fact, what the proof of the previous theorem shows is that if $\mathfrak{p}$ is a fixed prime end, then either one may find a prime chain for $\mathfrak{p}$ such that the corresponding cross-cuts are all trapping and their diameters tend to 0, or every *principal point* of the prime end $\mathfrak{p}$ is fixed (and this is particularly the case when f is area-preserving). A principal point is a point of ∂U that is the limit of a sequence of cross-cuts with diameter tending to 0 bounding the cross-sections of a prime chain of $\mathfrak{p}$. The set $\Pi(\mathfrak{p})$ of all principal points of $\mathfrak{p}$ is relevant because it can be characterized in the following alternative way. Consider the family $\mathfrak{F}$ of all arcs $\eta\colon [0,1) \to U$ such that $\lim_{t\to 1^-} \eta(t) = \mathfrak{p}$ in the topology of $c_{\mathcal{E}}U$. The same limit may fail to exist in the ambient space S (it will only exist if $\mathfrak{p}$ is accessible), but we may consider the limit set $L(\eta) = \bigcap_{0\leq t<1} \overline{\gamma([t,1))}$. The principal set is equal to

$\bigcap_{\eta\in\mathfrak{F}} L(\eta)$, and moreover it is realized as $L(\eta)$ for some η Mather [1982b]. In particular $\Pi(\mathfrak{p})$ is connected.

There have been other results regarding the realization of fixed points from fixed prime ends under different hypotheses, e.g. Ortega and Ruiz del Portal [2011] and Alligood and Yorke [1992] (see also Section 3). The converse of this result is more subtle, and was proved recently in Koropecki, Le Calvez, and Nassiri [2015]. In order to state it we need another definition.

2.2 The boundary condition. Note that a trapping cross-section can only occur if the prime ends rotation number vanishes. Here we introduce a natural generalization of the notion of trapping cross-section which poses no restrictions on the rotation number.

A (positive) *strong boundary trapping region* of U for f is an open set of the form $W = \bigcup_{D\in\mathfrak{F}} D$ with the following additional properties:

- $\mathfrak{F}$ is a family of pairwise disjoint cross-sections of U;
- The set $\{D \in \mathfrak{F} : \operatorname{diam}(\partial_U D) > c\}$ is finite for each $c > 0$;
- For each $D \in \mathfrak{F}$ there is $D' \in \mathfrak{F}$ such that $f(D) \subset D'$ and the cross-cuts α, α' bounding D and D' in U satisfy $f(\overline{\alpha}) \cap \overline{\alpha}' = \varnothing$.

These conditions in particular imply that $\mathrm{cl}_U\, f(W) \subset W$. The second condition always holds if $\partial_U W$ is contained in a compact arc. The last item guarantees that when one considers D and D' as subsets of the prime ends compactification $\mathrm{c}_{\mathcal{E}} U$, the closure of D is mapped into D'. For instance, a single trapping cross-section is a strong boundary trapping region.

We say that f has the *boundary condition* in U if there is a compact set $K \subset U$ such that, for each $n \in \mathbb{Z}$, $U \setminus K$ does not contain a set of the form $\partial_U W$ where W is a strong boundary trapping region of U for f^n.

The boundary condition is automatically satisfied if any of the following properties hold:

- f is area-preserving and U has finite area;
- f is nonwandering in U;
- there are no wandering cross-cuts of U;
- the dynamics induced in the circle of prime ends is transitive.

One of the reasons this condition is useful is that it is local (it can be verified by examining f in a neighborhood of ∂U), and therefore enables the use of our results in different settings by modifying f in a compact subset of U or outside of U. A particularly useful case is when one wants to extend results to non-simply connected open sets. Most of our results can be applied to isolated topological ends, and to do so one may use a surgery in the open set to “cut away” everything outside a collar neighborhood of the topological end and replace it by a fixed point.

Many results in Koropecki, Le Calvez, and Nassiri [2015] are stated under a slightly stronger condition called the ∂-*nonwandering condition*. However the only part of the article where this condition is crucial (Lemma 4.6 about maximal cross-cuts) remains valid if one assumes instead the boundary condition, and this is clear from the proof.

2.3 The irrational case. The proof of Theorem 1.2 requires a converse of Theorem 2.1. Assume $f: \mathbb{R}^2 \to \mathbb{R}^2$ is an orientation-preserving homeomorphism and $U = f(U) \subsetneq \mathbb{R}^2$ is an open topological disk.

Theorem 2.2 (Koropecki, Le Calvez, and Nassiri [ibid.]). *If f has the boundary condition in U and $\rho(f, U) \neq 0$ (mod $\mathbb{Z}$), then* $\operatorname{Fix}(f) \cap \partial U = \varnothing$. *If U is unbounded, then* $\operatorname{Fix}(f) \subset U$.

To prove this we rely on a general result on translation arcs. An n-*translation arc* for f is a simple arc α joining a point z to $f(z)$ such that the concatenation $\alpha * f(\alpha) * \cdots * f^n(\alpha)$ is also a simple arc. As typical example is when α is a fundamental domain of a stable or unstable branch of a hyperbolic saddle. Such α is an ∞-*translation arc* (i.e. it is an n-traslation arc for each $n \in \mathbb{N}$). A simplified version of the lemma on translation arcs Koropecki, Le Calvez, and Nassiri [ibid., Theorem E] states the following:

Lemma 2.3. *Under the hypotheses of the previous theorem, there exist $N \geq 1$ depending only on the rotation number $\rho = \rho(f, U)$ and a compact subset K of U such that any N-translation arc in $\mathbb{R}^2 \setminus K$ is disjoint from ∂U.*

The key is to prove this lemma is to compare the combinatorics of the orbits of cross-cuts in $\alpha \cap U$ seen in the cyclic order of the circle of prime ends $b_{\mathcal{E}} U$ (which corresponds to the combinatorics of the rigid rotation by $\rho(f, U)$ on the circle) with their respective positions in the linear order of the arc $\Gamma = \alpha * f(\alpha) * \cdots * f^N(\alpha)$. Using these two different approaches, if N is chosen sufficiently large, one is able to construct two simple loops which have nonzero algebraic intersection number, which is not possible in a surface of genus 0. In order to construct these simple loops, a fundamental step is to prove a "maximal cross-cut lemma" Koropecki, Le Calvez, and Nassiri [ibid., Lemma 4.6], which is the only part of Koropecki, Le Calvez, and Nassiri [ibid.] and Koropecki, Le Calvez, and Nassiri [2017] where the boundary condition plays a role (see Koropecki, Le Calvez, and Nassiri [2015, §4], Koropecki, Le Calvez, and Nassiri [2017], Koropecki, Le Calvez, and Tal [2017, Lemma 5]).

To illustrate the usefulness of Lemma 2.3, one may easily prove a particular case of Theorem 2.2, namely that *there is no hyperbolic (saddle) fixed point in ∂U*. Suppose that f is differentiable and p is a hyperbolic saddle. Because of the linear hyperbolic behavior near p, for any given N one may find an arbitrarily small neighborhood V of p such that

$V \setminus p$ is covered by N-translation arcs. If N is chosen as in Lemma 2.3, this means that V cannot intersect U and therefore p is not on the boundary of U.

The proof of Theorem 2.2 in the general case requires considerably more work. The main idea is to use a version of Brouwer's arc translation lemma to show that if there is a fixed point in ∂U then one may find arbitrarily close to this fixed point an N-translation arc intersecting ∂U, thus contradicting Lemma 2.3. But this cannot be done directly in $\mathbb{R}^2$; we need to work on a lift to the universal cover of $\mathbb{R}^2 \setminus X$ where X is a special set of fixed points. For the details we refer the reader to Koropecki, Le Calvez, and Nassiri [2015, §5]. We only mention here that an essential component of the proof is the existence of *maximally unlinked* sets of fixed points.

A closed set $X \subset \mathrm{Fix}(f)$ is *maximally unlinked* if it has the property that $f|_{\mathbb{R}^2 \setminus X}$ is isotopic to the identity in $\mathbb{R}^2 \setminus X$, and X is maximal among sets with this property with respect to inclusion. The existence of maximally unlinked sets was established by Jaulent [2014]. A stronger and more useful version of this result was recently obtained by Béguin, Crovisier, and Le Roux [2016]. These results about maximally unlinked sets in combination with the foliated version of Brouwer's plane translation theorem due to Le Calvez [2005] provide a powerful tool in two-dimensional dynamics, and led to many advances in recent years. We mention in particular the forcing theory of Le Calvez and Tal [2017], which produced several outstanding results in surface dynamics.

A version of Theorem 2.2 is also valid on an arbitrary surface, but one needs to make a special exception on the sphere, where it is easy to construct an area-preserving homeomorphism with an invariant disk U such that ∂U looks like a *hedgehog*: it has a single fixed point, with "hairs" which rotate around the fixed point with the combinatorics of an irrational rotation (see Figure 5 ahead). This kind of example has irrational prime ends rotation number but yet has a fixed point. It turns out that this is the only situation where a periodic point may exist if the rotation number is irrational:

Theorem 2.4 (Koropecki, Le Calvez, and Nassiri [2015]). *Let f be an orientation and area-preserving homeomorphism of a closed orientable surface S, and $U \subset S$ an open f-invariant topological disk whose complement has more than one point and such that $\rho(f, U) \notin \mathbb{Q}/\mathbb{Z}$. Then one of the following holds:*

- *(i) ∂U is an inessential annular continuum without periodic points;*
- *(ii) S is a sphere, U is dense in S, and $S \setminus U$ is a non-separating continuum with a unique fixed point and no other periodic points.*

Here by *inessential continuum* in S we mean a compact connected set $K \subset S$ which has a neighborhood D homeomorphic to a disk. By *annular* we mean that K is a decreasing intersection of closed topological annuli $A_1 \supset A_2 \supset \cdots$ such that A_{k+1} is essential in A_k for each k. Equivalently, this means that K has a neighborhood A homeomorphic to the annulus $\mathbb{A}$ such that K is essential in A and $A \setminus K$ has exactly two components. Even more general versions of Theorem 2.4 can be found in Koropecki, Le Calvez, and Nassiri [2015, §6-7].

Note how this result not only gives us dynamical information (no periodic points, with a single exception) but also topological information about the boundary. For instance, a disk as in Figure 6 cannot happen in Theorem 2.4 since its boundary is not annular.

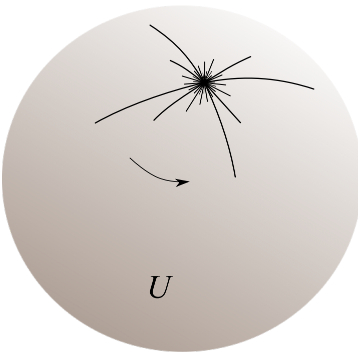


Figure 5: A hedgehog

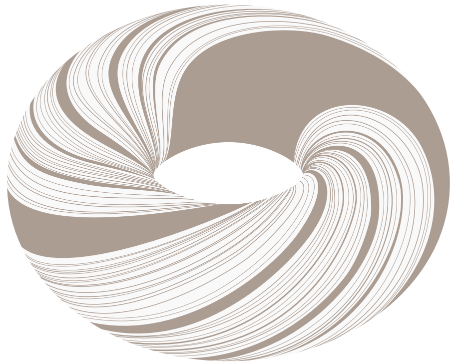


Figure 6: A homotopically unbounded disk

There is room for improvement in this case. We do not know much about the dynamics in ∂U when $\rho(f, U)$ is irrational and f is area-preserving, other than the fact that there are no periodic points (with the exception given in Theorem 2.4). For example, what type of minimal dynamics may appear in subsets of ∂U? Can one have more than one minimal set in ∂U when the map induced on the prime ends is minimal?

2.4 The rational case. We now consider part (i) of Theorem 1.2, which deals with the case when the rotation number is rational. This is contained in Koropecki, Le Calvez, and Nassiri [2017].

Instead of working in the plane, it is more convenient to work on the sphere, so we assume $f : \mathbb{S}^2 \to \mathbb{S}^2$ and $U \subset \mathbb{S}^2$ is an invariant open topological disk with more than one point in its complement. This is equivalent to working in the plane, since the Cartwright-Littlewood theorem Cartwright and Littlewood [1951] guarantees that there is a fixed point in $\mathbb{S}^2 \setminus U$, so by removing this point we are in a similar setting on the plane (and conversely, one may compactify $\mathbb{R}^2$ with one point).

Theorem 2.5 (Koropecki, Le Calvez, and Nassiri [2017]). *If f has the boundary condition in U and $\rho(f, U) = 0$, then the nonwandering set of $f|_{\partial U}$ is contained in* $\mathrm{Fix}(f)$.

The proof of this result has several components. We only mention that it relies on maximal isotopies Béguin, Crovisier, and Le Roux [2016], on a *real* prime ends rotation number associated to these isotopies, and on the study of the dynamics of a lift of f to the universal cover of $\mathbb{A} = \mathbb{S}^2 \setminus \{p, q\}$ where $p \in U$ and $q \notin U$ are fixed points of the maximal isotopy. If $\tilde{U}$ is a lift of $U \setminus \{p\}$ to this covering space, then we are able to show that every cross-cut of $\tilde{U}$ determined by a sufficiently small simple loop around a non-fixed point has endpoints located in a single fundamental domain of the *line of prime ends* (which can be thought as the universal cover of circle of prime ends of U). This essentially tells us that $\tilde{U}$ cannot have "tongues" that come close to two different non-fixed points in the same fiber of the covering map. Combining this with results from Brouwer-Le Calvez theory this leads to a contradiction if some non-fixed point of ∂U is nonwandering. But even without that assumption, this argument is what allows us to show that, unless ∂U has many fixed points, there are strong restrictions on the topology of ∂U. In particular, this is how we prove results such as Theorem 1.5.

A version of this result for arbitrary surfaces is also proved in Koropecki, Le Calvez, and Nassiri [2017], but only in the case where f is isotopic to the identity. The general case remains open and seems to be related to the problem of homotopical boundedness (see Section 2.7).

2.5 On C^r-generic area-preserving diffeomorphisms. Using surgery arguments, one may also obtain a version of Theorem 2.4 for an invariant *complementary domain* (i.e. a connected component of the complement of a continuum), which is not necessarily simply connected (see Koropecki, Le Calvez, and Nassiri [2015, §7]). This is particularly useful for C^r-generic area-preserving diffeomorphisms, since a variation of Mather's arguments Mather [1981] shows that rotation number associated to each topological end of an invariant complementary domain is irrational. This leads to the following (Koropecki, Le Calvez, and Nassiri [2015, Theorem B]):

Theorem 2.6. *If f is a C^r-generic area-preserving diffeomorphism ($r \geq 1$) and U is a periodic complementary domain, then there are no periodic points in ∂U. Moreover, ∂U is the union of finitely many pairwise disjoint annular continua.*

The last claim says that in a way $\overline{U}$ ressembles a surface with boundary. The generic condition required in this theorem can be given explicitly. The theorem holds whenever the following properties hold:

- every periodic point is either hyperbolic or elliptic, and there are no saddle connections;

- every neighborhood of an elliptic periodic point p contains a smaller neighborhood of p bounded by finitely many subarcs of the stable and unstable manifolds of some hyperbolic periodic point q, intersecting transversely.

Both conditions are C^r-generic for any $r \geq 1$ Robinson [1970] and Zehnder [1973].

Theorem 2.6 is a key part of the proof of Theorem 1.1. Let us briefly explain the idea behind that proof. For each area-preserivng diffeomorphism f we define the set K_f as the closure of the union of all stable manifolds of hyperbolic periodic points of f. Our aim is to prove that $K_f = S$ for a C^r-generic f.

Suppose that f satisfies the generic hypotheses listed above, and $K_f \neq S$. Let U be a connected component of $S \setminus K_f$, which must be periodic since $S \setminus K_f$ is invariant and f preserves area. Note that U cannot contain a periodic point, since it would then intersect a stable manifold of some hyperbolic periodic point (due to the generic assumptions). We claim that ∂U also contains no periodic points. Suppose instead that there is a periodic point $p \in \partial U$, and let C be the connected component of K_f containing p. Then the connected component U_0 of $S \setminus C$ containing U is a periodic complementary domain. By Theorem 2.6 there is no periodic point in ∂U_0. But clearly $p \in \partial U_0$, so this is a contradiction.

Thus $\overline{U}$ is aperiodic and compact. By the main theorem of Koropecki [2010] this means that either $\overline{U} = S = \mathbb{T}^2$ or $\overline{U}$ is an annular continuum. In the case that $S = \mathbb{T}^2$ it is known that a C^r-small perturbation creates a periodic point Addas-Zanata [2005, Corollary 2]. On the other hand if $\overline{U}$ is annular continuum, then it is easy to see that U must be homeomorphic to an annulus (otherwise it would contain a periodic point), and by a generalization of the Poincaré–Birkhoff theorem Franks [1988] one may find a C^r-small perturbation which creates a periodic point in U.

Using the lower-semicontinuity of the map $f \mapsto K_f$ in combination with these perturbative arguments one concludes that C^r-generically such a set U cannot exist. See Koropecki, Le Calvez, and Nassiri [2015, §8.5].

These results are useful for the study of dynamics of generic group actions on surfaces which still needs to be explored. For instance, one can prove existence of dense orbits for the semi-group generated by a pair of C^r generic area-preserving diffeomorphisms Koropecki and Nassiri [2010]. This is particularly useful as it has consequences on the instability problem of symplectic dynamics in higher dimensions Nassiri and Pujals [2012].

2.6 The smooth setting. Let us say a few words about Theorem 1.4. Note that the fact that there is a fixed point in ∂U implies that $\rho(f, U) = 0$ due to Theorem 2.2. Since U is bounded, by a simple argument we may reduce the problem to the analogous statement on the sphere with the additional assumption that $\rho(f, U) = 0$. Suppose that $f : \mathbb{S}^2 \to \mathbb{S}^2$

is an orientation-preserving and area-preserving C^1-diffeomorphism, and $U \subset \mathbb{S}^2$ is an f-invariant open simply connected set with $\rho(f, U) = 0$.

If $p \in \partial U$ is a fixed point of f, we may then blow-up p to a disk B, bounded by a circle C, where the dynamics corresponds to the map induced by $Df(p)$ on the unit circle by $v \mapsto Df(p)v/\|Df(p)v\|$. Denoting by g the new map obtained in this way, it still preserves a measure of full support on $\mathbb{S}^2 \setminus B$. Moreover, $\partial U \cap C$ is a nonempty compact invariant subset of C, so it should contain some recurrent point z. By Theorem 2.5, the point z must be a fixed point of $g|_C$. This means that $Df(p)$ has a 1-dimensional invariant subspace, with a positive eigenvalue. Since $\det(Df(p)) = 1$, both eigenvalues are positive as claimed in the first item of the theorem.

If f has no fixed point with eigenvalue 1 in ∂U then all fixed points in ∂U must be hyperbolic saddles, and there are finitely many of them, $p_1, \ldots, p_k$. Theorem 2.5 implies that every point of ∂U belongs to $W^s(p_i) \cap W^u(p_j)$ for some i, j. By a standard argument using the fact that f is area-preserving, if Γ is a stable or unstable branch of p_i (i.e. a connected component of $W^s(p_i) \setminus \{p_i\}$ or $W^u(p_i) \setminus \{p_i\}$) then $\Gamma \setminus \{p_i\}$ is either disjoint from ∂U or contained in ∂U. Thus ∂U is a union of stable and unstable branches of the points p_i. To prove that these branches are saddle connections, we show that if Γ^s, Γ^u are two branches in ∂U such that $\Gamma^s \cap \Gamma^u \neq \varnothing$ then $\Gamma^s = \Gamma^u$. Indeed, if this is not the case then there exists some simple loop γ bounded by the union of a compact subarc of Γ^s and a compact subarc of Γ^u. Since this loop is contained in ∂U, the set U must be in one of the connected components of $\mathbb{S}^2 \setminus \gamma$. If D denotes the remaining connected component, using the fact that f is area-preserving one may easily deduce that there exists $n > 0$ such that $f^n(\partial D) \cap D \neq \varnothing$, which means that $\partial U \cap D \neq \varnothing$ contradicting the fact that U is disjoint from D. Hence ∂U is a finite union of saddle connections as stated in Theorem 1.4.

We mention that Theorem 1.4 is potentially useful to study certain families of area-preserving maps, such as the standard family or conservative Hénon maps. These maps can be extended to bi-holomorphisms of $\mathbb{C}^2$, and it is known that whenever this happens saddle connections between periodic points cannot occur Ushiki [1980].

2.7 Homotopical boundedness. Consider an open topological disk $U \subset S$ where S is a closed orientable surface, invariant by a homeomorphism $f : S \to S$. The general idea that certain topological properties of ∂U force the presence of many fixed points has already appeared in Section 2.4. When S is an arbitrary surface, a general question inspired by the study of instability regions of area-preserving maps is the homotopical boundedness of U. If S is endowed with a metric of constant curvature we define the *covering diameter* $\mathfrak{D}(U) \in \mathbb{R}_+ \cup \{\infty\}$ as the diameter of any lift of U to the universal covering space of S. It is not difficult to produce examples where this number is infinite (for instance, see Figure 6). However, in the case that f is isotopic to the identiy and area-preserving (or more generally under a boundary condition) one can show that if $\mathfrak{D}(U) = \infty$ then the set of fixed points of f is *essential*.

Theorem 2.7 (Koropecki and Tal [2014, 2017]). *If f is area-preserving, isotopic to the identity, and its fixed point set is inessential, then there is a constant M independent of U such that $\mathfrak{D}(U) \leq M$.*

If f is not isotopic to the identity the result may fail to be true, but we conjecture that either $\mathfrak{D}(U) < \infty$ or $\mathrm{Fix}(f^n)$ is essential for some n (which is likely to depend only on the isotopy class of f). The case where f is in a pseudo-Anosov isotopy class seems to be easy to deal with (using a more direct argument), but the reducible case seems to need a new approach.

The proof Theorem 2.7 relied on Brouwer-Le Calvez theory, but recently a surprisingly simple proof was found using a "triple boundary lemma", which in its simplest form can be stated as follows:

Theorem 2.8 (Koropecki, Le Calvez, and Tal [2017]). *Suppose $f : \mathbb{S}^2 \to \mathbb{S}^2$ is orientation- and area-preserving and U_1, U_2, U_3 are pairwise disjoint open f-invariant topological disks. Then every $x \in \partial U_1 \cap \partial U_2 \cap \partial U_3$ is a fixed point.*

As an example of simple application of this theorem we state the following (which complements Theorem 1.5):

Theorem 2.9 (Koropecki, Le Calvez, and Tal [ibid.]). *If $K \subset \mathbb{R}^2$ is an invariant Wada-type continuum and f is area-preserving then $f^n|_K$ is the identity for some $n > 0$.*

We remark that the notion of homotopical boundedness can also be defined for non-simply connected open sets, and results on the same line as Theorem 2.7 are available Koropecki and Tal [2017].

3 Further results

3.1 Vanishing rotation numbers without fixed points. Consider a bounded open topological disk $U \subset \mathbb{R}^2$ invariant by some orientation-preserving homeomorphism f. We do not assume any additional condition on the dynamics. As we saw in Figure 3, even if $\rho(f, U) = 0$ it may be the case that $\mathrm{Fix}(f) \cap \partial U = \varnothing$. Theorem 2.1 tells us that this implies that there are boundary traps arbitrarily cose to ∂U; but this can be improved to the following statement, which is explicitly proved in Matsumoto and Nakayama [2011] but attributed to Cartwright-Littlewood.

Theorem 3.1. *Suppose there is no fixed point in ∂U and $\rho(f, U) = 0$. Then the dynamics induced by f on the circle of prime ends consists of alternating attracting and repelling fixed points. Moreover, each attracting fixed prime end is attracting in the disk $c_{\mathcal{E}} U$, and similarly for the repelling prime ends.*

What this tells us is that for the induced map f_e in $c_{\mathcal{E}} U$, a neighborhood of the circle of prime ends is covered by the basins of alternating attracting and repelling prime ends (see Figure 7). Although this is a beautiful result, it does not tell us anything about the dynamics of $f|_{\partial U}$. In Koropecki and Passeggi [2017], the authors obtained a translation to the boundary dynamics which tells us that this situation can only occur under very strict conditions. A simplified version of the main result states the following:

Theorem 3.2. *Suppose that $\rho(f, U) = 0$ and $\mathrm{Fix}(f) \cap \partial U = \varnothing$. Then there exists a finite pairwise disjoint family of rotational attractors and repellors (at least one of each) such that ∂U is contained in the union of their basins.*

By a *rotational attractor* we mean an invariant non-separating continuum A which is a topological attractor and has nonzero external rotation number (i.e. the prime ends rotation number of the disk $\mathbb{R}^2 \cup \{\infty\} \setminus A$ is nonzero).

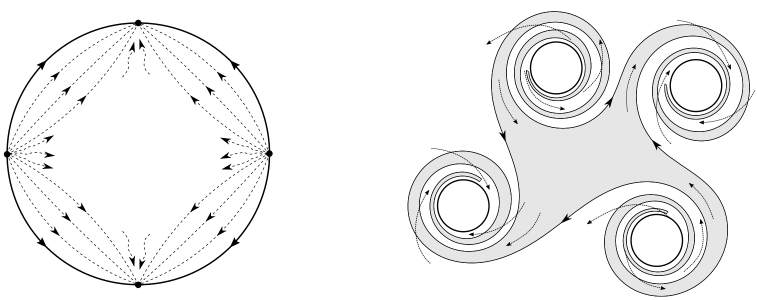


Figure 7: Prime ends and boundary dynamics as in Theorems 3.1 and 3.2

It is also shown that $f|_{\partial U}$ is topologically semiconjugate to a planar graph G where each vertex is an attractor or repellor and every edge is contained in the intersection of the basins of the corresponding vertices. Moreover, the semiconjugation extends to a monotone map from a neighborhood of G (contained in the union of the basins of the vertices) to a neighborhood of ∂U. We refer to Koropecki and Passeggi [ibid., §5.5] for further details. We only mention that a useful result introduced in that article, and essential to the proof of Theorem 3.2, is a Poincaré-Bendixson type theorem for translation lines Koropecki and Passeggi [ibid., Theorem A].

3.2 Rotation sets. As mentioned in the introduction, a common approach to generalize the notion of one-dimensional rotation number is to consider the rotation of orbits along homological directions. We consider the particularly simple case of the annulus $\mathbb{A} = \mathbb{T}^1 \times \mathbb{R}$. Suppose that $K \subset \mathbb{A}$ is an essential continuum and $f : \mathbb{A} \to \mathbb{A}$ is a homeomorphism isotopic to the identity. Given a lift $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}^2$ to the universal cover and $z \in K$, one defines its rotation number as

$$\rho(\tilde{f}, \tilde{z}) = \lim_{n \to \infty} (\tilde{f}^n(\tilde{z}) - \tilde{z})_1 / n,$$

for any $\tilde{z}$ in the fiber of z, where $(\cdot)_1$ denotes the first coordinate. This limit may fail to exist, and unlike in the circle, when it exists it generally depends on z. But considering the set of all such possible limits when they exist, one may define the (pointwise) *rotation set* $\rho(\tilde{f}, K) \subset \mathbb{R}$ associated to this lift. If K is a closed annulus, or more generally if K is *annular* (i.e. separates $\mathbb{A}$ into exactly two components), then the rotation set is compact and nonempty Handel [1990] and Koropecki [2016].

One may also define the rotation interval $\overline{\rho}(\tilde{f}, K) = [\inf \rho(\tilde{f}, K), \sup \rho(\tilde{f}, K)]$. It is natural then to ask whether having a rational element p/q in the rotation set or interval implies that one must have a corresponding periodic point of f of period q in K. The first result of this type is the classical Poincaré-Birkhoff theorem, which in a generalization due to Franks [1988] states that if f is area-preserving and K is a closed annulus then any rational element in its rotation interval is realized by a periodic point in K. Moreover, even without the area-preserving assumption, the result holds for elements of the rotation set Handel [1990] and Koropecki [2016].

If $K \subset \mathbb{A}$ is an essential continuum, we denote by $U_-(K)$ and $U_+(K)$ the two unbounded connected components of its complement (the latter being the one unbounded from above). Of particular interest are continua which are minimal with respect to the condition of being essential in $\mathbb{A}$. We call these continua *coboundaries*, and they are characterized by the property that $\partial U_-(K) = K = \partial U_+(K)$. Such a continuum may fail to be annular (as in Wada-type continua), but the continuum $K' = \mathbb{A} \setminus (U_-(K) \cup U_+(K))$, which can be thought as the “filling” of K, is annular and minimal with the property of being annular and essential. Any continuum with the latter property is called a *circloid*. Coboundaries and circloids are interesting because every essential continuum K contains a coboundary, and if K is invariant it also contains an invariant coboundary, which in turn has a corresponding invariant circloid.

One may then consider the (real) *upper* and *lower* prime ends rotation numbers $\rho_-(\tilde{f}, K)$ and $\rho_+(\tilde{f}, K)$ of K, which are defined in terms of the lift $\tilde{f}$ and lifts of the circles of prime ends of $U_\pm(K) \cup \{\pm\infty\}$. See Koropecki [2016] for more details. The relationship between these prime ends rotation numbers and the rotation set was studied in Matsumoto [2012] and Hernández-Corbato [2017], where it was proved that $\rho_\pm(\tilde{f}, K) \subset \overline{\rho}(\tilde{f}, K)$ when K is an annular continuum (see also Franks and Le Calvez [2003]).

3.3 No Birkhoff-like behavior for area-preserving maps. A circloid with empty interior is called a *cofrontier*. It is possible to produce an example of a cofrontier K where rotation interval has more than one point. For instance the Birkhoff attractor Le Calvez [1986] has a nontrivial interval as its rotation set (see also Boroński and Oprocha [2015]). When this happens, in general the dynamics in the cofrontier is rich; for instance there are infinitely many periodic points of arbitrarily large periods, uncountably many ergodic measures Koropecki [2016]. Moreover, if K is an attractor it implies that $f|_K$ has positive topological entropy Passeggi, Potrie, and Sambarino [2017]. It is conjectured that this is true even if K is not an attractor.

The results from Koropecki, Le Calvez, and Nassiri [2015] allow us to show that this kind of behavior is not possible in the area preserving setting: the rotation set is always a single point and coincides with the prime ends rotation numbers:

Theorem 3.3. *Suppose that f is area-preserving and $K \subset \mathbb{A}$ is an essential cofrontier. Then $\rho_-(\tilde{f}, K) = \rho_+(\tilde{f}, K)$ and this number is the only element of $\rho(\tilde{f}, K)$.*

The proof of this fact is explained in Koropecki [2016, Theorem 2.8] in a more general setting. This result also holds for arbitrary circloids (using the same argument combined with Koropecki, Le Calvez, and Tal [2017], for instance).

3.4 A Poincaré-like result for decomposable circloids. In the setting of the previous section, suppose that $K \subset \mathbb{A}$ is an f-invariant essential circloid whose boundary is *decomposable*, i.e. it can be written as the union of two proper subcontinua. The dynamics in a continuum of this type was studied in Jäger and Koropecki [2017], where the authors obtained a general Poincaré-type result without any additional hypothesis:

Theorem 3.4. *If K is an invariant essential circloid with decomposable boundary, then $\overline{\rho}(f, K)$ has a single element α, and*

- *α is rational if and only if there is a periodic point in K;*
- *α is irrational if and only if $f|_K$ is monotonically semiconjugate to the corresponding irrational rotation on the circle.*

Moreover, the semiconjugation in the last item is unique up to post-composition with a rotation.

The number α of course coincides with $\rho_-(f, K)$ and $\rho_+(f, K)$. This result emphasizes the fact that having points with different rotational behavior forces the topology of K to be complicated (which is a trait of indecomposable continua), something already noted in Barge and Gillette [1991].

Acknowledgments. This article was written during the stay of A. K. at the Institute of Mathematics - Friedrich Schiller University Jena, Germany, and he thanks the hospitality of the institution and his host Tobias Jäger.

References

S. Addas-Zanata (2005). "Some extensions of the Poincaré-Birkhoff theorem to the cylinder and a remark on mappings of the torus homotopic to Dehn twists". *Nonlinearity* 18.5, pp. 2243–2260. MR: 2164740 (cit. on p. 2026).

– (2015). "Uniform bounds for diffeomorphisms of the torus and a conjecture of Boyland". *J. Lond. Math. Soc. (2)* 91.2, pp. 537–553 (cit. on p. 2014).

K. T. Alligood and J. Yorke (1992). "Accessible saddles on fractal basin boundaries". *Ergod. Theory Dyn. Syst* 12.3, pp. 377–400 (cit. on pp. 2015, 2021).

M. Asaoka and K. Irie (2016). "A C^∞ closing lemma for Hamiltonian diffeomorphisms of closed surfaces". *Geom. Funct. Anal.* 26.5, pp. 1245–1254. MR: 3568031 (cit. on p. 2015).

M. Barge and R. M. Gillette (1991). "Rotation and periodicity in plane separating continua". *Ergod. Theory Dyn. Syst* 11.4, pp. 619–631 (cit. on pp. 2015, 2032).

F. Béguin, S. Crovisier, and F. Le Roux (Oct. 2016). "Fixed point sets of isotopies on surfaces". arXiv: 1610.00686 (cit. on pp. 2023, 2025).

J. Boroński and P. Oprocha (2015). "Rotational chaos and strange attractors on the 2-torus". *Math. Z.* 279.3-4, pp. 689–702 (cit. on pp. 2014, 2031).

P. Boyland, A. de Carvalho, and T. Hall (2016). "New rotation sets in a family of torus homeomorphisms". *Invent. Math.* 204.3, pp. 895–937. MR: 3502068 (cit. on p. 2014).

M. L. Cartwright and J. E. Littlewood (1951). "Some fixed point theorems". *Ann. of Math. (2)* 54. With appendix by H. D. Ursell, pp. 1–37. MR: 0042690 (cit. on pp. 2015, 2016, 2024).

A. Chéritat (2011). "Relatively compact Siegel disks with non-locally connected boundaries". *Math. Ann.* 349.3, pp. 529–542 (cit. on p. 2014).

P. Dávalos (Aug. 2016). "On annular maps of the torus and sublinear diffusion". *J. Inst. Math. Jussieu* FirstView, pp. 1–66 (cit. on p. 2014).

J. Franks (1988). "Generalizations of the Poincaré-Birkhoff theorem". *Ann. of Math. (2)* 128.1, pp. 139–151. MR: 951509 (cit. on pp. 2014, 2026, 2030).

– (1989). "Realizing rotation vectors for torus homeomorphisms". *Trans. Amer. Math. Soc.* 311.1, pp. 107–115. MR: 958891 (cit. on p. 2014).

J. Franks and P. Le Calvez (2003). “Regions of instability for non-twist maps”. *Ergod. Theory Dyn. Syst* 23.1, pp. 111–141 (cit. on pp. 2015, 2031).

M. Handel (1982). “A pathological area preserving C^∞ diffeomorphism of the plane”. *Proc. Amer. Math. Soc.* 86.1, pp. 163–168 (cit. on pp. 2014, 2017).

– (1990). “The rotation set of a homeomorphism of the annulus is closed”. *Comm. Math. Phys.* 127.2, pp. 339–349. MR: 1037109 (cit. on pp. 2014, 2030).

M. R. Herman (1986). “Construction of some curious diffeomorphisms of the Riemann sphere”. *J. London Math. Soc.* 2.2, pp. 375–384 (cit. on p. 2014).

L. Hernández-Corbato (2017). “An elementary proof of a theorem by Matsumoto”. *Nagoya Math. J.* 227, pp. 77–85 (cit. on p. 2030).

L. Hernández-Corbato, R. Ortega, and F. Ruiz del Portal (2012). “Attractors with irrational rotation number”. *Math. Proc. Camb. Phil. Soc.* 153.1, pp. 59–77 (cit. on p. 2015).

T. Jäger (2009). “Linearization of conservative toral homeomorphisms”. *Invent. Math.* 176.3, pp. 601–616. MR: 2501297 (cit. on p. 2014).

T. Jäger and A. Koropecki (2017). “Poincaré theory for decomposable cofrontiers”. *Ann. Henri Poincaré* 18.1, pp. 85–112. MR: 3592091 (cit. on p. 2031).

O. Jaulent (2014). “Existence d’un feuilletage positivement transverse à un homéomorphisme de surface”. *Ann. Inst. Fourier* 64.4, pp. 1441–1476 (cit. on p. 2023).

A. Kocsard (Nov. 2016). “On the dynamics of minimal homeomorphisms of $\mathbb{T}^2$ which are not pseudo-rotations”. arXiv: 1611.03784 (cit. on p. 2014).

A. Koropecki (2010). “Aperiodic invariant continua for surface homeomorphisms”. *Math. Z.* 266.1, pp. 229–236. MR: 2670681 (cit. on p. 2026).

– (2016). “Realizing rotation numbers on annular continua”. *Math. Z.* Pp. 1–16 (cit. on pp. 2030, 2031).

A. Koropecki, P. Le Calvez, and M. Nassiri (2015). “Prime ends rotation numbers and periodic points”. *Duke Math. J.* 164.3, pp. 403–472. MR: 3314477 (cit. on pp. 2015, 2016, 2018, 2019, 2021–2026, 2031).

– (2017). “Prime ends and boundary dynamics” (cit. on pp. 2016, 2017, 2022, 2024, 2025).

A. Koropecki, P. Le Calvez, and F. A. Tal (Nov. 2017). “A triple boundary lemma for surface homeomorphisms”. arXiv: 1711.00920 (cit. on pp. 2022, 2028, 2031).

A. Koropecki and M. Nassiri (2010). “Transitivity of generic semigroups of area-preserving surface diffeomorphisms”. *Math. Z.* 266.3, pp. 707–718. MR: 2719428 (cit. on p. 2026).

A. Koropecki and A. Passeggi (2017). “A Poincaré-Bendixson theorem for translation lines and applications to prime ends”. arXiv: 1701.04644 (cit. on p. 2029).

A. Koropecki, A. Passeggi, and M. Sambarino (Nov. 2016). “The Franks-Misiurewicz conjecture for extensions of irrational rotations”. arXiv: 1611.05498 (cit. on p. 2014).

A. Koropecki and F. A. Tal (2014). "Strictly toral dynamics". *Invent. Math.* 196.2, pp. 339–381. MR: 3193751 (cit. on pp. 2014, 2028).

– (2017). "Fully essential dynamics for area-preserving surface homeomorphisms". *Ergod. Theory Dyn. Syst*, pp. 1–46 (cit. on pp. 2014, 2028).

P. Le Calvez (1986). "Existence d'orbites quasi-periodiques dans les attracteurs de Birkhoff". *Comm. Math. Phys.* 106.3, pp. 383–394 (cit. on p. 2031).

– (2005). "Une version feuilletée équivariante du théorème de translation de Brouwer". *Publ. Math. Inst. Hautes Études Sci.* 102, pp. 1–98. MR: 2217051 (cit. on p. 2023).

P. Le Calvez and F. A. Tal (2017). "Forcing theory for transverse trajectories of surface homeomorphisms". To appear in Invent. Math. arXiv: 1503.09127 (cit. on pp. 2014, 2023).

J. Llibre and R. S. MacKay (1991). "Rotation vectors and entropy for homeomorphisms of the torus isotopic to the identity". *Ergod. Theory Dyn. Syst* 11.1, pp. 115–128. MR: 1101087 (cit. on p. 2014).

J. N. Mather (1981). "Invariant subsets for area preserving homeomorphisms of surfaces". In: *Mathematical analysis and applications, Part B*. Vol. 7. Adv. in Math. Suppl. Stud. New York: Academic Press, pp. 531–562. MR: 634258 (cit. on pp. 2015, 2016, 2025).

– (1982a). "Existence of quasi-periodic orbits for twist homeomorphisms of the annulus". *Topology* 21.4, pp. 457–467 (cit. on p. 2014).

– (1982b). "Topological proofs of some purely topological consequences of Carathéodory's theory of prime ends". In: *Selected studies: physics-astrophysics, mathematics, history of science*. Amsterdam: North-Holland, pp. 225–255. MR: 662863 (cit. on pp. 2018, 2021).

Sh. Matsumoto (2012). "Prime end rotation numbers of invariant separating continua of annular homeomorphisms". *Proc. Amer. Math. Soc.* 140.3, pp. 839–845. MR: 2869068 (cit. on p. 2030).

Sh. Matsumoto and H. Nakayama (2011). "Continua as minimal sets of homeomorphisms of S^2". *Enseign. Math. (2)* 57.3-4, pp. 373–392. MR: 2920733 (cit. on p. 2028).

J. Milnor (2011). *Dynamics in One Complex Variable*. Annals of Mathematics Studies 160, Third Edition. Princeton University Press (cit. on p. 2018).

M. Nassiri and E. R. Pujals (2012). "Robust transitivity in Hamiltonian dynamics". *Ann. Sci. Éc. Norm. Supér. (4)* 45.2, pp. 191–239. MR: 2977619 (cit. on p. 2026).

R. Ortega and F. Ruiz del Portal (2011). "Attractors with vanishing rotation number". *J. Eur. Math. Soc.* 13.6, pp. 1569–1590 (cit. on pp. 2015, 2021).

A. Passeggi (2014). "Rational polygons as rotation sets of generic homeomorphisms of the two torus". *J. Lond. Math. Soc. (2)* 89.1, pp. 235–254. MR: 3174742 (cit. on p. 2014).

A. Passeggi, R. Potrie, and M. Sambarino (2017). "Rotation intervals and entropy on attracting annular continua". *to appear in Geometry & Topology* (cit. on p. 2031).

R. Pérez-Marco (1997). "Fixed points and circle maps". *Acta Math.* 179.2, pp. 243–294. MR: 1607557 (cit. on p. 2014).

R. V. Plykin (1974). "Sources and sinks of A-diffeomorphisms of surfaces". *Mathematics of the USSR-Sbornik* 23.2, p. 233 (cit. on p. 2014).

M. Pollicott (1992). "Rotation sets for homeomorphisms and homology". *Trans. Amer. Math. Soc.* 331.2, pp. 881–894. MR: 1094554 (cit. on p. 2014).

C. Pommerenke (1992). *Boundary Behaviour of Conformal Maps*. Grundlehren der Mathematischen Wissenschaften 299. Springer (cit. on p. 2018).

Ch. Pugh and C. Robinson (1983). "The C^1 closing lemma, including Hamiltonians". *Ergod. Theory Dyn. Syst* 3.2, pp. 261–313. MR: 742228 (cit. on p. 2015).

C. Robinson (1970). "Generic properties of conservative systems". *Amer. J. Math.* 92, pp. 562–603. MR: 0273640 (cit. on p. 2026).

S. Ushiki (1980). "Sur les liasons-cols des systèmes dynamiques analytiques". *C. R. Acad. Sci. Paris Sér. A-B* 291.7, A447–A449. MR: 595912 (cit. on p. 2027).

R. B. Walker (1991). "Periodicity and decomposability of basin boundaries with irrational maps on prime ends". *Trans. Amer. Math. Soc.* 324.1, pp. 303–317 (cit. on pp. 2015, 2019).

Zhihong Xia (2006). "Area-preserving surface diffeomorphisms". *Comm. Math. Phys.* 263.3, pp. 723–735. MR: 2211821 (cit. on p. 2015).

E. Zehnder (1973). "Homoclinic points near elliptic fixed points". *Comm. Pure Appl. Math.* 26, pp. 131–182. MR: 0345134 (cit. on p. 2026).

Received 2017-11-28.

Andres Koropecki
Instituto de Matemática e Estatística
Universidade Federal Fluminense
Niterói
Brazil
ak@id.uff.br

Meysam Nassiri
School of Mathematics
Institute for Research in Fundamental Sciences (IPM)
Tehran
Iran
nassiri@ipm.ir

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2035–2052)

GEOMETRY OF TEICHMÜLLER CURVES

Martin Möller

Abstract

The study of polygonal billiard tables with simple dynamics led to a remarkable class of special subvarieties in the moduli of space of curves called Teichmüller curves, since they are totally geodesic submanifolds for the Teichmüller metric.

We survey the known methods to construct of Teichmüller curves and exhibit structure theorems that might eventually lead towards the complete classification of Teichmüller curves.

Introduction

The origin of the notion Teichmüller curve goes back to a remarkable discovery of Veech [1989] who constructed billard tables where the trajectories of a bouncing billiard ball have a remarkably simple dynamics, as simple as on a rectangular table. An unfolding construction of the billiard table yields a flat surface, that is a compact Riemann surface together with a flat metric and a finite number of cone-type singularities. Shearing such a flat surface by elements in $\mathrm{GL}_2^+(\mathbb{R})$ provides a whole family of flat surfaces. Only rarely is such an orbit closed in the moduli space $\Omega\mathcal{M}_g$ parametrizing flat surfaces. In this case the image of the $\mathrm{GL}_2^+(\mathbb{R})$-orbit in the moduli space of curves $\mathcal{M}_g$ is an algebraic curve, called a Teichmüller curve.

The moduli space of curves is not a locally homogeneous space and thus does not come naturally with a distinguished class of special algebraic subvarieties. Thanks to the $\mathrm{GL}_2^+(\mathbb{R})$-action, the moduli space of flat surfaces $\Omega\mathcal{M}_g$ inherits quite a bit of the properties of a homogeneous space. The special subvarieties there are affine invariant submanifolds and the smallest of them are Teichmüller curves. Part of the beauty of studying their geometry is reflected in the fact that Teichmüller curves admit a variety of different characterizations that may roughly be phrased as follows.

MSC2010: primary 32G15; secondary 37D50, 32G20, 57M50.
Keywords: Teichmüller curve, flat surfaces, real multiplication, torsion points, $\mathrm{GL}_2^+(\mathbb{R})$-action, affine invariant manifolds.

i) Teichmüller curves are immersed curves in the moduli space of curves $\mathcal{M}_g$ that are totally geodesic for the Teichmüller metric.

ii) Teichmüller curves are the images in $\mathcal{M}_g$ of closed $\mathrm{GL}_2^+(\mathbb{R})$-orbits in the moduli space $\Omega\mathcal{M}_g$ of flat surfaces.

iii) Teichmüller curves are curves in $\mathcal{M}_g$ whose variation of Hodge structures contains a rank two summand whose Kodaira-Spencer map is an isomorphism.

iv) Teichmüller curves are the images in $\mathcal{M}_g$ of two-dimensional subvarieties in strata of $\Omega\mathcal{M}_g$ cut out by torsion and real multiplication.

In all likelihood, Teichmüller curves should not exist, maybe except for low genus examples, and examples derived from them. The torsion and real multiplication conditions are just too restrictive. And yet they do exist!

The goal of this survey is to explain the above characterizations of Teichmüller curves and to summarize the current state of knowledge on the classification and geometry of Teichmüller curves.

1 Dynamically optimal billiard tables and flat surfaces

We start with a rational polygonal billiard table, that is, a planar polygon P all whose angles are rational multiples of π. The trajectories of a single ball bouncing in such a P might exhibit various types of long-term behavior. (If the trajectory hits a corner it just ends there and subsequently we disregard this measure zero set of cases.) The trajectory could be periodic. Second, trajectory might be dense, more precisely uniformly distributed all over the polygon, that is, visit every region with frequency proportional to the volume of the region. Last, the trajectory might be dense in some region strictly smaller than the whole polygon. For a rectangular table, the last possibility does not occur. Moreover which of the two first cases occurs depends on the initial direction only, not on the starting point. This simple trajectory behavior that rectangular tables exhibit is called *Veech dichotomy* or *optimal dynamics*. Polygons tiled by a rectangular table also exhibit this optimal dynamics.

Understanding whether a billiard table has optimal dynamics is simplified by performing the Katok-Zemlyakov unfolding construction Zemljakov and Katok [1975], as illustrated in Figure 1.

Instead of reflecting the trajectory at the boundary of the polygon we reflect the polygon and continue the trajectory by a straight line. Since the polygon is rational this process ends with a finite number of reflection copies. Gluing them together gives a *flat surface* (X, ω), that is, a compact Riemann surface X together with a holomorphic one-form ω

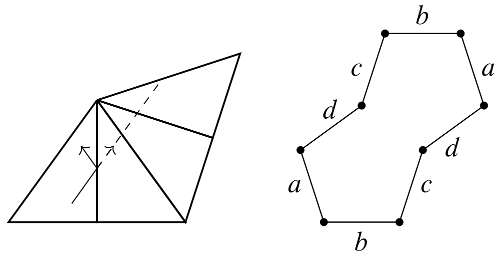


Figure 1: A $(\pi/2, \pi/5, 3\pi/10)$-triangle unfolds to a double pentagon

that provides us with a flat metric $|\omega|$ outside a finite number of cone-type singularities where the angle is a multiple of 2π. The set of all flat surfaces fit into a moduli space $\Omega\mathcal{M}_g$ with a natural forgetful map to the moduli space of curves $\mathcal{M}_g$. This moduli space is decomposed into *strata*

$$(1) \qquad \Omega\mathcal{M}_g = \bigcup_{\mu} \Omega\mathcal{M}_g(\mu) \quad \Big(\mu = (m_1, \ldots, m_n) \quad \sum_{i=1}^{n} m_i = 2g-2\Big),$$

according to the multiplicities μ of the zeros of ω.

In the example given in Figure 1 the surface has genus $g(X) = 2$ and one zero of cone angle 6π at the point marked with a dot. It is a Veech surface in the stratum $\Omega\mathcal{M}_2(2)$.

The moduli space of flat surfaces $\Omega\mathcal{M}_g$ carries a natural action of $\mathrm{GL}_2^+(\mathbb{R})$ induced by the linear action on planar polygons, see Figure 3. This action preserves the stratification (1). Moreover, the straight line flow on (X, ω) is dynamically optimal if and only if it is dynamically optimal on $A \cdot (X, \omega)$ for any $A \in \mathrm{GL}_2^+(\mathbb{R})$.

The initial observation of Veech [1989, 1991] was that if the $\mathrm{GL}_2^+(\mathbb{R})$-orbit of (X, ω) is closed in its stratum $\Omega\mathcal{M}_g(\mu)$, then (X, ω) has optimal dynamics. This is to say that for each direction θ one of two cases happen: Either all trajectories in the direction θ are uniformly distributed (hence dense) or the Veech surface is foliated in direction θ by closed geodesics and saddle connections between the saddle points, the zeros of ω. (The converse holds in low genus McMullen [2005b], but it is false in general Smillie and Weiss [2008].) Such a $\mathrm{GL}_2^+(\mathbb{R})$-orbit is closed if there is a lattice $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ stabilizing (X, ω) and the converse also holds Smillie and Weiss [2004]. Since the rotation images and the homothety images of (X, ω) are in the same fiber of the projection to $\mathcal{M}_g$, the images in $\mathcal{M}_g$ of closed $\mathrm{GL}_2^+(\mathbb{R})$-orbits are of the form $C = \Gamma\backslash\mathbb{H} \to \mathcal{M}_g$. They are immersed algebraic (but non-complete) curves in the moduli space of curves.

The action of $\mathrm{GL}_2^+(\mathbb{R})$ extends to an action on the moduli space $\mathcal{Q}_g$ that parameterizes *half-translation surfaces* (Y,q) consisting of a Riemann surface with a quadratic differential q and the above statements about images of orbits in $\mathcal{M}_g$ carry over verbatim. To summarize:

Definition 1.1. *A flat surface* (X,ω) *or a half-translation surface* (Y,q) *with closed* $\mathrm{GL}_2^+(\mathbb{R})$*-orbit is called* Veech surface *and the image curve* $C \to \mathcal{M}_g$ *of this orbit is called a* Teichmüller curve.

The name Teichmüller curve reflects that images of $\mathrm{GL}_2^+(\mathbb{R})$-orbits in the Teichmüller space $\mathcal{T}_g$ of any (Teichmüller-) marked half-translation surface (Y,q) are *Teichmüller discs*, i.e. discs $\Delta \to \mathcal{T}_g$ that are totally geodesic for the Teichmüller metric. By Teichmüller's fundamental theorems all such Teichmüller disc can be obtained as the orbits of half-translation surfaces (Y,q).

In some sense it is not strictly necessary to discuss the case of half-translation surfaces. Associated with any half-translation surface (Y,q) there is a canonical $\mathrm{GL}_2^+(\mathbb{R})$-equivariant double cover construction $\pi : X \to Y$ on which the quadratic differential $\pi^* q = \omega^2$ admits a square root. Consequently to each Teichmüller curve $C \to \mathcal{M}_{g(Y)}$ generated by (Y,q) there is a Teichmüller curve $C' \to \mathcal{M}_{g(X)}$ in a moduli space of somewhat larger genus, generated by a flat surface. Since most of the classification of Teichmüller curves works using the cohomology of the Veech surfaces and hence abelian differentials, we *restrict ourselves from now on to Teichmüller curves generated by flat Veech surfaces* (X,ω). (The reader might then check in each case at hand if the surface admits an involution that makes those surfaces arise as double coverings.) Also the itemized characterizations in the introduction are *equivalent only on this subclass of Teichmüller curves*.

The uniformizing group Γ is also called the *Veech group* of the Veech surface (X,ω). It can be characterized as the group of orientation-preserving homeomorphisms of X that are affine when expressed in the flat charts of $X \setminus Z(\omega)$ provided by ω. An important invariant of Γ and thus of any Teichmüller curve is the *trace field* $K = \mathbb{Q}[\mathrm{tr}(\gamma), \gamma \in \Gamma]$.

2 The list of known examples

The known examples of Veech surfaces and Teichmüller curves consist of a short list of series, up to a natural notion of primitivity. We present all these series and come back in the subsequent sections to the ideas behind their discovery.

A Veech surface (X,ω) is called *(geometrically) imprimitive* if there is a (branched) covering $\pi : X \to Y$ such that $\omega = \pi^*\eta$ for some one-form η on Y. Otherwise (X,ω) is called *geometrically primitive*. A Veech surface (X,ω) is called *algebraically primitive*

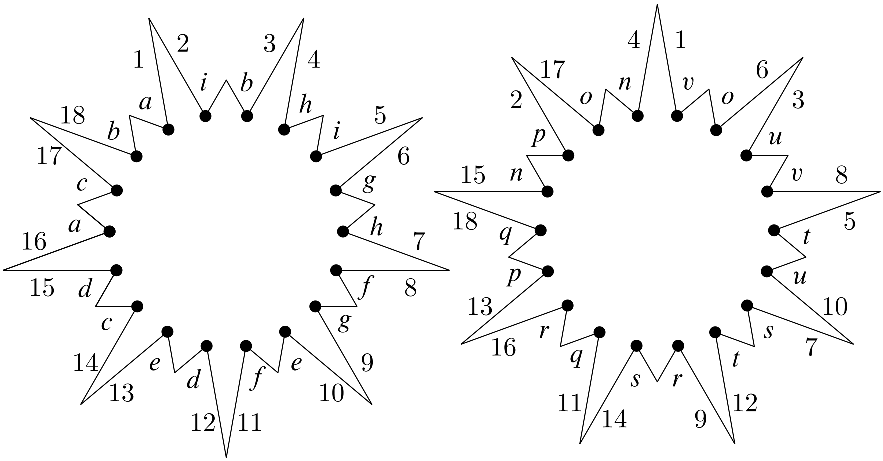


Figure 2: Veech surface with Veech group $\Delta(5, 9, \infty)$

if the trace field has degree $g = [K : \mathbb{Q}]$. Theorem 3.1 implies that algebraically primitive implies geometrically primitive. All these properties are $\mathrm{GL}_2^+(\mathbb{R})$-equivariant and we abuse the corresponding notions also for the Teichmüller curves the Veech surfaces generate.

The triangle group series. We currently know of a single series of primitive Teichmüller curves generated by Veech surfaces of unbounded genera, containing infinitely many algebraically primitive Teichmüller curves. This series is indexed by two parameters $m, n \in \mathbb{N} \cup \infty$ and constructed so that the Veech groups are the triangle groups $\Delta(m, n, \infty)$. The family was discovered in Bouw and Möller [2010b] and contains the original examples of Veech ($n = 2$ and $n = \infty$) and those of his student ($n = 3$, Ward [1998]), see Figure 2 for the case $m = 5$ and $n = 9$. Other polygonal presentations were given in the work of Hooper [2013] and Wright [2013].

The Weierstraß family and the Prym family. The Weierstraß family is generated by (nearly) L-shaped flat surfaces in the stratum $\Omega\mathcal{M}_2(2)$ of genus two as in Figure 3. The family was discovered independently by Calta [2004] and McMullen [2003]. Veech surfaces generating all the Teichmüller curves in this series are given by side length parameters $a = (0, \lambda)$, $c = (\lambda, 0)$, $b = (t, h)$, $c + d = (w, 0)$ where $t \in \mathbb{N}$, $w, h \in \mathbb{N}_{>0}$ and where $\lambda = (e + \sqrt{D})/2$ for $D = e^2 + 4wh$ is a quadratic irrational number.

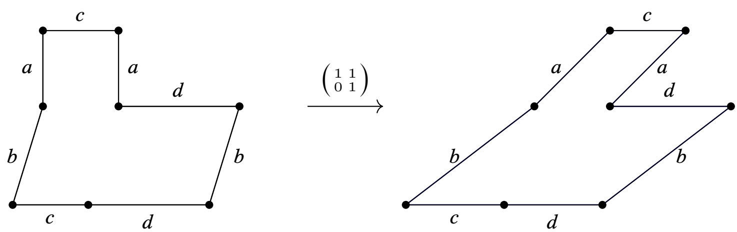


Figure 3: An L-shaped table (Calta [2004] and McMullen [2003] and the $\mathrm{GL}_2^+(\mathbb{R})$-action

The Prym family is generated by the S-shaped genus three surfaces and the X-shaped genus four surfaces (in the strata $\Omega\mathcal{M}_3(4)$ and $\Omega\mathcal{M}_4(6)$) in Figure 4, discovered by McMullen [2006a]. The trace field has degree $r = 2$ in all cases.

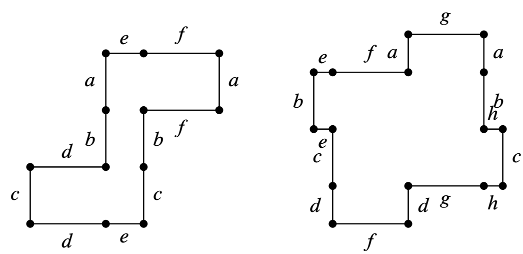


Figure 4: Veech surfaces in Prym family for $g = 3$ and $g = 4$ (McMullen [2006a])

The Gothic family. This is an infinite family of primitive Teichmüller curves discovered by McMullen, Mukamel, and Wright [2017] in the stratum $\Omega\mathcal{M}_4(2,2,2)$. It is generated by Veech surfaces that resemble Gothic cathedrals (see Figure 5), again with trace field of degree $r = 2$. Another infinite series, generated by Veech surfaces in the stratum $\Omega\mathcal{M}_4(3,3)$, has been announced by Eskin, McMullen, Mukamel and Wright.

The sporadic examples. There are two sporadic examples of Teichmüller curves. The Veech surfaces are constructed as the unfolding of the $(2\pi/9, \pi/3, 4\pi/9)$-triangle (in the stratum $\Omega\mathcal{M}_3(3,1)$, see Kenyon and Smillie [2000]) and as the unfolding of the $(\pi/3, \pi/5, 7\pi/15)$ triangle (in the stratum $\Omega\mathcal{M}_4(6)$, see Vorobets [1996]).

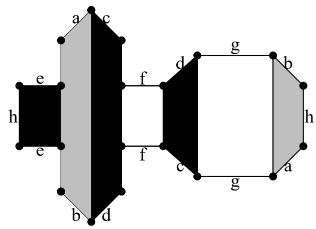


Figure 5: A Gothic Veech surface (McMullen, Mukamel, and Wright [2017])

3 Teichmüller curves and variations of Hodge structures

This section reveals the algebro-geometric nature of Teichmüller curves by two basic structure theorems. We also highlight similarities and differences to Shimura curves.

Theorem 3.1 (Möller [2006b]). *If (X, ω) is a Veech surface, then the Jacobian* $\mathrm{Jac}(X)$ *contains an abelian subvariety* $\mathrm{Jac}(X, \omega)$ *of dimension $r = [K : \mathbb{Q}]$ with real multiplication by the trace field K, i.e. the endomorphism ring of* $\mathrm{Jac}(X, \omega)$ *is an order in K.*

Here $\mathrm{Jac}(X, \omega)$ is the smallest abelian subvariety of $\mathrm{Jac}(X)$ whose tangent space contains ω via the canonical identification $T_{\mathrm{Jac}(X)} \cong \Gamma(X, \Omega^1_X)$.

For a Veech surface (X, ω) we let $z_1, \ldots, z_n$ be the zeros of ω, i.e. $\mathrm{div}(\omega) = \sum m_i z_i$.

Theorem 3.2 (Möller [2006a]). *For any i, j the divisor $[z_i - z_j]$ has finite order in* $\mathrm{Jac}(X, \omega)$.

To sketch the proof of the two theorems we consider the Teichmüller curve $C \to \mathcal{M}_g$ generated by the Veech surface. After passing to a finite unramified (in the orbifold sense) cover of C the universal curve over (a level cover of) $\mathcal{M}_g$ pulls back to a family of curves $f : \mathfrak{X} \to C$ that we may extend to a family of stable curves $\overline{f} : \overline{\mathfrak{X}} \to \overline{C}$ over a complete curve $\overline{C}$. The vector spaces $H^1(X, \mathbb{Q})$ glue to a locally constant bundle $\mathbb{V}_{\mathbb{Q}}$ over C. The Hodge bundle, the vector bundle with fiber $H^0(X, \Omega^1_X)$, is a subbundle of (the extension to $\overline{C}$ of) the vector bundle $\mathbb{V}_{\mathbb{C}}$. This vector bundle inclusion together with a polarization stemming from the symplectic pairings on the fibers of f defines a weight one variation of Hodge structures (VHS). The starting point for all theorems in this section is the decomposition (as variation of Hodge structures)

$$\mathbb{V}_{\mathbb{Q}} = \mathbb{W}_{\mathbb{Q}} \oplus \mathbb{M}_{\mathbb{Q}}, \quad \text{where} \quad \mathbb{W}_K = \mathbb{L}_1 \oplus \cdots \oplus \mathbb{L}_r \tag{2}$$

over any Teichmüller curve. Here $\mathbb{L}_1$ is the rank-two locally constant subbundle generated by $\mathrm{Re}(\omega)$ and $\mathrm{Im}(\omega)$, the *tautological plane*, and the $\mathbb{L}_i$ are the Galois conjugates of $\mathbb{L}_1$

over K. This decomposition follows from Deligne's semisimplicity of VHS and since the tautological plane is a sub-VHS essentially by definition of a Teichmüller curve as $\mathrm{GL}_2^+(\mathbb{R})$-orbit.

A $\mathbb{Q}$-decomposition of $\mathbb{V}_\mathbb{Q}$ defines a splitting of the family of Jacobian varieties. Moreover, for any $\lambda \in K$ the endomorphism $\oplus\sigma_i(\lambda)$, with σ_i running over the real embeddings of K, defines a rational endomorphism of the family of abelian subvarieties corresponding to $\mathbb{W}_\mathbb{Q}$, which provides the real multiplication claimed in Theorem 3.1.

The locally constant subbundle $\mathbb{L}_1$ is special, since by construction its monodromy representation is the standard representation of the uniformizing Fuchsian group Γ of C. As a consequence, the period map from the universal cover of C to the upper half plane, the period domain for such a rank two subsheaf, is an isomorphism. In particular, if we let $\mathcal{L}$ be the $(1,0)$-part of (the Deligne extension to $\overline{C}$ of) $\mathbb{L}_1$ then the Higgs field (also known as the Kodaira-Spencer map), that is the derivative

$$\tau : \mathcal{L} \to \mathcal{L}^{-1} \otimes \Omega^1_{\overline{C}}(\Delta), \qquad (\Delta = \overline{C} \setminus C) \tag{3}$$

of the period map, is an isomorphism. Those subbundles are called *maximal Higgs*, since for those subbundles the degree of $\mathcal{L}$ attains its maximum value $\frac{1}{2}\deg \Omega^1_{\overline{C}}(\Delta)$. The above translation can also be read backwards: maximal Higgs subbundles have period maps that are isomorphisms. Consequently, suppressing the omnipresent passages to finite unramified covers, we can summarize the discussion by the following characterization of Teichmüller curves in the language of complex geometry.

Proposition 3.3. *A Teichmüller curve is a curve $C \to \mathcal{M}_g$ such that the VHS of the family $f : \mathcal{X} \to C$ contains a rank two summand that is maximal Higgs. This maximal Higgs summand is unique.*

To illustrate the idea behind Theorem 3.2 note that the zeros $z_1, \dots, z_n$ on an individual Veech surface can be transported along the whole family $f : \mathcal{X} \to C$ without colliding, again by definition of $\mathrm{GL}_2^+(\mathbb{R})$-action. Passing to an unramified cover of C we may assume that they are the images of sections $z_i : C \to \mathcal{X}$. Using the theory of Néron models one can show that a finite index subgroup of the group of sections extends to the family of Jacobians. We may then project these sections to the family $\overline{\mathcal{A}} \to \overline{C}$ of the abelian subvarieties whose fibers are $\mathrm{Jac}(X,\omega)$. But this family does not have any non-zero sections. In fact, by the uniformization of the family $\overline{\mathcal{A}} \to \overline{C}$, sections can be identified with elements of $H^1(\overline{C}, \mathbb{W}_\mathbb{Q})$. This cohomology group naturally has a weight two Hodge structure and the sections provide elements of type $(1,1)$. The maximal Higgs direct summand of $\mathbb{W}_\mathbb{Q}$ prohibits the existence of such non-zero elements.

The two theorems can be recast to characterize Teichmüller curves purely using terms from algebraic geometry, as observed by Alex Wright.

Proposition 3.4. *A Teichmüller curve is the image in $\mathcal{M}_g$ of a two-dimensional suborbifold $\mathcal{M}$ of a stratum $\Omega\mathcal{M}_g(\mu)$ such that for each point $[(X,\omega)] \in \mathcal{M}$ the abelian variety* $\mathrm{Jac}(X,\omega)$ *has real multiplication by an order in a field of degree* $\dim(\mathrm{Jac}(X,\omega))$ *and such that for any two zeros z_1, z_2 of ω the difference $z_1 - z_2$ is torsion in* $\mathrm{Jac}(X,\omega)$.

Shimura curves are also defined as totally geodesic curves, but in the moduli space of Abelian varieties $\mathcal{A}_g$ (instead of $\mathcal{M}_g$) and for the Bergman-Siegel metric (instead of the Kobayashi metric). Usually Shimura curves are moreover required to have a CM point, and sometimes they are referred to as Kuga curves with this conditions relaxed. Shimura curves can also be defined as stemming from a homomorphism of a $\mathbb{Q}$-algebraic group into the symplectic group by quotienting the corresponding real groups by maximal compact subgroups and a lattice. Since $\mathcal{A}_g$ is a locally homogeneous space and since Shimura curves are defined by group theory, there are plenty of Shimura curves. However, since the Torelli-image of $\mathcal{M}_g$ in $\mathcal{A}_g$ is of large codimension for $g \to \infty$, most of the Shimura curves don't intersect the Torelli-image. The classification of Shimura curve in (the Torelli-image of) $\mathcal{M}_g$ is an open problem that is morally similar to the classification of Teichmüller curves, see e.g. Lu and Zuo [2014] for one of the latest results.

Shimura curves can also be characterized by a decomposition of the VHS like in (2), but now the bundle $\mathbb{M}_{\mathbb{Q}}$ has to have unitary monodromy and now all the bundles $\mathbb{L}_i$ have to be maximal Higgs (rather than just one of them), but they are not necessarily all Galois conjugates.

4 Constructing Veech surfaces and computing the Veech group

We revisit the known examples of primitive Teichmüller curves in the light of the previous structure results and sketch their method of construction.

Veech's and Ward's original examples were constructed by exhibiting two elements in the Veech group that jointly generate a Fuchsian triangle group. In the example of the double pentagon in Figure 1 this is the triangle group generated by the rotation $\begin{pmatrix} \cos 2\pi/5 & \sin 2\pi/5 \\ -\sin 2\pi/5 & \cos 2\pi/5 \end{pmatrix}$ and the vertical shear $\begin{pmatrix} 1 & 0 \\ 2\cot\pi/5 & 1 \end{pmatrix}$. The latter element belongs to the Veech group since the straight line flow in the vertical direction is periodic and the periodic orbits come in two homotopy classes, each sweeping out a cylinder bounded by saddle connections. Both cylinders have the same modulus equal to $2\cot\pi/5$.

Expanding on the previous remark we note that Teichmüller curves are never compact since any direction on the Veech surface admitting a saddle connection provides a parabolic element in the Veech group that has this direction as an eigenvector. However, the Veech groups are in general not generated by elliptic and parabolic elements, as we will prove in Section 7. In fact, none of the series of Teichmüller curves besides the original examples of Veech and Ward was detected by computing the Veech group! Only recently Mukamel [2017] gave an algorithm to compute the Veech group for a general Veech surface. His basic idea is to associate to each Veech surface over a Teichmüller curve the number of *girth directions* that contain a shortest saddle connection. This number provides a stratification of the $\mathrm{GL}_2^+(\mathbb{R})$-orbit of a Veech surface, since the generic number of girth directions is one. The algorithm proceeds by tracing along the spine of this stratification (consisting of surfaces with two girth directions) and testing if the Veech surfaces at vertices of the stratification are scissors congruent to each other.

The construction of the triangle group series started with the observation that families of cyclic coverings have rank two summands in their cohomology whose monodromy groups are triangle groups. However, these summands are not maximal Higgs in the sense of (3). The problem is that at the points where the monodromy has finite order the family of curves degenerates, but the period map can be continued over these points (after passing to a finite cover). Luckily, if we consider triangle groups (m, n, ∞), say with m, n odd and coprime for simplicity, the group $(\mathbb{Z}/2)^2$ acts on the family of cyclic covers. The quotient family still has the rank two summand in cohomology. Moreover, the fibers over the orbifold points are now smooth and Proposition 3.3 applies.

The Weierstraß series consists of surfaces $(X, \omega) \in \Omega\mathcal{M}_2(2)$ whose cohomology admits a self-adjoint endomorphism $\phi \in \mathrm{End}(H^1(X, \mathbb{Z}))$ such that $\phi^*\omega = \lambda\omega$ for some λ generating a fixed real quadratic extension K of $\mathbb{Q}$. Such a map ϕ defines an endomorphism of $\mathrm{Jac}(X)$, since it preserves the period lattice and it induces a consistent map on the tangent space of $\mathrm{Jac}(X)$ as it preserves the line $\mathbb{C} \cdot \omega$ by definition and another complex line by self-adjointness. The existence of such an endomorphism involves only the periods of ω and can thus be checked to hold for surfaces of the form in Figure 3 with the explicitly given parameters of the saddle connection vectors on the boundary. In the minimal stratum the torsion condition is void and thus Proposition 3.4 implies that such (X, ω) are Veech surfaces.

For the Prym series the crucial observation is that the same argument as for the Weierstraß series can be made for a four-dimensional part of cohomology that is (anti-)invariant by an involution rather than for the whole $H^1(X, \mathbb{Z})$.

For the Gothic series this observation is refined to work for endomorphisms acting on an even smaller part of $H^1(X, \mathbb{Z})$, the kernel of *two* projections to the first cohomology of smaller genus curves, provided that the ambient variety without imposing the real multiplication endomorphism behaves like the stratum $\Omega\mathcal{M}_2(2)$ in a sense made precise in the next section.

5 Affine invariant manifolds

Recall that Teichmüller curves are images of closed $\mathrm{GL}_2^+(\mathbb{R})$-orbits. The striking results of Eskin and Mirzakhani [2013] and Eskin, Mirzakhani, and Mohammadi [2015] imply that all the non-closed $\mathrm{GL}_2^+(\mathbb{R})$-orbits have very nice closures: They are manifolds, affine and $\mathbb{R}$-linear in a natural 'period' coordinate system, and more precisely quasi-projective varieties by Filip's results Filip [2016]. These orbit closures are thus called *affine invariant manifolds (AIM)*. Their classification is a very interesting question that created a lot of recent activity. We refer e.g. to Apisa [2015], Aulicino and Nguyen [2016], and Eskin, Filip, and Wright [n.d.] for some of the latest results and highlight here only the aspects connected with the classification of Teichmüller curves.

Suppose some stratum $\Omega\mathcal{M}_g(\mu)$ contains an infinite number of algebraically primitive Teichmüller curves C_i for $i \in \mathbb{N}$. The closure of their union is an AIM $\mathcal{M}$ by Eskin, Mirzakhani, and Mohammadi [2015]. The main observation of Matheus and Wright [2015] is that it is possible to spread out the decomposition information (2) from the union of the C_i to all of $\mathcal{M}$. Namely, they define a Hodge-Teichmüller plane over the moduli point of $(X.\omega)$ to be a $\mathbb{C}$-rank-two subspace $L \subset H^1(X,\mathbb{C})$ defined over $\mathbb{R}$ such that all its $\mathrm{GL}_2^+(\mathbb{R})$-translates intersect the $(1,0)$-part of the cohomology in a one-dimensional subspace. By (2) each point over each C_i has g orthogonal Hodge-Teichmüller planes and by a limiting argument each point of $\mathcal{M}$ has them. This leads to an immediate contradiction in many cases, e.g. when the monodromy representation on $H^1(X,\mathbb{C})$ over Teichmüller curves generated by Veech surfaces that are torus covers can be shown to not have that many Hodge-Teichmüller planes.

This idea was subsequently refined (by working with relative cohomology and by computing the algebraic hull of $\mathrm{GL}_2^+(\mathbb{R})$-cocycle for general AIM, hence in particular for those containing an infinite number of Teichmüller curves) to yield the following optimal (though ineffective) finiteness result.

Theorem 5.1 (Eskin, Filip, and Wright [n.d.]). *Each stratum $\Omega\mathcal{M}_g(\mu)$ contains only a finite number of Teichmüller curves with trace field of degree $r > 2$.*

In each stratum $\Omega\mathcal{M}_g(\mu)$ there are only a finite number of AIMs $\mathcal{M}_i$ 'like $\Omega\mathcal{M}_2(2)$' such that all primitive Teichmüller curves with $r = 2$ are contained in one of these $\mathcal{M}_i$. Conversely, any such AIM 'like $\Omega\mathcal{M}_2(2)$' contains infinitely many Teichmüller curves with $r = 2$.

To give a precise definition of an AIM 'like $\Omega\mathcal{M}_2(2)$' we recall that the tangent space of a stratum $\Omega\mathcal{M}_g(m_1,\ldots,m_n)$ at (X,ω) is modelled on the relative cohomology $H^1(X,Z(\omega),\mathbb{C})$, where $Z(\omega) = \{z_1,\ldots,z_n\}$ is the zero set of ω. The tangent space to an AIM $\mathcal{M}$ is by Eskin and Mirzakhani [2013] and Wright [2015] a linear subspace $T\mathcal{M} \subseteq H^1(X,Z(\omega),\mathbb{C})$, defined over a real number field K that generalizes the notion of the trace field. The *rank* of $\mathcal{M}$ is the integer $\frac{1}{2}\dim p(T\mathcal{M})$ where $p : H^1(X, Z(\omega), \mathbb{C}) \to H^1(X, \mathbb{C})$. It measures (half of) the number of moduli of $\mathcal{M}$ discounting those that stem from moving the zeros relative to each other. An AIM is 'like $\Omega\mathcal{M}_2(2)$' if it is rank two and $K = \mathbb{Q}$.

The quest for the classification of primitive Teichmüller curves is thus reduced to detecting the cases with exceptionally large trace field $r > 2$ (like most examples of the triangle series) and the AIMs 'like $\Omega\mathcal{M}_2(2)$'. Currently, we only know of the few examples mentioned in the previous section.

6 Finiteness and classification results

All presently known finiteness and classification results for Teichmüller curves are based on the study of their cusps.

The classification of Teichmüller curves in the Weierstraß series and the Prym series starts with listing all possible cusps, which amounts to a finite list of possible combinatorics for the saddle connections and a list of possible length data compatible with real multiplication by the order of a given discriminant D. The main problem is to detect when two cusps lie on the same Teichmüller curve. Sometimes it is possible to spot this, like for the cusps belonging to the horizontal and vertical direction in Figure 4. Spotting enough of those direction changes to connect any pair of cusps is the tedious step in the proof of the following theorem.

Theorem 6.1 (McMullen [2005a] and Lanneau and Nguyen [2014]). *In $\Omega\mathcal{M}_2(2)$ there is a unique primitive Teichmüller curve W_D with real multiplication by the order of discriminant D for each $D \neq 1 \mod 8$ and two such Teichmüller curves $W_D^{\pm}$ for each $D \equiv 1 \mod 8$.*

The Prym series in $g = 3$ consists of a unique primitive Teichmüller curve $W_D(4)$ with real multiplication by the order of discriminant D for each $D \equiv 0, 4 \mod 8$, it has two components $W_D^{\pm}(4)$ for $D \equiv 1 \mod 8$ and is empty for $D \equiv 5 \mod 8$.

A similar result for the $g = 4$-series is known in some cases Lanneau and Nguyen [2014] and the classification is open for Gothic curves. It would be very interesting to find a more conceptual argument for the classification of connected components. This classification is currently the only property of Teichmüller curves not accessible through the viewpoint of modular forms, see Section 7 below.

Around the time of discovery, it was puzzling that $\Omega\mathcal{M}_2(2)$ contains infinitely many primitive Teichmüller curves, while in $\Omega\mathcal{M}_2(1,1)$ there was only a single such curve known, the decagon in Veech's original family. Given Theorem 3.2 this should no longer come as a surprise: Finding two points z_1 and z_2 on a Riemann surface X whose difference is torsion is a very rare pick. It is equivalent to finding a map $p : X \to \mathbb{P}^1$ with

$p^{-1}(0) = \{z_1\}$ and $p^{-1}(\infty) = \{z_2\}$. Pushing this condition to the cusp of a primitive Teichmüller curve in $\Omega\mathcal{M}_2(1,1)$ amounts to detecting when ratios of sines at rational multiples of π belong to a quadratic number field. There are only finitely many possibilities and examining these cases, McMullen showed in McMullen [2006b] that the decagon is indeed the only example in $\Omega\mathcal{M}_2(1,1)$.

Even if there are no torsion constraints, i.e. in the minimal strata $\Omega\mathcal{M}_g(2g-2)$, there are many other constraints imposed by cusps on the existence of a primitive Teichmüller curve. We illustrate this in the smallest interesting case, the hyperelliptic connected component of the stratum $\Omega\mathcal{M}_3(4)$. The fiber over the cusp of an algebraically primitive Teichmüller curve is a projective line (by the real multiplication condition) and the limit of the generating one-form is a stable differential ω_∞ that we may normalize due to the hyperelliptic involution to have simple poles at $\pm x_i$ for $i = 1, 2, 3$ with residues $\pm r_i$ and a four-fold zero at zero. This amounts to the conditions

$$\sum_{i=1}^{3} r_i x_{i+1} x_{i+2} = 0 \quad \text{and} \quad \sum_{i=1}^{3} r_i x_i (x_{i+1}^2 + x_{i+2}^2) = 0, \tag{4}$$

where indices have to be read mod 3. We can moreover normalize $r_1 = 1$ and $x_1 = 1$. Algebraically primitive implies that the r_i are a $\mathbb{Q}$-basis of a totally real cubic number field $K \subset \mathbb{R}$ and we denote the two other real embeddings of K by σ and τ.

We now hint at two additional constraints. First, since the cusps lie on the boundary of Hilbert modular threefolds, whose boundary has been computed in terms of cross-ratios in Bainbridge and Möller [2012], the cross-ratio equation

$$R_{23}^{b_1} R_{13}^{b_2} R_{23}^{b_3} = 1, \quad \left(R_{ij} = \left(\frac{x_i + x_j}{x_i - x_j}\right)^2\right) \tag{5}$$

holds, where b_i are integers such that $\sum_{i=1}^{3} b_i / s_i = 0$, where $\{s_1, s_2, s_3\}$ is the $\mathbb{Q}$-basis of K trace dual to $\{r_1, r_2, r_3\}$. Note that it is already a very restrictive property for a $\mathbb{Q}$-basis of K that the reciprocals of dual basis are $\mathbb{Q}$-linearly dependent. Second, considerations of the Harder-Narasimhan filtration of the Hodge bundle over the Teichmüller curve imply that one of the two Galois conjugate forms also has to have a double zero in common with ω. This implies

$$\sum_{i=1}^{3} r_i^{\sigma} x_{i+1} x_{i+2} = 0. \tag{6}$$

We strongly suspect that the solution stemming from Veech's 7-gon

$$r_2 = v^2 + v - 2,\ r_3 = v^2 - 2,\ x_2 = -v^2 - v + 1,\ x_3 = v^2 + v - 2,\ (v = 2\cos(2\pi/7))$$

is the unique solution in some K up to permutation of variables of the equations (4), (5). and (6). However, the fact that equations are not algebraic, but involve Galois conjugates, makes the geometry of the solution set more interesting. Currently, we can only show that the set of solutions is finite. This is part of the following version of Theorem 5.1 for $g = 3$, that has the advantage to be at least in theory algorithmically implementable.

Theorem 6.2 (Bainbridge, Habegger, and Möller [2016]). *There are finitely many algebraically primitive Teichmüller curves in genus three.*

The proof in the minimal stratum uses a variant of the theory of *just likely intersections*. The rough statement of the main theorem of this theory is that all intersection points of an algebraic subvariety Y of a multiplicative torus $\mathbb{G}_m^n$ with all subtori of dimension $n - \dim(Y) - 1$ have bounded height, except for the anomalous locus $Y^{\mathrm{an}} \subseteq Y$ consisting of the subvarieties that intersect a translate of a subtorus of $\mathbb{G}_m^n$ in a larger subvariety than expected from the naive dimension count. The subtori this theory is applied to are those defined in (5), but the variant in Bainbridge, Habegger, and Möller [ibid.] uses coupled equations in $\mathbb{G}_m^n$ and the additive group $\mathbb{G}_a^n$.

7 Modular forms and Euler characteristics

The locus of abelian surfaces with real multiplication by an order of discriminant D is the Hilbert modular surface $X_D = \mathbb{H}^2/(\mathrm{SL}(\mathfrak{o}_D \oplus \mathfrak{o}_D^\vee)$. By Theorem 3.1 the Torelli image of a Teichmüller curve with quadratic trace field lands in X_D. We can thus use modular forms and other tools from number theory to approach the geometry of Teichmüller curves, e.g. in the Weierstraß, Prym and Gothic series.

The vanishing locus of a Hilbert modular form of weight (k, ℓ) descends to an algebraic curve in X_D. Not all curves on a Hilbert modular surface are the vanishing locus of a Hilbert modular form, not even linearly equivalent to such a vanishing locus. However, all the Teichmüller curves in the Weierstraß and the Prym series can be described using Hilbert modular forms. Concretely, let $\theta_{(m,m')}(\mathbf{z}, \mathbf{u})$ be the restriction of the classical Riemann theta function with characteristic $m, m' \in \frac{1}{2}\mathbb{Z}^2/\mathbb{Z}^2$ to $\mathbf{z} = (z_1, z_2) \in \mathbb{H}^2$, let $\mathbf{u} = (u_1, u_2)$ be coordinates of $\mathbb{C}^2$ that correspond to eigendirections of real multiplication, and let $D_2\theta_{(m,m')} = \frac{\partial}{\partial u_2}\theta_{(m,m')}(\mathbf{z}, \mathbf{u})$.

Theorem 7.1 (Bainbridge [2007] and Möller and Zagier [2016]). *The image of the Weierstraß Teichmüller curve W_D in X_D is the vanishing locus of the Hilbert modular form*

$$\mathfrak{D}\theta(\mathbf{z}) = \prod_{(m,m')\ \text{odd}} D_2\theta_{(m,m')}(\mathbf{z}) \tag{7}$$

of weight $(3,9)$. *The orbifold Euler characteristic of W_D is $\chi(W_D) = -\frac{9}{2}\chi(X_D)$.*

The Euler characteristic of Hilbert modular surfaces is explicitly computable, for fundamental discriminants D it is simply $\chi(X_D) = \zeta_{\mathbb{Q}(\sqrt{D})}(-1)$.

Note that the weight of the modular form $\mathfrak{D}\theta$ is non-parallel, while in the literature almost exclusively modular forms of parallel weight (k,k) appear. The reason for this can be explained as follows. Teichmüller curves are geodesic for the Teichmüller metric on $\mathcal{T}_g$, which is the same as the Kobayashi metric. Images of Teichmüller curves are still geodesic for the Kobayashi metric on the Hilbert modular surface X_D and in fact also on the moduli space of abelian surfaces $\mathcal{A}_2$. On X_D, the Kobayashi metric is the supremum of the Poincaré metrics on the two factors. In each point of the Teichmüller curve this supremum is attained precisely for the first factor of $\mathbb{H}^2$. This is a restatement of the fact that the maximal Higgs summand in Proposition 3.3 is unique.

The proof of Theorem 7.1 recasts in terms of modular forms the fact that the Abel-Jacobi map based at a Weierstrass point embeds the Veech surface in its Jacobian as the vanishing locus of a translate of the theta divisor. Consequently, the eigenform for real multiplication has a zero at the Weierstrass point (i.e. the Veech surface belongs to the stratum $\Omega\mathcal{M}_2(2)$) if and only if the theta divisor has a tangent in an eigendirection for real multiplication. This is expressed by the right hand side of (7).

There is an analogous theorem that expresses the Teichmüller curves in the Prym series as the vanishing locus of a determinantal expression in derivatives of theta functions Möller [2014]. It also yields an expression of the Euler characteristics $\chi(W_D(4))$ and $\chi(W_D(6))$ as a multiple (depending only on $D \mod 8$) of $\chi(X_D)$. The proof refines the above argument, using that the Prym-Abel-Jacobi image of a Veech surface (X,ω) in the Prym series is immersed in $\mathrm{Jac}(X,\omega)$.

8 Orbifold points and other connections to arithmetic geometry

Orbifold points of Teichmüller curves are, besides cusps and the Euler characteristic, the last missing piece in determining their topology. Orbifold points provide an additional automorphism of the Jacobian, besides the real multiplication on $\mathrm{Jac}(X,\omega)$. Consequently, orbifold points give points of complex multiplication (in the general sense, allowing endomorphism rings that are matrix rings). Recall that by proven versions of the André-Oort conjecture (see Edixhoven [2001]) Shimura curves in Hilbert modular surfaces are characterized by having infinitely many CM points. On the other hand, a Teichmüller curve is a Shimura curve at most for those generated by some torus covering Veech surfaces in genus three and four. We will thus find only finitely many CM points and hence finitely many orbifold points on a primitive Teichmüller curve.

Orbifold points are loosely connected to billiards. The unfolding construction (Figure 1) exhibits a Veech surface (X, ω) arising from a dynamically optimal billiard table P as $X = \cup_{g \in G} gP$ for some finite group G generated by reflections. The index two subgroup $G' \subset G$ that preserves the orientation belongs to the Veech group of (X, ω). However, if the billiard table has right angles only, this group G' might just consist of an involution. E.g. for L-shaped billiard tables it gives the hyperelliptic involution, common to all Veech surfaces in genus two rather than to special orbifold points. The locus of unfoldings of billiard tables is a real codimension one submanifold of the Teichmüller curves in this case.

Orbifold points on all but the most recently discovered series of Teichmüller curves have been determined by Mukamel and by Torres-Teigell and Zachhuber.

Theorem 8.1 (Mukamel [2014] and Torres-Teigell and Zachhuber [2015, 2016])**.** *The orbifold points on the Weierstraß Teichmüller curve W_D are a point of order five on W_5 and $\tilde{h}(-D)$ points of order two.*

The orbifold points on the Prym Teichmüller curve $W_D(4)$ for $D > 12$ are $H_2(D)$ points of order two and $H_3(D)$ points of order three.

Here $\tilde{h}(-D)$ are generalized class numbers, e.g. $\tilde{h}(-D) = h(-4D)/|\mathfrak{o}^*_{-4D}|$ for odd discriminants D and $H_2(D)$ and $H_3(D)$ are representation number for D by quadratic forms, with $H_2(D) = 0$ if D is odd. For $D \leq 12$ there are a finite number of exceptional cases with orbifold points. A similar statement also holds for the Prym Teichmüller curves $W_D(6)$, see Torres-Teigell and Zachhuber [2016].

We conclude this survey by addressing various aspects that emphasize the arithmetic nature of Teichmüller curves. They are defined over number fields, since the existence of the maximal Higgs subbundle $\mathbb{L}_1$ implies rigidity (Möller and Viehweg [2010] and McMullen [2009]). Since maximal Higgs is a numerical condition, the Galois conjugate of a Teichmüller curve is again a Teichmüller curve. This allows to search for natural integral models over number rings for Teichmüller curves and to study the primes of bad reduction of these models. Such models were computed in Bouw and Möller [2010a] and many more in Kumar and Mukamel [2014], providing an interesting conjectural picture of the bad primes.

Since Teichmüller curves are characterized by their uniformization, there is a natural notion of modular forms for the Veech group. Since the universal covering of the map

$C \to X_D$ of a Teichmüller curve to the Hilbert modular surface can be written as $z \mapsto (z, \varphi(z))$ for some holomorphic map φ, there is, besides the usual automorphy factor $(cz + d)$ also the twisted automorphy factor $(c^\sigma \varphi(z) + d^\sigma)$ where σ is a generator of $\mathrm{Gal}(K/\mathbb{Q})$. This leads to a theory of twisted modular forms, studied in Möller and Zagier [2016]. However, since the Veech group of a primitive Teichmüller curve is not arithmetic there is no theory of Hecke operators on twisted modular forms. It is an open problem if there is any replacement of the pivotal role usually played by Hecke eigenforms in this context.

Acknowledgments. Considerable part of the work presented in this survey was obtained in collaboration. The author thanks his collaborators for the pleasure to work jointly. The author is grateful to C. Salzmann for her help in the preparation of the pictures.

References

P. Apisa (2015). *GL(2, R)-Orbit Closures in Hyperelliptic Components of Strata*. arXiv: 1508.05438 (cit. on p. 2045).

D. Aulicino and D.-M. Nguyen (2016). *Rank Two Affine Manifolds in Genus Three*. arXiv: 1612.06970 (cit. on p. 2045).

M. Bainbridge (2007). "Euler characteristics of Teichmüller curves in genus two". *Geom. Topol.* 11, pp. 1887–2073. MR: MR2350471 (cit. on p. 2049).

M. Bainbridge, P. Habegger, and M. Möller (2016). "Teichmüller curves in genus three and just likely intersections in $\mathbf{G}_m^n \times \mathbf{G}_a^n$". *Publ. Math. Inst. Hautes Études Sci.* 124, pp. 1–98 (cit. on p. 2048).

M. Bainbridge and M. Möller (2012). "The Deligne-Mumford compactification of the real multiplication locus and Teichmüller curves in genus 3". *Acta Math.* 208.1, pp. 1–92. MR: 2910796 (cit. on p. 2047).

I. Bouw and M. Möller (2010a). "Differential equations associated with nonarithmetic Fuchsian groups". *J. Lond. Math. Soc. (2)* 81.1, pp. 65–90. MR: 2580454 (cit. on p. 2050).

– (2010b). "Teichmüller curves, triangle groups, and Lyapunov exponents". *Ann. of Math. (2)* 172.1, pp. 139–185. MR: 2680418 (cit. on p. 2039).

K. Calta (2004). "Veech surfaces and complete periodicity in genus two". *J. Amer. Math. Soc.* 17.4, pp. 871–908. MR: MR2083470 (cit. on pp. 2039, 2040).

B. Edixhoven (2001). "On the André-Oort conjecture for Hilbert modular surfaces". In: *Moduli of abelian varieties (Texel Island, 1999)*. Vol. 195. Progr. Math. Birkhäuser, Basel, pp. 133–155. MR: 1827018 (cit. on p. 2050).

A. Eskin, S. Filip, and A. Wright (n.d.). "The algebraic hull of the Kontsevich-Zorich cocycle". arXiv: 1702.02074 (cit. on p. 2045).

A. Eskin and M. Mirzakhani (2013). "Invariant and stationary measures for the SL(2,R) action on moduli space". arXiv: 1302.3320 (cit. on p. 2045).

A. Eskin, M. Mirzakhani, and A. Mohammadi (2015). "Isolation, equidistribution, and orbit closures for the $\mathrm{SL}(2,\mathbb{R})$ action on moduli space". *Ann. of Math. (2)* 182.2, pp. 673–721 (cit. on p. 2045).

S. Filip (2016). "Splitting mixed Hodge structures over affine invariant manifolds". *Ann. of Math. (2)* 183.2, pp. 681–713 (cit. on p. 2045).

W.P. Hooper (2013). "Grid graphs and lattice surfaces". *Int. Math. Res. Not. IMRN* 12, pp. 2657–2698. MR: 3071661 (cit. on p. 2039).

R. Kenyon and J. Smillie (2000). "Billiards on rational-angled triangles". *Comment. Math. Helv.* 75.1, pp. 65–108. MR: MR1760496 (cit. on p. 2040).

A. Kumar and R. Mukamel (2014). *Algebraic models and arithmetic geometry of Teichmüller curves in genus two*. arXiv: 1406.7057 (cit. on p. 2050).

E. Lanneau and D.-M. Nguyen (2014). "Teichmüller curves generated by Weierstrass Prym eigenforms in genus 3 and genus 4". *J. Topol.* 7.2, pp. 475–522. MR: 3217628 (cit. on p. 2046).

X. Lu and K. Zuo (2014). *The Oort conjecture on Shimura curves in the Torelli locus of curves*. arXiv: 1405.4751 (cit. on p. 2043).

C. Matheus and A. Wright (2015). "Hodge-Teichmüller planes and finiteness results for Teichmüller curves". *Duke Math. J.* 164.6, pp. 1041–1077 (cit. on p. 2045).

C. McMullen (2003). "Billiards and Teichmüller curves on Hilbert modular surfaces". *J. Amer. Math. Soc.* 16.4, pp. 857–885. MR: MR1992827 (cit. on pp. 2039, 2040).

– (2005a). "Teichmüller curves in genus two: Discriminant and spin". *Math. Ann.* 333.1, pp. 87–130 (cit. on p. 2046).

– (2005b). "Teichmüller curves in genus two: the decagon and beyond". *J. Reine Angew. Math.* 582, pp. 173–199. MR: MR2139715 (cit. on p. 2037).

– (2006a). "Prym varieties and Teichmüller curves". *Duke Math. J.* 133.3, pp. 569–590. MR: MR2228463 (cit. on p. 2040).

– (2006b). "Teichmüller curves in genus two: torsion divisors and ratios of sines". *Invent. Math.* 165.3, pp. 651–672. MR: MR2242630 (cit. on p. 2047).

– (2009). "Rigidity of Teichmüller curves". *Math. Res. Lett.* 16.4, pp. 647–649. MR: 2525030 (cit. on p. 2050).

C. McMullen, R. Mukamel, and A. Wright (2017). "Cubic curves and totally geodesic subvarieties of moduli space". *Ann. of Math. (2)* 185.3, pp. 957–990 (cit. on pp. 2040, 2041).

M. Möller (2006a). "Periodic points on Veech surfaces and the Mordell-Weil group over a Teichmüller curve". *Invent. Math.* 165.3, pp. 633–649. MR: MR2242629 (cit. on p. 2041).

– (2006b). "Variations of Hodge structures of a Teichmüller curve". *J. Amer. Math. Soc.* 19.2, pp. 327–344. MR: MR2188128 (cit. on p. 2041).

– (2014). "Prym covers, theta functions and Kobayashi curves in Hilbert modular surfaces". *Amer. J. Math.* 136.4, pp. 995–1021. MR: 3245185 (cit. on p. 2049).

M. Möller and E. Viehweg (2010). "Kobayashi geodesics in $\mathfrak{A}_g$". *J. Differential Geom.* 86.2, pp. 355–579. MR: 2772554 (cit. on p. 2050).

M. Möller and D. Zagier (2016). "Modular embeddings of Teichmüller curves". *Compos. Math.* 152.11, pp. 2269–2349. MR: 3577896 (cit. on pp. 2049, 2051).

R. Mukamel (2014). "Orbifold points on Teichmüller curves and Jacobians with complex multiplication". *Geom. Topol.* 18, pp. 779–829 (cit. on p. 2050).

– (2017). "Fundamental domains and generators for lattice Veech groups". *Comment. Math. Helv.* 92.1, pp. 57–83. MR: 3615035 (cit. on p. 2044).

J. Smillie and B. Weiss (2004). "Minimal sets for flows on moduli space". *Israel J. Math.* 142, pp. 249–260 (cit. on p. 2037).

– (2008). "Veech's dichotomy and the lattice property". *Ergodic Theory Dynam. Systems* 28.6, pp. 1959–1972 (cit. on p. 2037).

D. Torres-Teigell and J. Zachhuber (2015). *Orbifold points on Prym-Teichmüller curves in genus three* . arXiv: 1502.05381. to appear in Int. Math. Res. Not. (Cit. on p. 2050).

– (2016). *Orbifold points on Prym-Teichmüller curves in genus four* . arXiv: 1609.00144. to appear in J. Math. Inst. Jussieu (cit. on p. 2050).

W. Veech (1989). "Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards". *Invent. Math.* 97.3, pp. 553–583. MR: MR1005006 (cit. on pp. 2035, 2037).

– (1991). "Erratum: "Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards"". *Invent. Math.* 103.2, p. 447. MR: MR1085115 (cit. on p. 2037).

Ya. B. Vorobets (1996). "Plane structures and billiards in rational polygons: the Veech alternative". *Uspekhi Mat. Nauk* 51.5(311), pp. 3–42. MR: 1436653 (cit. on p. 2040).

C. Ward (1998). "Calculation of Fuchsian groups associated to billiards in a rational triangle". *Ergodic Theory Dynam. Systems* 18.4, pp. 1019–1042. MR: MR1645350 (cit. on p. 2039).

A. Wright (2013). "Schwarz triangle mappings and Teichmüller curves: the Veech-Ward-Bouw-Möller curves". *Geom. Funct. Anal.* 23.2, pp. 776–809. MR: 3053761 (cit. on p. 2039).

– (2015). "Cylinder deformations in orbit closures of translation surfaces". *Geom. Topol.* 19.1, pp. 413–438. MR: 3318755 (cit. on p. 2045).

A. N. Zemljakov and A. B. Katok (1975). "Topological transitivity of billiards in polygons". *Mat. Zametki* 18.2, pp. 291–300. MR: 0399423 (cit. on p. 2036).

Received 2017-11-22.

Martin Möller
Institut für Mathematik, Goethe–Universität Frankfurt
Robert-Mayer-Str. 6–8
60325 Frankfurt am Main
Germany
moeller@math.uni-frankfurt.de

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2053–2080)

GROUP ACTIONS ON 1-MANIFOLDS: A LIST OF VERY CONCRETE OPEN QUESTIONS

Andrés Navas

Abstract

Over the last four decades, group actions on manifolds have deserved much attention by people coming from different fields, as for instance group theory, low-dimensional topology, foliation theory, functional analysis, and dynamical systems. This text focuses on actions on 1-manifolds. We present a (non exhaustive) list of very concrete open questions in the field, each of which is discussed in some detail and complemented with a large list of references, so that a clear panorama on the subject arises from the lecture.

From the very beginning, groups were recognized as mathematical objects endowed with a certain "dynamics". For instance, Cayley realized every group as a group of permutations via left translations:

$$G \longrightarrow \mathcal{P}(G), \quad g \mapsto L_g : G \to G, \quad L_g(h) = gh.$$

For a finitely-generated group G, this action has a geometric realization: We can consider the so-called *Cayley graph* of G whose vertices are the elements of G, two of which f, g are relied by an edge whenever $g^{-1} f$ is a generator (or the inverse of a generator). Then G becomes a subgroup of the group of automorphisms of this graph.

In general, the group of such automorphisms is larger than the group G. A classical result of Coxeter, Frucht, and Powers [1981] consists on a slight modification of this construction so that the automorphisms group of the resulting graph actually coincides with G. In fact, there are uncountably many such modifications, even for the trivial group. The smallest nontrivial regular graph of degree 3 with trivial automorphism group is known as *the Frucht graph*, and is depicted below.

MSC2010: primary 37C85; secondary 20F60, 22F50, 37B05, 37E10, 57R30.
Keywords: Group actions, dynamics, diffeomorphisms, 1-manifolds, circle.

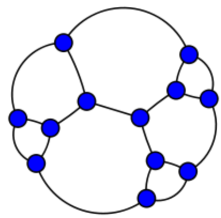


Figure 1: The Frucht graph.

Another modification of Cayley's construction allows realizing every countable group as a group of homeomorphisms of the Cantor set. Assume for a while that such a group G is infinite, and endow the space $M := \{0,1\}^G = \{\varphi : G \to \{0,1\}\}$ with the product topology (metric). Then M becomes a Cantor set, and G faithfully acts on M by shifting coordinates: $L_g(\varphi)(h) := \varphi(g^{-1}h)$. Despite of its apparent simplicity, this shift action is fundamental in several contexts, and has attracted the attention of many people over the last decades Boyle, Lind, and Rudolph [1988], Cohen [2017], Hochman [2010], and Popa [2006b].

In the case where G is finite, one can modify the previous construction just by adding extra coordinates to the space $\{0,1\}^G$ on which the action is trivial. More interestingly, there is a single "small" group of homeomorphisms of the Cantor set that contains all finite groups. To properly define it, we see the Cantor set as the boundary at infinite $\partial\Gamma$ of a regular tree Γ of degree 3. Every proper, clopen ball in $\partial\Gamma$ can be canonically seen as the boundary at infinite of a rooted tree. We then consider the set of automorphisms of $\partial\Gamma$ that arise by cutting $\partial\Gamma$ into finitely many clopen balls and sending them into the pieces of another partition of $\partial\Gamma$ into clopen balls (with the same number of pieces), so that the restriction to each such ball is nothing but the canonical identification between these pieces viewed as boundaries of rooted trees. This yields the so-called *Thompson's group V*, which, among many remarkable properties, is finitely presented and simple. (See Cannon, Floyd, and Parry [1996] for more on this.) It is easy to see that V contains all finite groups.

Having realized every countable group as a group of homeomorphisms of a 0-dimensional space, one can ask whether some restriction arises when passing to higher dimension. Certainly, there are number of other motivations for considering this framework, perhaps the most transparent one coming from foliation theory. Indeed, to every group action of a finitely-generated group by homeomorphisms of a manifold M, one can associate a foliation by the classical procedure of suspension as follows: Letting $g_1, \ldots .g_k$ be a system of generators of G, we consider S_k, the surface of genus k, with fundamental group

$$\pi_1(S_k) = \left\langle a_1, \dots, a_k, b_1, \dots, b_k : \prod_{i=1}^{k} [a_i, b_i] = id \right\rangle.$$

In there, the generators a_i are freely related, hence there is a homomorphism $\phi \colon \pi_1(S_k) \to G$ sending a_i into g_i and b_i into the identity. We then consider the product space $\Delta \times M$ endowed with the action of $\pi_1(S_k)$ given by $h(x, y) = (\bar{h}(x), \phi(h)(y))$, where $\bar{h}$ stands for the deck transformation on the Poincaré disc Δ associated to h. The quotient under this action is naturally a foliated, fibrated space with basis S_k and fiber M, and the holonomy group of this foliation coincides with G. (See Candel and Conlon [2000] for more on this construction.)

By the discussion above, and despite some remarkable recent progress Brown, Fisher, and Hurtado [2016, 2017], it seems impossible to develop a full theory of groups acting on manifolds. Here we restrict the discussion to the simplest case, namely, actions on 1-dimensional spaces. In this context, the ordered structure of the phase space allows developing a very complete theory for actions by homeomorphisms, and the techniques coming from 1-dimensional dynamics allow the same for actions by diffeomorphisms. For each of such settings there are good references with very complete panoramas of the developments up to recent years: see Deroin, Navas, and Rivas [2017] and Ghys [2001b] and Navas [2011b], respectively. This is the reason why we prefer to focus on challenging problems that remain unsolved, hoping that the reader will become motivated to work on some of them.

1 Actions of Kazhdan's groups

In 1967, Kazhdan introduced a cohomological property and proved that it is satisfied by higher-rank simple Lie groups and their lattices, as for instance $\mathrm{SL}(n, \mathbb{Z})$ for $n \geq 3$ and their finite index subgroups Každan [1967]. Since discrete groups satisfying this property are necessarily finitely generated, he proved finite generation for these lattices, thus solving a longstanding question. Since then, the so-called *Kazhdan's property (T)* has become one of the most important tools for studying actions and representations of Lie groups.

Although Kazhdan's original definition is somewhat technical, there is a more geometric property later introduced by Serre which turns out to be equivalent in the locally compact setting: a group satisfies *Serre's property (FH)* if every action by (affine) isometries on a Hilbert space has an invariant vector. (See Bekka, de la Harpe, and Valette [2008] and G. A. Margulis [1991] for a full discussion on this.)

Property (T) has very strong consequences for the dynamics of group actions in different settings; see for instance Bader, Furman, Gelander, and Monod [2007], Furman [1999], Popa [2006a], Shalom [2006], and Zimmer [1984]. In what concerns actions on

1-dimensional spaces, a classical result pointing in this direction, due to Watatani and Alperin, states that every action of a group with property (T) by isometries of a real tree has a fixed point. In a more dynamical framework, Witte–Morris proved in Witte [1994] the following remarkable result: For $n \geq 3$, every action of a finite-index subgroup of $\mathrm{SL}(n, \mathbb{Z})$ by orientation-preserving homeomorphisms of the interval (resp. the circle) is trivial (resp. has finite image).

This theorem is even more remarkable because of its proof, which is amazingly elementary. However, it strongly relies on the existence of certain nilpotent subgroups inside the lattice, which do not arise in other (*e.g.* cocompact) cases. Despite some partial progress in this direction Lifschitz and Morris [2004, 2008] and Morris [2014] (see Morris [2011] for a full panorama on this), the following question remains open.

Question 1. Does there exist a lattice in a higher-rank simple Lie group admitting a nontrivial action by orientation-preserving homeomorphisms of the interval ?

Notice that the statement above doesn't deal with actions on the circle. This is due to a theorem of Ghys [1999], which reduces the general case to that on the interval: Every action of a lattice in a higher-rank simple Lie group by orientation-preserving homeomorphisms of the circle has a finite orbit (hence a finite-index subgroup -which is still a lattice-fixes some interval).

The question above can be rephrased in the more general setting of Kazhdan groups.

Question 2. Does there exist an infinite, finitely-generated Kazhdan group of circle homeomorphisms ?

Question 3. Does there exist a nontrivial (hence infinite) Kazhdan group of orientation-preserving homeomorphisms of the interval ?

A concrete result on this concerns actions by diffeomorphisms: If a finitely-generated group of $C^{3/2}$ circle diffeomorphisms satisfies property (T), then it is finite Navas [2002, 2011b] (see Cornulier [2017] for the piecewise-smooth case). However, the situation is unclear in lower regularity. For instance, the group $G := \mathrm{SL}(2, \mathbb{Z}) \ltimes \mathbb{Z}^2$ has the *relative property (T)* (in the sense that for every action of G by isometries of a Hilbert space, there is a vector that is invariant by $\mathbb{Z}^2$), yet it naturally embeds into the group of circle homeomorphisms. Indeed, the group $\mathrm{SL}(2, \mathbb{Z})$ acts projectively on the 2-fold covering of S^1 -which is still a circle-, and blowing up an orbit one can easily insert an equivariant $\mathbb{Z}^2$-action. However, no action of this group is C^1 smoothable Navas [2005, 2010a].

The example above can be easily modified as follows: Letting $\mathbb{F}_2 \subset \mathrm{SL}(2, \mathbb{Z})$ be a finite-index subgroup, the semidirect product $G := \mathbb{F}_2 \ltimes \mathbb{Z}^2$ still has the relative property (T) (with respect to $\mathbb{Z}^2$). Moreover, starting with a free group of diffeomorphisms of the interval and using the blowing up procedure along a countable orbit, one can easily embed

G into the group of orientation-preserving homeomorphisms of the interval. Because of these examples, the answers to both Question 2 and 3 remain unclear.

There is a different, more dynamical approach to Question 1 above. Indeed, when dealing with the continuous case, actions on the interval and actions on the real line are equivalent. In the latter context, an easy argument shows that for every action of a finitely-generated group G by orientation-preserving homeomorphisms of the real line without global fixed points, one of the following three possibilities occurs:

(i) There is a σ-finite measure μ that is invariant under the action.

(ii) The action is semiconjugate to a minimal action for which every small-enough interval is sent into a sequence of intervals that converge to a point under well-chosen group elements, but this property does not hold for every bounded interval. (Here, by a *semiconjugacy* we roughly mean a factor action for which the factor map is a continuous, non-decreasing, proper map of the real line.)

(iii) The action is semiconjugate to a minimal one for which the contraction property above holds for all bounded intervals.

Observe that a group may have actions of different type. (A good exercise is to build actions of $\mathbb{F}_2$ of each type.) In case (i), the translation number homomorphism $g \mapsto \mu([x, g(x)[)$ provides a nontrivial homomorphism from G into $\mathbb{R}$. In case (ii), it is not hard to see that, when looking at the minimal semiconjugate action, the map φ that sends x into the supremum of the points $y > x$ for which the interval $[x, y]$ can be contracted along group elements is an orientation-preserving homeomorphism of the real line that commutes with all elements of G and satisfies $\varphi(x) > x$ for all x. Therefore, there is an induced G-action on the corresponding quotient space $\mathbb{R}/\sim$, where $x \sim \varphi(x)$, which is a topological circle.

By the discussion above, case (i) cannot arise for infinite groups with property (T). As a direct consequence of Ghys' theorem stated above, case (ii) can neither arise for faithful actions of lattices in higher-rank simple Lie groups. Hence, if such a group admits an action (without global fixed points) on the real line, the action must satisfy property (iii).

Question 4. Does there exist an infinite, finitely-generated group that acts on the real line all of whose actions by orientation-preserving homeomorphisms of the line without global fixed points are of type (iii) ?

2 Cones and orders on groups

Groups of orientation-preserving homeomorphisms of the real line are left orderable, that is, they admit total order relations that are invariant under left multiplication. Indeed, such a group can be ordered by prescribing a dense sequence (x_n) of points in the line, and letting $f \prec g$ if the smallest n for which $f(x_n) \neq g(x_n)$ is such that $f(x_n) < g(x_n)$. Conversely, for a countable left-orderable group, it is not hard to produce an action on the line. (See Deroin, Navas, and Rivas [2017] and Ghys [2001b] for more on this.) This may fail, however, for uncountable groups with cardinality equal to that of $\mathrm{Homeo}_+(\mathbb{R})$; see Mann [2015].

The characterization above yields to a dynamical approach for the theory of left-orderable groups (which goes back to Dedekind and Hölder). In this view, a useful idea independently introduced by Ghys [2001a] and Sikora [2004] consists in endowing the space $\mathcal{LO}(G)$ of all left-orders of a given left-orderable group G with the Chabauty topology. (Two orders are close if they coincide over a "large" finite subset.) This provides a totally disconnected, compact space, which is metrizable in case G is countable. (One can let $\mathrm{dist}(\prec, \prec') = 1/n$, where n is the largest integer such that $\preceq$ and $\preceq'$ coincide over the set A_n of a prescribed exhaustion of $G = \cup_i A_i$ by finite subsets.)

A result of Linnell establishes that spaces of left orders are either finite or uncountable Linnell [2011] (see also Clay, Mann, and Rivas [2017]). Left-orderable groups with finitely many orders were classified by Tararin (see Deroin, Navas, and Rivas [2017] and Kopytov and Medvedev [1996]): they are all solvable, the simplest examples being $\mathbb{Z}$ and the Klein bottle group $\langle a, b : bab = a \rangle$. In an opposite direction, for some classes of groups G, it is known that no left order is isolated in $\mathcal{LO}(G)$: solvable groups with infinitely many left orders Rivas and Tessera [2016], free groups Hermiller and Šunić [2017], Kielak [2015], McCleary [1985], Navas [2010b], and Rivas [2012], free products of groups Rivas [2012], and surface groups Alonso, Brum, and Rivas [2017]. The following question remains, however, open.

Question 5. Does there exist a finitely-generated, amenable, left-orderable group having an isolated order inside an infinite space of left orders ?

It is somewhat surprising that several classes of groups with infinitely many left orders do admit isolated left orders. Constructions have been proposed by different authors using quite distinct techniques: dynamical, group theoretical and combinatorial (see for instance Dehornoy [2014], Dubrovina and Dubrovin [2001], Ito [2013, 2016, 2017], Matsumoto [2017], and Navas [2011a]). However, the most striking examples remain the first ones, namely, the braid groups B_n. To be more precise, let

$$B_n := \langle \sigma_1, \ldots, \sigma_{n-1} : \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}, \sigma_i \sigma_j = \sigma_j \sigma_i \text{ for } |i - j| \geq 2 \rangle$$

be the standard presentation of B_n. Denote $a_i := (\sigma_i \cdots \sigma_{n-1})^{(-1)^{i-1}}$, where $i \in \{1, \ldots, n-1\}$. Building on seminal work of Dehornoy [2000], Dubrovina and Dubrovin showed in Dubrovina and Dubrovin [2001] that B_n admits the disjoint decomposition

$$B_n = \langle a_1, \ldots, a_{n-1} \rangle^+ \cup \langle a_1^{-1}, \ldots, a_{n-1}^{-1} \rangle^+ \cup \{id\},$$

where $\langle \cdot \rangle^{+}$ stands for the semigroup generated by the corresponding set of elements. An easy argument then shows that the order whose elements larger than the identity are those in $\langle a_1, \ldots, a_{n-1} \rangle^{+}$ is well defined, total and left invariant; more importantly, it is isolated, since it is the only left order for which the elements $a_1, \ldots, a_{n-1}$ are all larger than the identity Linnell [2011].

Despite the apparent simplicity of the previous decomposition into finitely-generated positive and negative *cones*, no elementary proof is available. Finding an elementary approach is a challenging problem. The only nontrivial case that is well understood is that of $n = 3$, where the decomposition is evident from the picture below.

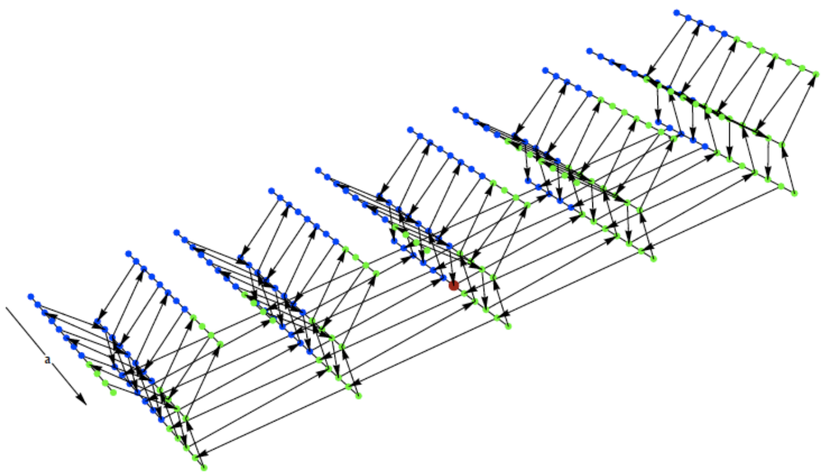


Figure 2: The positive and negative cones of an isolated order on the Cayley graph of $B_3 = \langle a, b \colon ba^2b = a \rangle$ with respect to the generators $a := \sigma_1 \sigma_2$ and $b := \sigma_2^{-1}$.

Notice that a general left-orderable group G acts on $\mathcal{LO}(G)$ by conjugacy: given a left order $\preceq$ and $g \in G$, the conjugate of $\preceq$ under g is the left order $\preceq_g$ for which $f_1 \preceq_g f_2$ if and only if $g f_1 g^{-1} \prec g f_2 g^{-1}$, which is equivalent to $f_1 g^{-1} \prec f_2 g^{-1}$.

Question 6. Does there exist a finitely-generated, left-orderable group for which the conjugacy action on its space of left orders is minimal (that is, all the orbits are dense) ?

It is not hard to show that free groups do admit left orders with a dense orbit under the conjugacy action McCleary [1985] and Rivas [2012]. However, this action is not minimal. Indeed, free groups are bi-orderable (that is, they admit left orders that are also invariant under right multiplication), and a bi-order is, by definition, a fixed point for the conjugacy action.

The conjugacy action was brilliantly used by Morris [2006] to settle a question of Linnell [2001], which was priorly raised -in the language of foliations- by Thurston [1974]: Every finitely-generated, left-orderable, amenable group admits a nontrivial homomorphism into $\mathbb{Z}$. Indeed, amenability provides an invariant probability measure for the conjugacy action. (Remind that one of the many definitions of *amenability* is that every action on a compact space admits an invariant probability measure.) Then the key idea is that, by the Poincaré recurrence theorem, left orders in the support of such a measure must satisfy a certain recurrence property, which reads as an algebraic property that is close to the Archimedean one (the so-called *Conradian property*; see Navas and Rivas [2009]). This property allows obtaining the desired homomorphism. The next extension of the question (also proposed by Linnell), which is reminiscent of the Tits alternative, remains open.

Question 7. Does there exist a finitely-generated, left-orderable group without free subgroups and admitting no nontrivial homomorphism into $\mathbb{Z}$?

It is worth stressing that a negative answer to Question 4 above would imply a negative one to Question 7. Indeed, on the one hand, an action of type (i) provides a group homomorphism into $\mathbb{R}$ (via translation numbers), hence into $\mathbb{Z}$ for finitely-generated groups. On the other hand, as explained before, an action of type (ii) factors through a locally contracting action on the circle, which implies the presence of a free subgroup by a theorem of G. Margulis [2000] (see also Ghys [2001b]).

A closely related question is the following.

Question 8. Does there exist a finitely-generated, left-orderable group G with no nontrivial homomorphism into $\mathbb{Z}$ and trivial group of bounded cohomology $\mathrm{H}^2_b(G, \mathbb{R})$?

Again, a negative answer to Question 4 would also imply a negative one to this question. Indeed, locally contracting actions on S^1 are parameterized (up to semiconjugacy) by a nontrivial cohomological class taking values in $\{0, 1\}$, according to a seminal work of Ghys [1987] (see also Ghys [2001b]).

Besides these questions addressed for particular families of left-orderable groups, obstructions to left-orderability that go beyond torsion-freenes or the so-called *unique product property* UPP (namely, for each finite subset of the group, there is at least one product of two elements in this set that cannot be represented as another product of two elements in the set) are poorly understood. Although this goes beyond the scope of this text (and is one of the main lines of research of the theory of left-orderable groups), for specific families of dynamically defined groups, this has wide interest. A particular question on this was raised and nicely discussed by Calegari [2009].

Question 9. Is the group of orientation-preserving homeomorphisms of the 2-disk that are the identity at the boundary left orderable ?

Notice that the group in discussion is torsion free, due to a classical result of de Kerékjártó [1941] (see also Kolev [2006]). It is worth to stress that it is even unknown whether this group satisfies the UPP.

3 Groups of piecewise-projective homeomorphisms

Groups of piecewise-affine homeomorphisms have been deeply studied in relation to Thompson's groups. Remind that Thompson's group T is defined as the subgroup of (the previously introduced group) V formed by the elements that respect the cyclic order in $\partial\Gamma$, the boundary of the homogeneous tree of degree-3. Besides, Thompson's group F is the subgroup of T formed by the elements that fix a specific point in $\partial\Gamma$ (say, the left-most point of the boundary of a clopen ball). Another realization arises in dimension 1: T (resp. F) is a group of orientation-preserving, piecewise-affine homeomorphisms of the circle (rep. interval). Both groups are finitely presented Cannon, Floyd, and Parry [1996] and have a dyadic nature, in the sense that the slopes of elements are integer powers of 2, and break points are dyadic rationals. One of the most challenging questions on these groups is the following.

Question 10. Is Thompson's group F amenable ?

A beautiful result of Brin and Squier establishes that the group of piecewise-affine homeomorphisms of the interval (hence F) doesn't contain free subgroups Brin and Squier [1985]. However, despite much effort by several people over the last decades (which includes several mistaken announcements pointing in the two possible directions), Question 10 remains as a kind of nightmare for the mathematical community; see Ghys [2009].

Besides the well-known question above, the algebraic structure of certain groups of piecewise-affine homeomorphisms, mainly generalizations of Thompson's groups Stein [1992], is quite interesting. A concrete problem on them deals with distorted elements. To properly state it, remind that, given a group G with a finite generating system $\mathcal{G}$, the *word length* $\|g\|$ of $g \in G$ is the minimum number of factors needed to write g as a product of elements in $\mathcal{G}$ (and their inverses). An element of infinite order $g \in G$ is said to be *distorted* if

$$\lim_{n\to\infty} \frac{\|g^n\|}{n} = 0.$$

More generally, an element is said to be distorted in a general group whenever it is distorted inside a finitely-generated subgroup of this group.

Distorted elements naturally appear inside nilpotent groups, and have been extensively used to study rigidity phenomena of group actions on 2-manifolds Calegari and Freedman [2006] and Franks and Handel [2006]. In the 1-dimensional setting, Avila proved that irrational rotations are distorted (in a very strong way) in the group of C^∞ diffeomorphism

Avila [2008]. However, despite some partial progress Guelman and Liousse [2017], the following question is open.

Question 11. Does the group of piecewise-affine circle homeomorphisms contain distorted elements ?

Beyond the piecewise-affine setting, the group of piecewise-projective homeomorphisms is a larger source of relevant examples of finitely-generated groups. In this direction, one could ask whether examples yielding to an affirmative answer to Question 3 may arise inside the group of piecewise-projective homeomorphisms of the line.

As a concrete example of an interesting group, remind that Thompson's group T itself has a natural piecewise-projective, non piecewise-affine realization (which goes back to Thurston and, independently, to Ghys and Sergiescu): just replace dyadic rationals by rationals via the Minkowsky mark function, and change piecewise-affine maps by maps that are piecewise in $\mathrm{PSL}(2,\mathbb{Z})$.

Among the new examples of groups constructed via this approach, the most remarkable is, with no doubt, the group G_{LM} introduced by Lodha and Moore [2016], which -as a group acting on the line- is generated by the homeomorphisms f, g, h below (notice that f, g generate a group isomorphic -actually, conjugate- to Thompson's group F): $f(t) := t+1$,

$$g(t) := \begin{cases} t & \text{if } t \leq 0, \\ \frac{t}{1-t} & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 3 - \frac{1}{t} & \text{if } \frac{1}{2} \leq t \leq 1, \\ t+1 & \text{if } t \geq 1, \end{cases} \quad \text{and} \quad h(t) := \begin{cases} \frac{2t}{1+t} & \text{if } t \in [0,1], \\ t & \text{otherwise.} \end{cases}$$

Among its many remarkable properties, G_{LM} has no free subgroup (by an easy extension of Brin–Squier's theorem mentioned above), it is non-amenable (due to prior work of Ghys and Carrière [1985] and Monod [2013]), and has a finite presentation (this is the main technical contribution of Lodha and Moore [2016]). Although it is not the first example of a group with these properties (see Ol'shanskii and Sapir [2003]), it has several other properties, as being torsion-free and of type F_∞ (a property which is much stronger than being finitely presented; see Lodha [2014]). Recently, based on previous work relating smoothness with 1-dimensional hyperbolic dynamics in the solvable context Bonatti, Monteverde, Navas, and Rivas [2017], this group was proven to be non C^1 smoothable in Bonatti, Lodha, and Triestino [2017] (see Lodha [2017] for a related result concerning a group closely related to T). This is to be compared with a classical result of Ghys and

Sergiescu, according to which T is topologically conjugate to a group of C^∞ diffeomorphisms Ghys and Sergiescu [1987]. The following general question becomes natural.

Question 12. What are the groups of piecewise-projective homeomorphisms of the interval/circle that are topologically conjugate to groups of C^1 diffeomorphisms ?

4 The spectrum of sharp regularities for group actions

Regularity issues appear as fundamental when dealing with both the dynamical properties of a given action and the algebraic constraints of the acting group. The original source of this goes back to Denjoy's classical theorem: Every C^2 circle diffeomorphism without periodic points is minimal. The C^2 hypothesis or, at least, a derivative with bounded variation, is crucial for this result. (The theorem is false in the $C^{1+\alpha}$ setting Herman [1979] and Tsuboi [1995], and remains unknown for diffeomorphisms whose derivatives are τ-continous with respect to the modulus of continuity $\tau(x) = |x \log(x)|$.) This is the reason why, when dealing with group actions (and, more generally, codimension-1 foliations), such an hypothesis is usually made. Nevertheless, in recent years, many new phenomena have been discovered in different regularities, thus enriching the theory.

One of the main problems to deal with in this direction is that of the optimal regularity. This problem is twofold. On the one hand, one looks for the maximal regularity that can be achieved, under topological conjugacy, of a given action. On the other hand, one asks for the maximal regularity in which a given group can faithfully act by varying the topological dynamics. A very concrete question in the latter direction is the following.

Question 13. Given $0 < \alpha < \beta < 1$, does there exist a finitely-generated group of $C^{1+\alpha}$ diffeomorphisms of the circle/interval that does not embed into the group of $C^{1+\beta}$ diffeomorphisms ?

There are concrete reasons for restricting this problem only to regularities between C^1 and C^2. On the one hand, Kim and Koberda have recently settled the analog of Question 13 for regularities larger than C^2, whereas the (discrete) Heisenberg group faithfully acts by $C^{1+\alpha}$ diffeomorphisms for any $\alpha < 1$ Castro, Jorquera, and Navas [2014], but it does not embed into the group of C^2 diffeomorphisms Plante and Thurston [1976]. On the other hand, Thurston gave the first examples of groups that are non C^1 smoothable via his remarkable stability theorem Thurston [1974] (see also Bonatti, Monteverde, Navas, and Rivas [2017] and Navas [2010a]), while examples of groups of C^1 diffeomorphisms that are non $C^{1+\alpha}$ smoothable arise in relation to growth of groups Navas [2008]. In an opposite direction, every countable group of circle homeomorphisms is topologically conjugate to a group of Lipschitz homeomorphisms, as it is shown below following the arguments of Deroin, Kleptsyn, and Navas [2007] and Navas [2014].

Example 1. Let G be a group with a finite, symmetric generating system $\mathcal{G}$ which acts by homeomorphisms of a compact 1-manifold M. Let Leb denote the normalized Lebesgue measure on M. Given $\varepsilon > 0$, let $\bar{\mu}_\epsilon$ be the measure on M defined as

$$\bar{\mu}_\varepsilon := \sum_{f \in G} e^{-\varepsilon \| f \|} f_*(Leb),$$

where $\| f \|$ denotes the word-length of g (with respect to $\mathcal{G}$). The measure $\bar{\mu}_\varepsilon$ has finite mass for ε large enough. Indeed,

$$\bar{\mu}_\varepsilon(M) = \sum_{n \geq 0} e^{-n\varepsilon} \big| S(n) \big| \leq \sum_{n \geq 0} e^{-n\varepsilon} |\mathcal{G}| \big(|\mathcal{G}| - 1 \big)^{n-1} = \frac{|\mathcal{G}|}{|\mathcal{G}| - 1} \sum_{n \geq 0} \left(\frac{|\mathcal{G}| - 1}{e^\varepsilon} \right)^n,$$

where $|S(n)|$ stands for the cardinal of the set $S(n)$ of elements having word-length equal to n. Moreover, since for every $g \in \mathcal{G}$ and all $f \in G$ it holds $\|gf\| \leq \|f\| + 1$, we have

$$\text{(1)} \qquad g_*(\bar{\mu}_\varepsilon) = \sum_{f \in G} e^{-\varepsilon \| f \|} (gf)_*(Leb) \leq e^\varepsilon \sum_{f \in G} e^{-\varepsilon \| gf \|} (gf)_*(Leb) = e^\varepsilon \bar{\mu}_\varepsilon.$$

Let μ_ε be the normalization of $\bar{\mu}_\varepsilon$. The probability measure μ_ε has total support and no atoms. It is hence topologically equivalent to Leb (dimension 1 is crucial here; see Harrison [1975, 1979] for examples of non-smoothable homeomorphisms in higher dimension). By a change of coordinates sending μ_ε into Leb, relation (1) becomes, for each interval $I \subset M$,

$$|g^{-1}(I)| = g_*(Leb)(I) \leq e^\varepsilon Leb(I) = e^\varepsilon |I|.$$

This means that, in these new coordinates, g^{-1} is Lipschitz with constant $\leq e^\varepsilon$.

It is interesting to specialize Question 12 to nilpotent group actions. Indeed, these actions are known to be C^1 smoothable Farb and Franks [2003], Jorquera [2012], and Parkhe [2016], though they are non C^2 smoothable unless the group is abelian Plante and Thurston [1976]. Moreover, the only settled case for Question 12 is that of a nilpotent group, namely, the group G_4 of 4×4 upper-triangular matrices with integer entries and 1's in the diagonal. In concrete terms, G_4 embeds into the group of $C^{1+\alpha}$ diffeomorphisms of the interval for every $\alpha < 1/2$, though it does not embed for $\alpha > 1/2$ Jorquera, Navas, and Rivas [2018]. (The case $\alpha = 1/2$ remains open; compare Navas [2013].)

5 Zero Lebesgue measure for exceptional minimal sets

Differentiability issues are crucial in regard to ergodic type properties for actions. It follows from Denjoy's theorem quoted above that a single C^2 circle diffeomorphism cannot admit an *exceptional minimal set*, that is, a minimal invariant set homeomorphic to the Cantor set. Although the mathematical community took some time to realize that these sets may actually appear for group actions (the first explicit example appears in Sacksteder [1964]), it is worth pointing out that these sets naturally arise, for instance, for Fuchsian groups (and, more generally, for groups with a Schottky dynamics), as well as for certain semiconjugates of Thompson's group T. The following question is due to Ghys and Sullivan.

Question 14. Let G be a finitely-generated group of C^2 circle diffeomorphisms. Assume that G admits an exceptional minimal set Λ. Is the Lebesgue measure of Λ equal to zero ?

An important recent progress towards the solution of this question is made in Deroin, Kleptsyn, and Navas [n.d.], where it is answered in the affirmative for groups of real-analytic diffeomorphisms. Besides, an affirmative answer is provided, also in the real-analytic context, to another important question due to Hector.

Question 15. Let G be a finitely-generated group of C^2 circle diffeomorphisms. Assume that G admits an exceptional minimal set Λ. Is the set of orbits of intervals of $\mathrm{S}^1 \setminus \Lambda$ finite ?

Beyond having settled these two questions in the real-analytic context, the main contribution of Deroin, Kleptsyn, and Navas [ibid.] consists in proposing new ideas yielding to structure results for groups of circle diffeomorphisms admitting an exceptional minimal set (see Deroin, Filimonov, Kleptsyn, and Navas [2017] and the references therein for a full discussion on this). Indeed, so far, positive answers to these questions were known only in the expanding case (that is, whenever for every $x \in \Lambda$ there is $g \in G$ such that $Dg(x) > 1$; see Navas [2004b]), and for Markovian like dynamics Cantwell and Conlon [1989, 1988]. What is clear now is that, in the non expanding case, a certain Markovian structure must arise (see Deroin [2017] for a precise result in this direction in the conformal case). This view should also be useful to deal with the following classical question (conjecture) of Dippolito [1978].

Question 16. Let G be a finitely-generated group of C^2 circle diffeomorphisms. Assume that G admits an exceptional minimal set Λ. Is the restriction of the action of G to Λ topologically conjugated to the action of a group of piecewise-affine homeomorphisms ?

It should be pointed out that Questions 14, 15 and 16 have natural analogs for codimension-1 foliations. In this broader context, they all remain widely open, even in the (transversely) real-analytic setting. However, the ideas and techniques from Deroin, Kleptsyn, and Navas [n.d.] show that, also in this generality, structural issues are the right tools to deal with them.

6 Ergodicity of minimal actions

In case of minimal actions, a subtle issue concerns *ergodicity* (with respect to the Lebesgue measure), that is, the nonexistence of measurable invariant sets except for those having zero or full (Lebesgue) measure. The original motivation for this comes from a theorem independently proved by Katok and Hasselblatt [1995] and Herman [1979]: The action of a C^2 circle diffeomorphism without periodic points is ergodic (with respect to the Lebesgue measure). Notice that this result does not follow from Denjoy's theorem (which only ensures minimality), since we know from the seminal work of Arnold [1961] that the (unique) invariant measure may be singular with respect to the Lebesgue measure.

Katok's proof is performed via a classical *control of distortion* technique, which means that there is a uniform control on the ratio $\sup Df_n / \inf Df_n$ for the value of the derivatives on certain intervals along a well-chosen sequence of compositions f_n. This allows transferring geometric data from micro to macro scales, so that the proportion of the measures of different sets remains controlled when passing from one scale to another. Clearly, this avoids the existence of invariant sets of intermediate measure, thus proving ergodicity.

Herman–Katok's theorem deals with an "elliptic" context, whereas several classical ergodicity-like results (going back to Poincaré's linearization theorem) hold in an hyperbolic context. One hopes that a careful combination of both techniques would yield to an affirmative answer to the next question, also due to Ghys and Sullivan.

Question 17. Let G be a finitely-generated group of C^2 circle diffeomorphisms. If the action of G is minimal, is it necessarily ergodic with respect to the Lebesgue measure ?

So far, an affirmative answer to this question is known in the case where the group is generated by elements that are C^2 close to rotations Navas [2004b], for expanding actions Navas [2004b] and Deroin, Kleptsyn, and Navas [2009], and for groups of real-analytic diffeomorphisms which are either free Deroin, Kleptsyn, and Navas [n.d.] or have infinitely many ends Alvarez, Filimonov, Kleptsyn, Malicet, Cotón, Navas, and Triestino [2015]. It is worth pointing out that the C^2 regularity hypothesis is crucial here; see for instance Kodama and Matsumoto [2013].

Again, Question 17 has a natural extension to the framework of codimension-1 foliations, where it remains widely open.

7 Absolute continuity of the stationary measure

Due to the absence of invariant measures for general groups actions, a useful tool to consider are the *stationary measures*, which correspond to probability measures that are invariant in mean. More precisely, given a probability distribution p on a (say, finitely generated) group G that acts by homeomorphisms of a compact metric space M, a probability measure μ on M is said to be stationnary with respecto to p if for every measurable subset $A \subset M$, one has

$$\mu(A) = \sum_{g \in G} p(g)\mu(g^{-1}(A)).$$

There are always stationary measures: this follows from a fixed-point argument or from Krylov–Bogoliuvob's argument consisting in taking means and passing to the limit in the (compact) space of probabilities on M. A crucial property is that, in case of uniqueness of the stationary measure (with respect to a given p), the action is ergodic with respect to this measure Navas [2011b].

It is shown in Antonov [1984] and Deroin, Kleptsyn, and Navas [2007] that, for a group of orientation-preserving circle homeomorphisms G acting minimally, the stationary measure is unique with respect to each probability distribution p on G that is *non-degenerate* (*i.e.* the support of the measure generates G as a semigroup). Hence, the next problem becomes relevant in relation to Question 17 above.

Question 18. Let G be a finitely-generated group of C^2 orientation-preserving circle diffeomorphisms. Does there exist a non-degenerate probability distribution on G for which the stationary measure is absolutely continuous with respect to the Lebesgue measure ?

A classical argument of "balayage" due to Furstenberg [1971] solves this question for lattices in $\mathrm{PSL}(2,\mathbb{R})$. However, this strongly relies on the geometry of the Poincaré disk, and does not extend to general groups. Moreover, it is shown in Deroin, Kleptsyn, and Navas [2009] and Guivarc'h and Le Jan [1990] (see also Blachère, Haïssinsky, and Mathieu [2011]) that the resulting probability measure is singular with respect to the Lebesgue measure for non cocompact lattices whenever the distribution p is *symmetric* (*i.e.* $p(g) = p(g^{-1})$ for all $g \in G$) and finitely supported (and, more generally, for distributions with finite first moment). This also holds for groups with a Markovian dynamics for which there are *non-expandable points* (*i.e.* points x such that $Dg(x) \leq 1$ for all $g \in G$), as for instance (the smooth realizations of) Thompson's group T. (Notice that, for the canonical action of $\mathrm{PSL}(2,\mathbb{Z})$, the point $[1:0]$ is non-expandable.)

Once again, Question 18 extends to the framework of codimension-1 foliations, where it remains widely open. (Uniqueness of the stationary measure in this setting and, more generally, in a transversely conformal framework, is the main content of Deroin and Kleptsyn [2007].)

A probability distribution on a group induces a random walk on it, many of whose properties reflect algebraic features of the group and translate into particular issues of the stationary measures. Remind that to every probability distribution one can associate a "maximal boundary", which, roughly, is a measurable space endowed with a "contracting" action having a unique stationary measure so that any other space of this type is a measurable factor of it Furstenberg [1973]. The study of this *Poisson–Furstenberg boundary* is one of the main topics in this area, and explicit computations are, in general, very hard Erschler [2010]. In our framework, a valuable result in this direction was obtained by Deroin, who proved in Deroin [2013] that for every group of smooth-enough circle diffeomorphisms with no finite orbit and whose action is locally discrete in a strong (and very precise) sense, the Poisson–Furstenberg boundary identifies with the circle endowed with the corresponding stationary measure provided the probability distribution on the group satisfies a certain finite-moment condition. Extending this result to more general groups is a challenging problem. In particular, the next question remains unsolved.

Question 19. Given a symmetric, finitely-supported, non-degenerate probability distribution on Thompson's group T, does the Poisson–Furstenberg boundary of T with respect to it identifies with the circle endowed with the corresponding stationary measure ?

Last but not least, random walks are also of interest for groups acting on the real line. In this setting, a nontrivial result is the existence of a (nonzero) σ-finite stationary measure for symmetric distributions on finitely-generated groups Deroin, Kleptsyn, Navas, and Parwani [2013]. This is closely related to general recurrence type results for symmetric random walks on the line. One hopes that these ideas may be useful in dealing with Question 1, though no concrete result in this direction is known yet.

8 Structural stability and the space of representations

Another aspect in which differentiability issues crucially appear concerns stability. Remind that, given positive numbers $r \leq s$, an action by C^s diffeomorphisms is said to be C^r structurally stable if every perturbation that is small enough in the C^s topology is C^r conjugate to it. (In the case $s = 0$, we allow semiconjugacies instead of conjugacies.) Usually, structural stability arises in hyperbolic contexts, and the situation in an elliptic type framework is less clear. The next question was formulated by Rosenberg more than 40 years ago (see for instance Rosenberg and Roussarie [1975]).

Question 20. Does there exist a faithful action of $\mathbb{Z}^2$ by C^∞ orientation-preserving circle diffeomorphisms that is C^∞ structurally stable ?

A closely related question, also due to Rosenberg, concerns the topology of the space of $\mathbb{Z}^2$-actions.

Question 21. Given $r \geq 1$, is the subset of $\mathrm{Diff}^r_+(\mathrm{S}^1)^2$ consisting of pairs of commuting diffeomorphisms locally connected ?

These two questions have inspired very deep work of many people, including Herman and Yoccoz, who devoted their thesis to closely related problems. However, despite all these efforts, they remain widely open. Among some recent progress concerning them, we can mention the proof of the connectedness of the space of commuting C^∞ diffeomorphisms of the closed interval Bonatti and Eynard-Bontemps [2016] (which, in its turn, has important consequences for codimension-1 foliations Eynard-Bontemps [2016]), and that of the path connectedness of the space of commuting C^1 diffeomorphisms of either the circle or the closed interval Navas [2014]. These two results apply in general to actions of $\mathbb{Z}^n$.

In a non-Abelian context, several other questions arise in relation to the structure of the space of actions. Among them, we can stress a single one concerning actions with an exceptional minimal set, for which the results from Deroin, Kleptsyn, and Navas [n.d.] point in a positive direction.

Question 22. Given a faithful action ϕ_0 of a finitely generated group G by C^∞ circle diffeomorphisms admitting an exceptional minimal set, does there exist a path ϕ_t of faithful actions of G that is continuous in the C^∞ topology and starts with $\phi_0 = \phi$ so that each ϕ_t admits an exceptional minimal set for $t < 1$ and ϕ_1 is minimal ?

Quite surprisingly, structural stability is interesting even in the continuous setting. Indeed, the dictionary between left orders and actions on the interval shows that such an action is structurally stable if and only if a certain canonical left order arising from it is isolated in whole the space of left orders. Similarly, an action on the circle is structurally stable if and only if a natural "cyclic order" induced from it is an isolated point in the corresponding space of cyclic orders (endowed with the appropriate Chabauty topology; see Mann and Rivas [n.d.]). In this regard, we may ask the following. (Compare Question 5.)

Question 23. Let G be a finitely-generated group of circle homeomorphisms whose action is C^0 structurally stable. Suppose that G admits infinitely many non semiconjugate actions on the circle. Does G contain a free subgroup in two generators ?

9 Approximation by conjugacy and single diffeomorphisms

Some of the connectedness results discussed above are obtained by constructing paths of conjugates of a given action. This idea is particularly simple and fruitful in very low regularity. We next give a quite elementary example to illustrate this.

Example 2. As is well known, every circle homeomorphism has zero topological entropy. In most textbooks, this is proved by an easy counting argument of separated orbits for such an f. However, to show this, we can also follow the arguments of Example 1 for $G = \langle f \rangle$ and any $\varepsilon > 0$. Indeed, the outcome is that f is topologically conjugate to a Lipschitz homeomorphism with Lipschitz constant $\leq e^{\varepsilon}$. By the invariance of entropy under topological conjugacy and its classical estimate in terms of the logarithm of the Lipschitz constant, we obtain that $h_{\text{top}}(f) \leq \varepsilon$. Since this holds for all $\varepsilon > 0$, we must have $h_{\text{top}}(f) = 0$.

The naive argument above still works for groups of subexponential growth, as for instance nilpotent groups Navas [2014]. Therefore, for all actions of these groups on 1-manifolds, the *geometric entropy* (as defined by Ghys, Langevin, and Walczak [1988]) always equals zero. More interestingly, a similar strategy should be useful to deal with groups of C^1 diffeomorphisms. Indeed, in this context, Hurder has shown in Hurder [2000] that zero entropy is a consequence of the absence of *resilient pairs*, which means that there are no elements f, g such that $x < f(x) < f(y) < g(x) < g(y) < y$ for certain points x, y. (Notice that the converse holds even for homeomorphisms, as is follows from a classical counting argument.) The proof of this fact is quite involved, and one hopes for an affirmative answer to the question below, which would immediately imply this result.

Question 24. Let G be a finitely-generated group of C^1 circle diffeomorphisms. Suppose that G has no resilient pairs. Given $\varepsilon > 0$, can G be conjugated (by a homeomorphism) into a group of Lipschitz homeomorphisms for which the Lipschitz constants of the generators are all $\leq e^{\varepsilon}$?

A particularly clarifying example on this concerns conjugates of C^1 diffeomorphisms without periodic points (that is, with irrational rotation number), as explained below.

Example 3. Remind that every cocycle $\varphi : M \to \mathbb{R}$ with respect to a continuous map $f : M \to M$ is cohomologous to each of its Birkhoff means. Indeed, letting

$$\psi_n := \frac{1}{n} \sum_{i=0}^{n-1} S_i \varphi, \qquad \text{where} \qquad S_n \varphi := \sum_{i=0}^{n-1} \varphi \circ f^i \quad \text{and} \quad S_0 \varphi := 0,$$

one easily checks the identity

$$\varphi - \frac{S_n \varphi}{n} = \psi_n - \psi_n \circ f.$$

If f belongs to $\mathrm{Diff}^1_+(\mathrm{S}^1)$, we can specialize this remark to $\varphi := \log(Df)$. Besides, if f has irrational rotation number ρ, then an easy argument shows that $S_n(\log Df)/n \to 0$. Therefore,

$$\psi_n \circ f + \log Df - \psi_n \longrightarrow 0.$$

Adding a constant c_n to ψ_n, we may assume that ψ_n coincides with $\log Dh_n$ for a C^1 diffeomorphism h_n. The relation above then becomes $\log D(h_n f h_n^{-1}) \circ h_n \to 0$, which shows that $h_n f h_n^{-1}$ converges to R_ρ in the C^1 topology.

The argument above extends to actions of nilpotent groups, thus giving an affirmative answer to Question 25 for these groups Navas [2014]. However, this idea strongly uses the additive nature of the logarithm of the derivative, and it seems hard to directly extend it to higher regularity. Despite this, one expects that the use of a Schwarzian-like derivative (cocycle) would yield to an affirmative answer to the following question.

Question 25. Let f be a C^2 circle diffeomorphism of irrational rotation number ρ. Does the set of C^2 conjugates of f contain the rotation R_ρ in its C^2-closure ?

The discussion above reveals that many natural questions still remain unsolved for single diffeomorphisms. Below we state two more of them.

Question 26. For which values of $r > 1$ there exists $s \geq r$ such that for every C^s circle diffeomorphism f of irrational rotation number ρ, the sequence f^{q_n} converges to the identity in the C^r topology, where (q_n) is the sequence of denominators of the rational approximations of ρ ?

This question is inspired by a fundamental result of Herman [1979], according to which one has the convergence $f^{q_n} \to Id$ in the C^1 topology for C^2 circle diffeomorphisms of irrational rotation number (see also Navas and Triestino [2013] and Yoccoz [1984a]). The answer to this question should consider a result of Yoccoz [1995], who constructed a C^∞ circle diffeomorphism with irrational rotation number and trivial centralizer.

Question 27. Let f be a C^2 circle diffeomorphism of irrational rotation number ρ. Given $\varepsilon > 0$, let M_ε be the *mapping torus* of f over $S^1 \times [0, \varepsilon]$, that is, the surface obtained by identifying $(x, 0) \sim (f(x), \varepsilon)$. Let $\rho(\varepsilon) \in \mathbb{R}/\mathbb{Z}$ be such that M_ε corresponds to the elliptic curve $\mathbb{C}/(\mathbb{Z} + i\rho(\varepsilon)\mathbb{Z})$. Does $\rho(\varepsilon)$ converges to ρ as $\varepsilon \to 0$?

This question is due to Arnold. One hopes that recent progress on fine properties of circle diffeomorphisms should lead to a positive solution of it.

10 Topological invariance of the Godbillon–Vey class

The group of circle diffeomorphisms supports a remarkable cohomology class, namely, the *Godbillon–Vey class*, which is represented by the cocycle

$$(f, g) \mapsto \int_{S^1} \log(Df)\, D(\log D(g \circ f)).$$

Notice that, though this formula requires two derivatives, it can be naturally extended to $C^{3/2+\varepsilon}$ diffeomorphisms (just pass half of the derivative from right to left; see Hurder and Katok [1990] and Tsuboi [1992]). However, no extension to $C^{1+\alpha}$ diffeomorphisms is possible for α small Tsuboi [1989, 1995].

According to a well-known result of Gelfand and Fuchs, the continuous cohomology of the whole group of C^∞ circle diffeomorphisms is generated by two classes: the Euler class (which is the single generator in the C^1 setting), and the Godbillon–Vey class. Obviously, the Godbillon–Vey class induces (by restriction) a class in $H^2(G,\mathbb{R})$ for every group G of $C^{3/2+\varepsilon}$ circle diffeomorphisms. We refer to Ghys [1989] and Hurder [2002] and the references therein for a panorama on this, including a full discussion on the next open question.

Question 28. Is the (restriction of the) Godbillon–Vey class invariant under topological conjugacy for groups of C^2 diffeomorphisms ?

A first result in the positive direction was established by Raby, who proved invariance under conjugacy by C^1 diffeomorphisms Raby [1988]. Very soon after that, an alternative proof for this fact was proposed by Ghys and Tsuboi [1988]. Some years later, in Hurder and Katok [1990]. Hurder and Katok proved invariance under conjugacies that are absolutely continuous (with an absolutely continuous inverse); see Hilsum [2015] for a recent result in the same direction.

Ghys–Tsuboi's proof of Raby's theorem is of a dynamical nature. Indeed, in the most relevant cases of this framework, what it is proved is that C^1 conjugacies between groups of C^r diffeomorphisms are automatically C^r provided $r \geq 2$. This applies for instance to non-Abelian groups whose action is minimal.

It is not hard to extend Ghys–Tsuboi's theorem to (bi-)Lipschitz conjugacies Navas [2007, 2011b]. However, absolutely continuous conjugacies are harder to deal with.

Question 29. What are the groups of C^2 circle diffeomorphisms acting minimally for which the normalizer inside the group of absolutely continuous homeomorphisms coincides with that inside the group of diffeomorphisms ?

11 On groups of real-analytic diffeomorphisms

The real-analytic framework offers new problems of wide interest even in the classical context. In this regard, remind that a celebrated result proved by Yoccoz [1984b] establishes that every real-analytic circle homeomorphism is minimal provided it has an irrational rotation number, thus extending Denjoy's theorem to this setting. (The same holds for C^∞ homeomorphisms with non-flat singularities.) Ghys has asked whether this extends to the case where singularities may also arise for the inverse of the map.

Question 30. Does Denjoy's theorem hold for circle homeomorphisms whose graphs are real-analytic ?

In a cohomological setting, another question concerns the validity of Geldfand–Fuch's theorem in the real-analytic case.

Question 31. Is the continuous cohomology of the group $\mathrm{Diff}^{\omega}_{+}(\mathrm{S}^1)$ of orientation-preserving, real-analytic circle diffeomorphisms generated by the Euler and the Godbillon–Vey classes ?

A negative answer to this question would require the construction of a cocycle that uses real-analyticity in a crucial way. This would be somehow similar to Mather's homomorphism defined on the group of C^1 circle diffeomorphisms with derivatives having bounded variation. Remind that this is defined as

$$f \mapsto \int_{\mathrm{S}^1} [D(\log Df)]_{\mathrm{reg}},$$

where $[D(\log Df)]_{\mathrm{reg}}$ stands for the regular part of the signed measure obtained as the derivative (in the sense of distributions) of the (finite total variation) function $\log Df$; see Mather [1985]. Such a homomorphism cannot exist in other regularities, because the corresponding groups of diffeomorphisms are known to be simple Mather [1974, 1975], except for class C^2 Mather [1984]. By the way, though this is not related to real-analytic issues, this critical case must appear in any list of selected problems on the subject.

Question 32. Is the group of orientation-preserving C^2 circle diffeomorphisms simple ?

Finally, we would like to focus on finitely-generated subgroups of $\mathrm{Diff}^{\omega}_{+}(\mathrm{S}^1)$. These have a tendency to exhibiting a much more rigid behavior than groups of diffeomorphisms. For example, though Thompson's groups act by C^∞ diffeomorphisms, the group F (hence T) does not faithfully act by real-analytic diffeomorphisms. One way to see this is by looking at solvable subgroups: F contains such groups in arbitrary degree of solvability, though solvable groups of real-analytic diffeomorphisms of either the interval or the circle are matabelian Ghys [1993]. (See however Navas [2004a] for algebraic constraints that apply to solvable groups of C^2 diffeomorphisms.)

Quite surprisingly, many algebraic issues that are known to hold or not to hold in the setting of C^∞ diffeomorphisms are open in the real-analytic setting. For instance, it is unknown whether irrational rotations are distorted elements in $\mathrm{Diff}^{\omega}_{+}(\mathrm{S}^1)$. A more striking open question concerns the famous *Tits alternative*.

Question 33. Does the Tits alternative hold in $\mathrm{Diff}^{\omega}_{+}(\mathrm{S}^1)$? More precisely, does every non-metabelian subgroup of this group contain a free subgroup ?

We refer to Farb and Shalen [2002] for a partial result that reduces the general case to that of the interval. Notice that F provides a negative answer to this question for groups of C^∞ diffeomorphisms because of the aforementioned Ghys–Sergiescu's C^∞ realization and Brin–Squier's theorem.

Added in proof: Question 13 has been recently solved by Kim and Koberda [2017]. I wish to thank Bassam Fayad and Sang-Hyun Kim for their remarks and corrections to an earlier version of this text.

References

Juan Alonso, Joaquín Brum, and Cristóbal Rivas (2017). "Orderings and flexibility of some subgroups of $Homeo_+(\mathbb{R})$". *J. Lond. Math. Soc. (2)* 95.3, pp. 919–941. MR: 3664524 (cit. on p. 2058).

Sébastien Alvarez, Dmitry Filimonov, Victor Kleptsyn, Dominique Malicet, Carlos Meniño Cotón, Andrés Navas, and Michele Triestino (2015). "Groups with infinitely many ends acting analytically on the circle". arXiv: 1506.03839 (cit. on p. 2066).

V. A. Antonov (1984). "Modeling of processes of cyclic evolution type. Synchronization by a random signal". *Vestnik Leningrad. Univ. Mat. Mekh. Astronom.* vyp. 2, pp. 67–76. MR: 756386 (cit. on p. 2067).

V. I. Arnold (1961). "Small denominators. I. Mapping the circle onto itself". *Izv. Akad. Nauk SSSR Ser. Mat.* 25, pp. 21–86. MR: 0140699 (cit. on p. 2066).

Artur Avila (2008). "Distortion elements in $Diff^\infty(R/Z)$". arXiv: 0808.2334 (cit. on p. 2062).

Uri Bader, Alex Furman, Tsachik Gelander, and Nicolas Monod (2007). "Property (T) and rigidity for actions on Banach spaces". *Acta Math.* 198.1, pp. 57–105. MR: 2316269 (cit. on p. 2055).

Bachir Bekka, Pierre de la Harpe, and Alain Valette (2008). *Kazhdan's property (T)*. Vol. 11. New Mathematical Monographs. Cambridge University Press, Cambridge, pp. xiv+472. MR: 2415834 (cit. on p. 2055).

Sébastien Blachère, Peter Haïssinsky, and Pierre Mathieu (2011). "Harmonic measures versus quasiconformal measures for hyperbolic groups". *Ann. Sci. Éc. Norm. Supér. (4)* 44.4, pp. 683–721. MR: 2919980 (cit. on p. 2067).

C. Bonatti, I. Monteverde, A. Navas, and C. Rivas (2017). "Rigidity for C^1 actions on the interval arising from hyperbolicity I: solvable groups". *Math. Z.* 286.3-4, pp. 919–949. MR: 3671566 (cit. on pp. 2062, 2063).

Christian Bonatti and Hélène Eynard-Bontemps (2016). "Connectedness of the space of smooth actions of $\mathbb{Z}^n$ on the interval". *Ergodic Theory Dynam. Systems* 36.7, pp. 2076–2106. MR: 3568972 (cit. on p. 2069).

Christian Bonatti, Yash Lodha, and Michele Triestino (2017). "Hyperbolicity as an obstruction to smoothability for one-dimensional actions". arXiv: 1706.05704 (cit. on p. 2062).

Mike Boyle, Douglas Lind, and Daniel Rudolph (1988). "The automorphism group of a shift of finite type". *Trans. Amer. Math. Soc.* 306.1, pp. 71–114. MR: 927684 (cit. on p. 2054).

Matthew G. Brin and Craig C. Squier (1985). "Groups of piecewise linear homeomorphisms of the real line". *Invent. Math.* 79.3, pp. 485–498. MR: 782231 (cit. on p. 2061).

Aaron Brown, David Fisher, and Sebastian Hurtado (2016). "Zimmer's conjecture: Subexponential growth, measure rigidity, and strong property (T)". arXiv: 1608.04995 (cit. on p. 2055).

– (2017). "Zimmer's conjecture for actions of $\mathrm{SL}(m, \mathbb{Z})$". arXiv: 1710.02735 (cit. on p. 2055).

Danny Calegari (2009). "Orderability, and groups of homeomorphisms of the disk". *Geometry and the Imagination* (cit. on p. 2060).

Danny Calegari and Michael H. Freedman (2006). "Distortion in transformation groups". *Geom. Topol.* 10. With an appendix by Yves de Cornulier, pp. 267–293. MR: 2207794 (cit. on p. 2061).

Alberto Candel and Lawrence Conlon (2000). *Foliations. I*. Vol. 23. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xiv+402. MR: 1732868 (cit. on p. 2055).

J. W. Cannon, W. J. Floyd, and W. R. Parry (1996). "Introductory notes on Richard Thompson's groups". *Enseign. Math. (2)* 42.3-4, pp. 215–256. MR: 1426438 (cit. on pp. 2054, 2061).

John Cantwell and Lawrence Conlon (1988). "Foliations and subshifts". *Tohoku Math. J. (2)* 40.2, pp. 165–187. MR: 943817 (cit. on p. 2065).

– (1989). "Leaves of Markov local minimal sets in foliations of codimension one". *Publ. Mat.* 33.3, pp. 461–484. MR: 1038484 (cit. on p. 2065).

Gonzalo Castro, Eduardo Jorquera, and Andrés Navas (2014). "Sharp regularity for certain nilpotent group actions on the interval". *Math. Ann.* 359.1-2, pp. 101–152. MR: 3201895 (cit. on p. 2063).

Adam Clay, Kathryn Mann, and Cristóbal Rivas (2017). "On the number of circular orders on a group". arXiv: 1704.06242 (cit. on p. 2058).

David Bruce Cohen (2017). "The large scale geometry of strongly aperiodic subshifts of finite type". *Adv. Math.* 308, pp. 599–626. MR: 3600067 (cit. on p. 2054).

Yves Cornulier (2017). "Commensurated actions for groups with piecewise properties". Preprint (cit. on p. 2056).

Harold Scott Macdonald Coxeter, Roberto Frucht, and David L. Powers (1981). *Zero-symmetric graphs*. Trivalent graphical regular representations of groups. Academic

Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, pp. ix+170. MR: 658666 (cit. on p. 2053).

Patrick Dehornoy (2000). *Braids and self-distributivity*. Vol. 192. Progress in Mathematics. Birkhäuser Verlag, Basel, pp. xx+623. MR: 1778150 (cit. on p. 2058).

– (2014). "Monoids of O-type, subword reversing, and ordered groups". *J. Group Theory* 17.3, pp. 465–524. MR: 3200370 (cit. on p. 2058).

B. Deroin, V. Kleptsyn, A. Navas, and K. Parwani (2013). "Symmetric random walks on $\mathrm{Homeo}^+(\mathbf{R})$". *Ann. Probab.* 41.3B, pp. 2066–2089. MR: 3098067 (cit. on p. 2068).

B. Deroin, A. Navas, and C. Rivas (2017). "Groups, Orders, and Dynamics". arXiv: 1408.5805 (cit. on pp. 2055, 2058).

Bertrand Deroin (2013). "The Poisson boundary of a locally discrete group of diffeomorphisms of the circle". *Ergodic Theory Dynam. Systems* 33.2, pp. 400–415. MR: 3035290 (cit. on p. 2068).

– (2017). "Rigidity of self-similar sets of the complex line". In: Preprint (cit. on p. 2065).

Bertrand Deroin, Dmitry Filimonov, Victor Kleptsyn, and Andrés Navas (2017). "A paradigm for codimension one foliations". In: *Geometry, dynamics, and foliations 2013*. Vol. 72. Adv. Stud. Pure Math. Math. Soc. Japan, Tokyo, pp. 59–69. MR: 3726705 (cit. on p. 2065).

Bertrand Deroin and Victor Kleptsyn (2007). "Random conformal dynamical systems". *Geom. Funct. Anal.* 17.4, pp. 1043–1105. MR: 2373011 (cit. on p. 2067).

Bertrand Deroin, Victor Kleptsyn, and Andrés Navas (n.d.). "On the ergodic theory of free group actions by real-analytic circle diffeomorphisms". In: *Invent. Math*. arXiv: 1312.4133 (cit. on pp. 2065, 2066, 2069).

– (2007). "Sur la dynamique unidimensionnelle en régularité intermédiaire". *Acta Math.* 199.2, pp. 199–262. MR: 2358052 (cit. on pp. 2063, 2067).

– (2009). "On the question of ergodicity for minimal group actions on the circle". *Mosc. Math. J.* 9.2, 263–303, back matter. MR: 2568439 (cit. on pp. 2066, 2067).

Paul R. Dippolito (1978). "Codimension one foliations of closed manifolds". *Ann. of Math. (2)* 107.3, pp. 403–453. MR: 0515731 (cit. on p. 2065).

T. V. Dubrovina and N. I. Dubrovin (2001). "On braid groups". *Mat. Sb.* 192.5, pp. 53–64. MR: 1859702 (cit. on p. 2058).

Anna Erschler (2010). "Poisson-Furstenberg boundaries, large-scale geometry and growth of groups". In: *Proceedings of the International Congress of Mathematicians. Volume II*. Hindustan Book Agency, New Delhi, pp. 681–704. MR: 2827814 (cit. on p. 2068).

Hélène Eynard-Bontemps (2016). "On the connectedness of the space of codimension one foliations on a closed 3-manifold". *Invent. Math.* 204.2, pp. 605–670. MR: 3489706 (cit. on p. 2069).

Benson Farb and John Franks (2003). "Groups of homeomorphisms of one-manifolds. III. Nilpotent subgroups". *Ergodic Theory Dynam. Systems* 23.5, pp. 1467–1484. MR: 2018608 (cit. on p. 2064).

Benson Farb and Peter Shalen (2002). "Groups of real-analytic diffeomorphisms of the circle". *Ergodic Theory Dynam. Systems* 22.3, pp. 835–844. MR: 1908556 (cit. on p. 2074).

John Franks and Michael Handel (2006). "Distortion elements in group actions on surfaces". *Duke Math. J.* 131.3, pp. 441–468. MR: 2219247 (cit. on p. 2061).

Alex Furman (1999). "Gromov's measure equivalence and rigidity of higher rank lattices". *Ann. of Math. (2)* 150.3, pp. 1059–1081. MR: 1740986 (cit. on p. 2055).

Harry Furstenberg (1971). "Random walks and discrete subgroups of Lie groups". In: *Advances in Probability and Related Topics, Vol. 1*. Dekker, New York, pp. 1–63. MR: 0284569 (cit. on p. 2067).

– (1973). "Boundary theory and stochastic processes on homogeneous spaces", pp. 193–229. MR: 0352328 (cit. on p. 2068).

É Ghys (2001a). *Personal comunnication* (cit. on p. 2058).

– (2009). *Ce groupe m'agace* (cit. on p. 2061).

É. Ghys, R. Langevin, and P. Walczak (1988). "Entropie géométrique des feuilletages". *Acta Math.* 160.1-2, pp. 105–142. MR: 926526 (cit. on p. 2070).

É. Ghys and T. Tsuboi (1988). "Différentiabilité des conjugaisons entre systèmes dynamiques de dimension 1". *Ann. Inst. Fourier (Grenoble)* 38.1, pp. 215–244. MR: 949007 (cit. on p. 2072).

Étienne Ghys (1987). "Groupes d'homéomorphismes du cercle et cohomologie bornée". In: *The Lefschetz centennial conference, Part III (Mexico City, 1984)*. Vol. 58. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 81–106. MR: 893858 (cit. on p. 2060).

– (1989). "L'invariant de Godbillon-Vey". *Astérisque* 177-178. Séminaire Bourbaki, Vol. 1988/89, Exp. No. 706, 155–181. MR: 1040572 (cit. on p. 2072).

– (1993). "Sur les groupes engendrés par des difféomorphismes proches de l'identité". *Bol. Soc. Brasil. Mat. (N.S.)* 24.2, pp. 137–178. MR: 1254981 (cit. on p. 2073).

– (1999). "Actions de réseaux sur le cercle". *Invent. Math.* 137.1, pp. 199–231. MR: 1703323 (cit. on p. 2056).

– (2001b). "Groups acting on the circle". *Enseign. Math. (2)* 47.3-4, pp. 329–407. MR: 1876932 (cit. on pp. 2055, 2058, 2060).

Étienne Ghys and Yves Carrière (1985). "Relations d'équivalence moyennables sur les groupes de Lie". *C. R. Acad. Sci. Paris Sér. I Math.* 300.19, pp. 677–680. MR: 802650 (cit. on p. 2062).

Étienne Ghys and Vlad Sergiescu (1987). "Sur un groupe remarquable de difféomorphismes du cercle". *Comment. Math. Helv.* 62.2, pp. 185–239. MR: 896095 (cit. on p. 2063).

Nancy Guelman and Isabelle Liousse (2017). "Distortion in groups of affine interval exchange transformations". arXiv: 1705.00144 (cit. on p. 2062).

Yves Guivarc'h and Yves Le Jan (1990). "Sur l'enroulement du flot géodésique". *C. R. Acad. Sci. Paris Sér. I Math.* 311.10, pp. 645–648. MR: 1081425 (cit. on p. 2067).

Jenny Harrison (1975). "Unsmoothable diffeomorphisms". *Ann. of Math. (2)* 102.1, pp. 85–94. MR: 0388458 (cit. on p. 2064).

– (1979). "Unsmoothable diffeomorphisms on higher dimensional manifolds". *Proc. Amer. Math. Soc.* 73.2, pp. 249–255. MR: 516473 (cit. on p. 2064).

Michael-Robert Herman (1979). "Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations". *Inst. Hautes Études Sci. Publ. Math.* 49, pp. 5–233. MR: 538680 (cit. on pp. 2063, 2066, 2071).

Susan Hermiller and Zoran Šunić (2017). "No positive cone in a free product is regular". *Internat. J. Algebra Comput.* 27.8, pp. 1113–1120. MR: 3741005 (cit. on p. 2058).

Michel Hilsum (2015). "Fonctions à variation bornée et classe de Godbillon-Vey". *Q. J. Math.* 66.2, pp. 547–562. MR: 3356837 (cit. on p. 2072).

Michael Hochman (2010). "On the automorphism groups of multidimensional shifts of finite type". *Ergodic Theory Dynam. Systems* 30.3, pp. 809–840. MR: 2643712 (cit. on p. 2054).

S. Hurder and A. Katok (1990). "Differentiability, rigidity and Godbillon-Vey classes for Anosov flows". *Inst. Hautes Études Sci. Publ. Math.* 72, 5–61 (1991). MR: 1087392 (cit. on p. 2072).

Steven Hurder (2000). *Entropy and dynamics of C^1 foliations*. Preprint (cit. on p. 2070).

– (2002). "Dynamics and the Godbillon-Vey class: a history and survey". In: *Foliations: geometry and dynamics (Warsaw, 2000)*. World Sci. Publ., River Edge, NJ, pp. 29–60. MR: 1882764 (cit. on p. 2072).

Tetsuya Ito (2013). "Dehornoy-like left orderings and isolated left orderings". *J. Algebra* 374, pp. 42–58. MR: 2998793 (cit. on p. 2058).

– (2016). "Construction of isolated left orderings via partially central cyclic amalgamation". *Tohoku Math. J. (2)* 68.1, pp. 49–71. MR: 3476136 (cit. on p. 2058).

– (2017). "Isolated orderings on amalgamated free products". *Groups Geom. Dyn.* 11.1, pp. 121–138. MR: 3641836 (cit. on p. 2058).

Eduardo Jorquera (2012). "A universal nilpotent group of C^1 diffeomorphisms of the interval". *Topology Appl.* 159.8, pp. 2115–2126. MR: 2902746 (cit. on p. 2064).

Eduardo Jorquera, Andrés Navas, and Cristóbal Rivas (2018). "On the sharp regularity for arbitrary actions of nilpotent groups on the interval: the case of N_4". *Ergodic Theory Dynam. Systems* 38.1, pp. 180–194. MR: 3742542 (cit. on p. 2064).

Anatole Katok and Boris Hasselblatt (1995). *Introduction to the modern theory of dynamical systems*. Vol. 54. Encyclopedia of Mathematics and its Applications. With a supplementary chapter by Katok and Leonardo Mendoza. Cambridge University Press, Cambridge, pp. xviii+802. MR: 1326374 (cit. on p. 2066).

D. A. Každan (1967). "On the connection of the dual space of a group with the structure of its closed subgroups". *Funkcional. Anal. i Priložen.* 1, pp. 71–74. MR: 0209390 (cit. on p. 2055).

B. de Kerékjártó (1941). "Sur les groupes compacts de transformations topologiques des surfaces". *Acta Math.* 74, pp. 129–173. MR: 0006549 (cit. on p. 2061).

Dawid Kielak (2015). "Groups with infinitely many ends are not fraction groups". *Groups Geom. Dyn.* 9.1, pp. 317–323. MR: 3343355 (cit. on p. 2058).

Sang-hyun Kim and Thomas Koberda (Nov. 2017). "Diffeomorphism groups of critical regularity". arXiv: 1711.05589 (cit. on p. 2074).

Hiroki Kodama and Shigenori Matsumoto (2013). "Minimal C^1-diffeomorphisms of the circle which admit measurable fundamental domains". *Proc. Amer. Math. Soc.* 141.6, pp. 2061–2067. MR: 3034431 (cit. on p. 2066).

Boris Kolev (2006). "Sous-groupes compacts d'homéomorphismes de la sphère". *Enseign. Math. (2)* 52.3-4, pp. 193–214. MR: 2300606 (cit. on p. 2061).

ValeriĭM. Kopytov and NikolaĭYa. Medvedev (1996). *Right-ordered groups*. Siberian School of Algebra and Logic. Consultants Bureau, New York, pp. x+250. MR: 1393199 (cit. on p. 2058).

Lucy Lifschitz and Dave Witte Morris (2004). "Isotropic nonarchimedean S-arithmetic groups are not left orderable". *C. R. Math. Acad. Sci. Paris* 339.6, pp. 417–420. MR: 2092755 (cit. on p. 2056).

– (2008). "Bounded generation and lattices that cannot act on the line". *Pure Appl. Math. Q.* 4.1, Special Issue: In honor of Grigory Margulis. Part 2, pp. 99–126. MR: 2405997 (cit. on p. 2056).

Peter A. Linnell (2001). "Left ordered groups with no non-abelian free subgroups". *J. Group Theory* 4.2, pp. 153–168. MR: 1812322 (cit. on p. 2060).

– (2011). "The space of left orders of a group is either finite or uncountable". *Bull. Lond. Math. Soc.* 43.1, pp. 200–202. MR: 2765563 (cit. on pp. 2058, 2059).

Yash Lodha (2014). "A nonamenable type F_∞ group of piecewise projective homeomorphisms". arXiv: 1408.3127 (cit. on p. 2062).

– (2017). "A finitely presented infinite simple group of homeomorphisms of the circle". arXiv: 1710.06220 (cit. on p. 2062).

Yash Lodha and Justin Tatch Moore (2016). "A nonamenable finitely presented group of piecewise projective homeomorphisms". *Groups Geom. Dyn.* 10.1, pp. 177–200. MR: 3460335 (cit. on p. 2062).

Kathryn Mann (2015). "Left-orderable groups that don't act on the line". *Math. Z.* 280.3-4, pp. 905–918. MR: 3369358 (cit. on p. 2058).

Kathryn Mann and Cristobal Rivas (n.d.). "Group orderings, dynamics, and rigidity". In: Ann. Inst. Fourier (Grenoble). arXiv: 1607.00054 (cit. on p. 2069).

G. A. Margulis (1991). *Discrete subgroups of semisimple Lie groups*. Vol. 17. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, pp. x+388. MR: 1090825 (cit. on p. 2055).

Gregory Margulis (2000). "Free subgroups of the homeomorphism group of the circle". *C. R. Acad. Sci. Paris Sér. I Math.* 331.9, pp. 669–674. MR: 1797749 (cit. on p. 2060).

John N. Mather (1974). "Commutators of diffeomorphisms". *Comment. Math. Helv.* 49, pp. 512–528. MR: 0356129 (cit. on p. 2073).

– (1975). "Commutators of diffeomorphisms. II". *Comment. Math. Helv.* 50, pp. 33–40. MR: 0375382 (cit. on p. 2073).

– (1984). "A curious remark concerning the geometric transfer map". *Comment. Math. Helv.* 59.1, pp. 86–110. MR: 743944 (cit. on p. 2073).

– (1985). "Commutators of diffeomorphisms. III. A group which is not perfect". *Comment. Math. Helv.* 60.1, pp. 122–124. MR: 787665 (cit. on p. 2073).

Shigenori Matsumoto (2017). "Dynamics of isolated orders". Preprint. arXiv: 1701.05987 (cit. on p. 2058).

Stephen H. McCleary (1985). "Free lattice-ordered groups represented as o-2 transitive l-permutation groups". *Trans. Amer. Math. Soc.* 290.1, pp. 69–79. MR: 787955 (cit. on pp. 2058, 2059).

Nicolas Monod (2013). "Groups of piecewise projective homeomorphisms". *Proc. Natl. Acad. Sci. USA* 110.12, pp. 4524–4527. MR: 3047655 (cit. on p. 2062).

Dave Witte Morris (2006). "Amenable groups that act on the line". *Algebr. Geom. Topol.* 6, pp. 2509–2518. MR: 2286034 (cit. on p. 2060).

– (2011). "Can lattices in $\mathrm{SL}(n,\mathbb{R})$ act on the circle?" In: *Geometry, rigidity, and group actions*. Chicago Lectures in Math. Univ. Chicago Press, Chicago, IL, pp. 158–207. MR: 2807831 (cit. on p. 2056).

– (2014). "Some arithmetic groups that do not act on the circle". In: *Geometric group theory*. Vol. 21. IAS/Park City Math. Ser. Amer. Math. Soc., Providence, RI, pp. 207–248. MR: 3329729 (cit. on p. 2056).

Andrés Navas (2002). "Actions de groupes de Kazhdan sur le cercle". *Ann. Sci. École Norm. Sup. (4)* 35.5, pp. 749–758. MR: 1951442 (cit. on p. 2056).

– (2004a). "Groupes résolubles de difféomorphismes de l'intervalle, du cercle et de la droite". *Bull. Braz. Math. Soc. (N.S.)* 35.1, pp. 13–50. MR: 2057043 (cit. on p. 2073).

– (2004b). "Sur les groupes de difféomorphismes du cercle engendrés par des éléments proches des rotations". *Enseign. Math. (2)* 50.1-2, pp. 29–68. MR: 2084334 (cit. on pp. 2065, 2066).

– (2005). "Quelques nouveaux phénomènes de rang 1 pour les groupes de difféomorphismes du cercle". *Comment. Math. Helv.* 80.2, pp. 355–375. MR: 2142246 (cit. on p. 2056).
– (2007). "Three remarks on one dimensional bi-Lipschitz conjugacies". Preprint. arXiv: 0705.0034 (cit. on p. 2072).
– (2008). "Growth of groups and diffeomorphisms of the interval". *Geom. Funct. Anal.* 18.3, pp. 988–1028. MR: 2439001 (cit. on p. 2063).
– (2010a). "A finitely generated, locally indicable group with no faithful action by C^1 diffeomorphisms of the interval". *Geom. Topol.* 14.1, pp. 573–584. MR: 2602845 (cit. on pp. 2056, 2063).
– (2010b). "On the dynamics of (left) orderable groups". *Ann. Inst. Fourier (Grenoble)* 60.5, pp. 1685–1740. MR: 2766228 (cit. on p. 2058).
– (2011a). "A remarkable family of left-ordered groups: central extensions of Hecke groups". *J. Algebra* 328, pp. 31–42. MR: 2745552 (cit. on p. 2058).
– (2011b). *Groups of circle diffeomorphisms*. Spanish. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, pp. xviii+290. MR: 2809110 (cit. on pp. 2055, 2056, 2067, 2072).
– (2013). "On centralizers of interval diffeomorphisms in critical (intermediate) regularity". *J. Anal. Math.* 121, pp. 1–30. MR: 3127376 (cit. on p. 2064).
– (2014). "Sur les rapprochements par conjugaison en dimension 1 et classe C^1". *Compos. Math.* 150.7, pp. 1183–1195. MR: 3230850 (cit. on pp. 2063, 2069–2071).
Andrés Navas and Cristóbal Rivas (2009). "A new characterization of Conrad's property for group orderings, with applications". *Algebr. Geom. Topol.* 9.4. With an appendix by Adam Clay, pp. 2079–2100. MR: 2551663 (cit. on p. 2060).
Andrés Navas and Michele Triestino (2013). "On the invariant distributions of C^2 circle diffeomorphisms of irrational rotation number". *Math. Z.* 274.1-2, pp. 315–321. MR: 3054331 (cit. on p. 2071).
Alexander Yu Ol'shanskii and Mark V Sapir (2003). "Non-amenable finitely presented torsion-by-cyclic groups". *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* 96.1, pp. 43–169 (cit. on p. 2062).
Kiran Parkhe (2016). "Nilpotent dynamics in dimension one: structure and smoothness". *Ergodic Theory Dynam. Systems* 36.7, pp. 2258–2272. MR: 3568980 (cit. on p. 2064).
J. F. Plante and W. P. Thurston (1976). "Polynomial growth in holonomy groups of foliations". *Comment. Math. Helv.* 51.4, pp. 567–584. MR: 0436167 (cit. on pp. 2063, 2064).
Sorin Popa (2006a). "Some computations of 1-cohomology groups and construction of non-orbit-equivalent actions". *J. Inst. Math. Jussieu* 5.2, pp. 309–332. MR: 2225044 (cit. on p. 2055).

Sorin Popa (2006b). "Some rigidity results for non-commutative Bernoulli shifts". *J. Funct. Anal.* 230.2, pp. 273–328. MR: 2186215 (cit. on p. 2054).

Gilles Raby (1988). "Invariance des classes de Godbillon-Vey par C^1-diffeomorphismes". *Ann. Inst. Fourier (Grenoble)* 38.1, pp. 205–213. MR: 949006 (cit. on p. 2072).

Cristóbal Rivas (2012). "Left-orderings on free products of groups". *J. Algebra* 350, pp. 318–329. MR: 2859890 (cit. on pp. 2058, 2059).

Cristóbal Rivas and Romain Tessera (2016). "On the space of left-orderings of virtually solvable groups". *Groups Geom. Dyn.* 10.1, pp. 65–90. MR: 3460331 (cit. on p. 2058).

Harold Rosenberg and Robert Roussarie (1975). "Some remarks on stability of foliations". *J. Differential Geometry* 10, pp. 207–219. MR: 0380826 (cit. on p. 2068).

Richard Sacksteder (1964). "On the existence of exceptional leaves in foliations of co-dimension one". *Ann. Inst. Fourier (Grenoble)* 14.fasc. 2, pp. 221–225. MR: 0174060 (cit. on p. 2065).

Yehuda Shalom (2006). "The algebraization of Kazhdan's property (T)". In: *International Congress of Mathematicians. Vol. II*. Eur. Math. Soc., Zürich, pp. 1283–1310. MR: 2275645 (cit. on p. 2055).

Adam S. Sikora (2004). "Topology on the spaces of orderings of groups". *Bull. London Math. Soc.* 36.4, pp. 519–526. MR: 2069015 (cit. on p. 2058).

Melanie Stein (1992). "Groups of piecewise linear homeomorphisms". *Trans. Amer. Math. Soc.* 332.2, pp. 477–514. MR: 1094555 (cit. on p. 2061).

William P. Thurston (1974). "A generalization of the Reeb stability theorem". *Topology* 13, pp. 347–352. MR: 0356087 (cit. on pp. 2060, 2063).

Takashi Tsuboi (1989). "On the foliated products of class C^1". *Ann. of Math. (2)* 130.2, pp. 227–271. MR: 1014925 (cit. on p. 2072).

– (1992). "Area functionals and Godbillon-Vey cocycles". *Ann. Inst. Fourier (Grenoble)* 42.1-2, pp. 421–447. MR: 1162569 (cit. on p. 2072).

– (1995). "Homological and dynamical study on certain groups of Lipschitz homeomorphisms of the circle". *J. Math. Soc. Japan* 47.1, pp. 1–30. MR: 1304186 (cit. on pp. 2063, 2072).

Dave Witte (1994). "Arithmetic groups of higher $\mathbf{Q}$-rank cannot act on 1-manifolds". *Proc. Amer. Math. Soc.* 122.2, pp. 333–340. MR: 1198459 (cit. on p. 2056).

J.-C. Yoccoz (1984a). "Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne". *Ann. Sci. École Norm. Sup. (4)* 17.3, pp. 333–359. MR: 777374 (cit. on p. 2071).

– (1995). *Petits diviseurs en dimension 1*. Vol. 231. Astérisque. Société Mathématique de France (cit. on p. 2071).

Jean-Christophe Yoccoz (1984b). "Il n'y a pas de contre-exemple de Denjoy analytique". *C. R. Acad. Sci. Paris Sér. I Math.* 298.7, pp. 141–144. MR: 741080 (cit. on p. 2072).

R. J. Zimmer (1984). "Kazhdan groups acting on compact manifolds". *Invent. Math.* 75.3, pp. 425–436. MR: 735334 (cit. on p. 2055).

Received 2017-12-02.

Andrés Navas
Dpto de Matemática y Ciencia de la Computación
Universidad de Santiago de Chile
Alameda 3363, Estación Central
Santiago, Chile
andres.navas@usach.cl

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2081–2104)

ROBUST DYNAMICS, INVARIANT STRUCTURES AND TOPOLOGICAL CLASSIFICATION

Rafael Potrie

Abstract

This text is about geometric structures imposed by robust dynamical behaviour. We explain recent results towards the classification of partially hyperbolic systems in dimension 3 using the theory of foliations and its interaction with topology. We also present recent examples which introduce a challenge in the classification program and we propose some steps to continue this classification. Finally, we give some suggestions on what to do after classification is achieved.

1 Introduction

A major goal of dynamics is to be able to predict long term behaviour of a system via the knowledge of the rules that govern the way it is transformed. In this context, whatever can be said *a priori* of a system is relevant. It is important to search for conditions that can be detected by observing the system evolve in a finite amount of time that will result in consequences on the asymptotic behaviour of it. A beautiful example of this interaction is Shub's entropy conjecture (see e.g. Shub [2006]) which states that the knowledge of how a manifold wraps around itself (action in homology, detectable in just one iterate) is enough to find lower bounds on the complexity (entropy) of the system.

The unifying theme of this paper is the dynamical implications of invariant geometric structures and the interaction of the latter with topological and geometric structures on the phase space of the dynamical system. Hyperbolicity has been quite successful in the following sense[1]:

- it is possible (via cone-fields) to detect if a system is hyperbolic (with given constants) using only finitely many iterates,

The author was partially supported by CSIC group 618.
MSC2010: primary 37D30; secondary 37C20, 57R30.
Keywords: Differentiable dynamics, Partial hyperbolicity, foliations, topology of 3-manifolds.

[1] We refer the reader to Bowen [1975], Katok and Hasselblatt [1995], Shub [1987], and Franks [1982] for a more complete mathematical and historical account of hyperbolic theory.

- hyperbolicity is strongly tied with robust dynamical behaviour, namely, structural stability. (Hyperbolic maps are well known to be structurally (or Ω-)stable and the converse direction is also known to hold in the C^1-topology.) This says that one may expect hyperbolicity when the dynamical system, up to change of coordinates, is stable under small perturbations,
- it is possible to obtain very precise dynamical information of hyperbolic systems from the topological, symbolic and ergodic points of view,
- for globally hyperbolic systems (Anosov systems) and hyperbolic attractors there is a strong (yet incomplete) classification theory from the topological point of view.

As hyperbolic systems fail to describe all systems one is lead to encounter, it is natural to see what of this theory can extend to some weaker settings and weakenings of the notion of hyperbolicity have appeared in many different ways in the literature since the early 70's. The most ubiquitous generalisations of hyperbolicity are the notions of *non-uniform hyperbolicty* and *partial hyperbolicity*.

The first notion is a relaxation from the point of view of uniformity which forbids to detect this structure with information of finitely many iterates (though there are many important results that detect this property with positive probability in parametric families of dynamics). Non-uniform hyperbolicity is a property of certain invariant measures and provides very strong implications on the dynamics (see e.g. Sarig [2013] and references therein). Even in the case of non-uniform hyperbolicity of 'large measures' (such as volume), this structure does not impose topological restrictions on the manifolds that admit it Dolgopyat and Pesin [2002]. Let us remark that in the case of surfaces, there are some soft ways (Ruelle inequality) to detect non-uniform hyperbolicity of certain measures and this allows very strong description of dynamics of smooth diffeomorphisms of surfaces (see e.g. the recent Buzzi, Crovisier, and Sarig [n.d.]).

Partial hyperbolicity, the main object of this article, has some advantages over non-uniform hyperbolicity, though the study of its dynamics is far from being so developed[2]. A diffeomorphism $f : M \to M$ is said to admit a *dominated splitting* if it admits a Df-invariant continuous splitting $TM = E_1 \oplus \ldots \oplus E_\ell$ with $2 \le \ell \le \dim M$ into non-trivial subbundles and such that there exists $N > 0$ so that for any $x \in M$ and unit vectors $v_i \in E_i(x)$, $v_j \in E_j(x)$ with $i < j$ one has that:

$$\|Df^N v_i\| < \|Df^N v_j\|.$$

[2] There is an exception though, which is the case of surfaces where there is a quite complete understanding of the dynamics implied by a dominated splitting Pujals and M. Sambarino [2009] (see also Gourmelon and Potrie [n.d.]).

This condition can be detected via cone-fields (and therefore, given the strength, N, by knowing finitely many iterates of f, see e.g. Bonatti, Díaz, and Viana [2005, Appendix B] or Crovisier and Potrie [2015, Chapter 2]). It is easy to see that in some contexts, this structure already imposes constraints on the topology of manifolds that can admit such diffeomorphisms (e.g. in surfaces, the Euler characteristic must vanish).

A continuous Df-invariant bundle E is said to be *uniformly contracted* (resp. *uniformly expanded*) if there exists $N > 0$ so that for every unit vector $v \in E$ one has that:

$$\|Df^N v\| < 1 \text{ (resp. } \|Df^{-N} v\| < 1 \text{).}$$

If a diffeomorphism $f : M \to M$ admits a dominated splitting of the form $TM = E_1 \oplus E_2 \oplus E_3$ (where E_2 may be trivial) one says that:

- f is *Anosov* if E_1 is uniformly contracted, $E_2 = \{0\}$ and E_3 is uniformly expanded,
- f is *strongly partially hyperbolic* if both E_1 is uniformly contracted and E_3 is uniformly expanded.
- f is *partially hyperbolic* if either E_1 is uniformly contracted or E_3 is uniformly expanded,

We typically put all the uniformly contracting bundles together and denote the resulting bundle as E^s (and symmetrically we denote E^u to the sum of all uniformly expanded subbundles). The rest of the bundles are typically called *center bundles* and it is their existence that makes this structure, on the one hand more flexible and ubiquitous, and on the other, harder to understand.

As already explained, the study of Anosov systems from a topological point of view is still incomplete, still some remarkable progress has been achieved, notably through the work of Franks, Manning and Newhouse (see Hammerlindl and Potrie [2018, Section 3] for a fast account with references). Anosov flows have received lot of attention, but even in dimension 3 their classification is still far from complete (see e.g. the introductions of Barbot and Fenley [2015b] and Beguin, Bonatti, and Yu [2016]). It may seem hopeless to try to attack a classification of partially hyperbolic systems, even in dimension 3.

Pujals has proposed to try to classify strongly partially hyperbolic systems by comparing them with Anosov systems. This would yield relevant information in the quest to understand its dynamics even if the understanding of Anosov systems is incomplete. Later we will try to expand on this. The proposal was undertaken in the pioneering works of Bonatti and Wilkinson [2005] and Brin, Burago, and Ivanov [2004] and Burago and Ivanov [2008] and has spurred several results in the subject (see Carrasco, F. Hertz, J. Hertz, and Ures [2018] and Hammerlindl and Potrie [2018] for recent surveys).

The progress in this program has been intense and several unexpected features started to appear Bonatti, Parwani, and Potrie [2016] and Bonatti, Gogolev, and Potrie [2016] leading to the recent realisation of new features of strongly partially hyperbolic diffeomorphisms in dimension 3 in Bonatti, Gogolev, Hammerlindl, and Potrie [2017]. There is still lot of work ahead, but several positive results give hope that a more precise program can be attacked. We will survey these examples as well as the recent positive results in Barthelmé, Fenley, Frankel, and Potrie [2017] and we refer the reader to Hammerlindl and Potrie [2018] for a more complete account on the classification of strongly partially hyperbolic systems in 3-manifolds with solvable fundamental group.

Acknowledgments. I thank the generosity and support of great mathematicians who I admire: C. Bonatti, E. Pujals, A. Wilkinson and specially S. Crovisier and M. Sambarino. I was lucky to work with all my collaborators and students, I cannot explicitly acknowledge all, but I am extremely grateful to each of them, what appears here is full of their contributions. A. Passeggi and A. Sambarino have shared the way since the very beginning. Nati and Amalia make it worth it.

2 Robust dynamics

Partially hyperbolic dynamics arose as a natural generalisation of hyperbolicity Hirsch, Pugh, and Shub [1977] and as a way to deal with some systems arising naturally in other contexts such as frame flows in negatively curved manifolds Brin and Pesin [1974]. It also provided examples of structurally stable higher dimensional Lie group actions on manifolds. In analogy to the hyperbolic setting, these systems were shown to be robust and enjoy some stability properties, at least in the case where they admit an invariant foliation tangent to the center direction Berger [2010] and Hirsch, Pugh, and Shub [1977]. Also, as in the case of hyperbolic systems, the strong bundles always integrate into invariant foliations. We refer the reader to Bonatti, Díaz, and Viana [2005] and Crovisier and Potrie [2015] for recent expositions of these properties.

Structural stability was early conjectured to imply hyperbolicity (Palis and Smale [1970]) and this was famously proven true by R. Mañé (Mañé [1982, 1988]) in the C^1-topology. The ideas introduced by Mañé involved the use of *dominated splittings* and have since allowed to characterise other robust dynamical behaviour.

A system is said to be *transitive* if it has a dense orbit. In Mañé [1982] it is shown that a C^1-robustly[3] transitive surface diffeomorphism must be Anosov (in particular, the surface must be the torus!). These ideas were shown to extend to obtain (weaker) geometric

[3] i.e. its C^1 perturbations are transitive.

structures in a series of works Díaz, Pujals, and Ures [1999] and Bonatti, Díaz, and Pujals [2003].

Theorem 2.1 (Bonatti-Diaz-Pujals-Ures)**.** *Let $f : M \to M$ be a robustly transitive diffeomorphism. Then, f is* volume hyperbolic[4]. *In particular, for M a 3-dimensional manifold, f is partially hyperbolic.*

Parallel to the theory of robust transitivity, another theory deeply related to partial hyperbolicity was developed; it is the theory of *stable ergodicity*, we refer the reader to Wilkinson [2010] for a survey of this theory. The analog of the above theorem also exists: if a diffeomorphism is stably ergodic in dimension 3, then it is partially hyperbolic Bochi, Fayad, and Pujals [2006]. These results are sharp, see the examples in Bonatti and Díaz [1996] and Bonatti and Viana [2000].

Theorem 2.1 imposes some restrictions on manifolds admitting robustly transitive (or stably ergodic) diffeomorphisms; e.g. it is known that even-dimensional spheres cannot admit a pair of transverse continuous sub-bundles (see e.g. footnote 1 in Avila and Bochi [2012]) and other manifolds may also have such obstructions. However, these obstructions are far from being as sharp as in the case of surfaces. For instance, the following is still an open problem:

Question 1. *Is there a 3-manifold which does not admit robustly transitive (or stably ergodic) diffeomorphisms? Does the sphere S^3 admit robustly transitive (or stably ergodic) diffeomorphisms? Are there robustly transitive diffeomorphisms of $\mathbb{T}^3$ homotopic to the identity?*

This seems to be a difficult question. It was shown that the 3-sphere cannot admit strongly partially hyperbolic systems Burago and Ivanov [2008]. For endomorphisms in dimension 2 (which can be seen as a toy model for Question 1) a complete description of topological obstructions for robust transitivity was recently obtained Lizana and Ranter [2017]. In higher dimensions, even obstructions to the existence of Anosov diffeomorphisms are far from well understood (see Gogolev and Lafont [2016]).

The ultimate goal would be to provide a topological classification of partially hyperbolic systems which would allow to understand finer dynamical properties that they poses and in that way deduce that if a system has some robust property then its dynamics can be precisely understood.

[4]This is a weaker notion than partial hyperbolicity. It requires both extremal bundles to verify that the differential uniformly expands/contracts the volume in the bundle, see e.g. Crovisier and Potrie [2015] for this and more definitions. When the bundle is one-dimensional this implies that the bundle is uniformly expanded or contracted.

3 Examples

To the present, we know of the following mechanisms to construct examples (we refer the reader to Crovisier and Potrie [2015, Section 3] for a rather large list of examples)

- Algebraic and geometric constructions. Including linear automorphism of tori and nilmanifolds and geodesic and frame flows on negative curvature.
- Skew-products.
- Examples arising from h-transversalities, including Mañé type examples (Mañé [1978]) and small C^1-perturbations of partially hyperbolic systems.
- Surgery constructions.

None of the four mechanisms is completely understood. Algebraic examples are not even completely understood for the construction of Anosov diffeomorphisms and for geometric examples it is not completely clear which manifolds admit metrics of negative curvature.

Skew products where the base is more hyperbolic than the fiber can be said to be well understood, but when they work the other way around (the expansion and contraction is seen in the fibers), this is just starting to be studied and several exciting examples are starting to appear (see F. Hertz, J. Hertz, and Ures [2016], Farrell and Gogolev [2016], and Gogolev, Ontaneda, and F. Hertz [2015]).

The concept of h-transversality was just recently introduced Bonatti, Gogolev, Hammerlindl, and Potrie [2017] though it appeared implicitly in several examples.

Surgery constructions are only partly understood for Anosov flows (see Franks and Williams [1980], Handel and Thurston [1980], Goodman [1983], Fried [1983], Fenley [1994], and Béguin, Bonatti, and Yu [2017]). For partially hyperbolic diffeomorphisms, this kind of construction is in its infancy Gogolev [2018] and Bonatti and Wilkinson [2005].

Question 2. *Are there other ways to construct examples?*

In the rest of this section we extend a bit on the last two kind of examples.

3.1 h-Transversalities.

An h-transversality between two partially hyperbolic diffeomorphisms $f, g : M \to M$ is a diffeomorphism $h : M \to M$ which verifies that $Dh(E^u_f)$ is transverse to $E^s_g \oplus E^c_g$ and $Dh^{-1}(E^s_g)$ is transverse to $E^c_f \oplus E^u_f$.

The key property of this condition is that if f is h-transverse to itself, then $f^n \circ h$ will be partially hyperbolic for large n (see Bonatti, Gogolev, Hammerlindl, and Potrie [2017, Section 2]) so it provides a nice way to construct new examples once one gets enough control on the bundles of a partially hyperbolic diffeomorphism.

Using some detailed study of these bundles for certain Anosov flows (see Bonatti, Parwani, and Potrie [2016], Bonatti, Gogolev, and Potrie [2016], and Bonatti and Zhang [2017]) in Bonatti, Gogolev, Hammerlindl, and Potrie [2017] we were able to construct a large family of new partially hyperbolic examples. Several questions remain (see in particular Bonatti, Gogolev, Hammerlindl, and Potrie [ibid., Section 1.4]) but we emphasise on some that involve the h-transversality itself:

Question 3. *Describe the set of h-transversalities from an Anosov flow to itself. In particular, is it possible to construct h-transversalities from an Anosov flow to itself so that h is isotopic to identity but not through h-transversalities?*

If the last question admits a positive answer one could hope to construct partially hyperbolic diffeomorphisms isotopic to identity behaving very differently from an Anosov flows (see the discussion after Theorem 5.7 and Barthelmé, Fenley, Frankel, and Potrie [2017]).

Also, one can wonder if this notion may help creating new examples in higher dimensions, this is completely unexplored territory.

3.2 Surgeries. A conjecture attributed to Ghys (see Dehornoy [2013]) asserts that every transitive Anosov flow can be obtained (up to topological equivalence) from a given one by performing a finite number of simple operations that consist essentially on making finite lifts or quotients and surgeries (of Fried's type Fried [1983]).

We can propose a very vague question in the setting of partially hyperbolic diffeomorphisms:

Question 4. *If one add operations such as h-transversalities and some new type of surgeries to partially hyperbolic diffeomorphisms*[5]*, can one obtain a classification up to this equivalence?*

In particular, let us mention that we see all skew-products in dimension 3 as equivalent (the surgery is well explained in Bonatti and Wilkinson [2005, Proposition 4.2]). Also, the product of an Anosov in $\mathbb{T}^2$ and the identity in the circle can be easily 'surgered' to obtain the time one map of a suspension Anosov flow. On the other hand, we believe that partially hyperbolic diffeomorphisms of DA-type should not be in the same equivalence class as skew products (see Potrie [2015b] for evidence even if the question is not well posed).

[5] One should also change topological equivalence by some 'conjugacy modulo centers' which should be weaker than leaf conjugacy to allow non-dynamical coherence, see also Question 12.

4 Partial hyperbolicity in other contexts

Here we give a small glimpse of other contexts on which one encounters partially hyperbolic dynamics. The choice of topics is certainly biased by the author's interests.

4.1 Dynamics far from tangencies. We refer the reader to Crovisier and Potrie [2015] for a more complete account on the relations between partial hyperbolicity and dynamics far from homoclinic tangencies and Crovisier [2014] for a survey on its recent progress. We wish to emphasise the following point though: homoclinic tangencies are a semilocal phenomena, so the partial hyperbolicity obtained by being far from homoclinic tangencies only holds on the chain-recurrent set. In this setting, it makes sense to work and try to analyse the classes independently and so many global arguments (i.e. that depend on the topology of the manifold, or the isotopy class of the diffeomorphism) are lost.

Probably the main remaining open question in this setting is the following:

Conjecture 1 (Bonatti [2011])**.** *Generic diffeomorphisms far from homoclinic tangencies have finitely many chain-recurrence classes.*

See Pujals and M. Sambarino [2000], Crovisier, M. Sambarino, and D. Yang [2015], Croviser, Pujals, and M. Sambarino [n.d.], and Crovisier [2013] for some progress in this direction. A natural object of study that may combine well global and geometric arguments with the semi-local ones is the study of *attractors*, see Crovisier and Pujals [2015] and Crovisier, Potrie, and M. Sambarino [2017].

4.2 Skew-products. Skew products appear everywhere, as iterated function systems, as so called *fast-slow dynamics* or even as random perturbations of dynamics. The idea is to couple some dynamics with a random, or chaotic behaviour in the base which typically can be modelled by a hyperbolic system. This way, one naturally obtains a partially hyperbolic system (when the coupling is 'more random' that the dynamics on the fibres).

This point of view appears several places in the literature, for example in the notion of *fiber bunching* introduced in Bonatti, Gómez-Mont, and Viana [2003] by extending ideas of Ledrappier [1986] (see also Avila and Viana [2010]). Fiber bunching allows one to see the fibered dynamics as a partially hyperbolic system and construct invariant holonomies which are essentially lifts of strong stable and unstable manifolds to the fibered dynamics.

We have used this idea to give a somewhat different approach to the study of the Livšic problem for non-conmutative groups Kocsard and Potrie [2016] (see Kalinin [2011] and references therein for an introduction to the problem). This was later used also in recent results such as Hurtado [2016] and Avila, Kocsard, and Liu [2017].

4.3 Discrete subgroups of Lie groups. We refer the reader to Benoist [1997], Bridgeman, Canary, Labourie, and A. Sambarino [2015], and Guéritaud, Guichard, Kassel, and Wienhard [2017] and reference therein for a more detailed presentation of the subject. We just mention here that in Bochi, Potrie, and A. Sambarino [2017] we re-interpreted an interesting family of representations of certain groups into Lie groups, known as *Anosov representations* and introduced by Labourie (and later extended to general word-hyperbolic groups by Guichard and Wienhard) in terms of *dominated splitting* and partial hyperbolicity. We refer the reader to Bochi, Potrie, and A. Sambarino [ibid.] but we pose here the following question which we believe to be in the same spirit as Theorem 2.1:

Question 5. *Let Γ be a word hyperbolic group, G a semisimple Lie group and $\rho : \Gamma \to G$ a representations which is robustly[6] faithful and discrete. Is it Anosov for some parabolic of G? The same question makes sense if one demands ρ to be robustly quasi-isometric and G of real rank ≥ 2.*

The question is open even for robustly quasi-isometric representations of the free group in two generators into $\mathrm{SL}(3,\mathbb{R})$. This specific question can be posed in the language of linear cocycles as follows:

Question 6. *Let $A_0, B_0 \in \mathrm{SL}(3,\mathbb{R})$ be two matrices such that for every A, B close to A_0, B_0 one has that the linear cocycle over the subshift of $\{A, B, A^{-1}, B^{-1}\}^{\mathbb{Z}}$ which does not allow products AA^{-1}, $A^{-1}A$, BB^{-1}, $B^{-1}B$ verifies that it has positive Lyapunov exponents for every invariant measure. Does this linear cocycle admit a partially hyperbolic splitting?*

As in the case of Theorem 2.1 this question can be divided in two: show that periodic orbits are *uniformly hyperbolic* at the period, and show that this is enough to obtain a dominated splitting. The second part is Bochi, Potrie, and A. Sambarino [ibid., Question 4.10] and we were recently able to solve it Kassel and Potrie [n.d.].

5 Strong partial hyperbolicity in 3-manifolds

In this section we will be concerned with diffeomorphisms $f : M \to M$ where M is a 3-dimensional closed manifold and f is a (strongly) partially hyperbolic diffeomorphism admitting an invariant splitting of the form $TM = E^s \oplus E^c \oplus E^u$ into one-dimensional bundles. For some of the progress in higher dimensions we refer the reader to Hammerlindl and Potrie [2018, Section 14].

The topological study of these systems can be divided intro three main problems:

- to find topological obstructions for M to admit such diffeomorphisms,

[6]The topology in the space of representations is given by pointwise convergence.

- to study the integrability of the center bundle,
- to classify these systems up to what happens in the center direction.

We present in this section the state of the art in these problems. It is relevant to remark that in dimension 3 there is a quite advanced knowledge on the topology of closed manifolds (see e.g.Hatcher [n.d.]) and its interactions with geometry and foliations (we refer the reader to Calegari [2007] for a nice account). This allows to pursue a one-by-one methodology to deal with classes of manifolds with increasing complexity. In higher dimensions, different approaches need to be explored and one possibility is to start by making assumptions on the center foliations (in the spirit of Bonatti and Wilkinson [2005]) to work from there instead of studying specific classes of manifolds.

5.1 Topological obstructions. In this section we are interested with the following:

Question 7. *Which 3-manifolds support partially hyperbolic diffeomorphisms? If M admits a partially hyperbolic diffeomorphism, which isotopy classes of diffeomorphisms of M admit partially hyperbolic representatives?*

This question has a complete answer for 3-manifolds whose fundamental group has subexponential growth. In Burago and Ivanov [2008] it is shown that if the fundamental group of M is abelian, then the action on the homology of M has to be partially hyperbolic (in Potrie [2015a, Appendix A] it is shown that it has to be strongly partially hyperbolic). This in particular gives that manifolds such as S^3 or $S^2 \times S^1$ do not support partially hyperbolic diffeomorphisms. This was extended in Parwani [2010] to manifolds with subexponential growth of fundamental group obtaining a similar result. By now, we have a complete classification of partially hyperbolic diffeomorphisms in 3-manifolds with (virtually) solvable fundamental group (see Hammerlindl and Potrie [2014, 2015, 2018]) and we know exactly which isotopy classes admit partially hyperbolic representatives.

On the other hand, in all generality, we do not even know which manifolds admit Anosov flows. For example, the following question is open [7]:

Question 8. *Does every hyperbolic 3-manifold M admit a finite lift supporting an Anosov flow?*

However, one can pose the following question in the spirit of Pujals' conjecture:

[7] The classification of Anosov flows is completely open in hyperbolic 3-manifolds; when the manifold is toroidal there has been important recent progress towards classification Barbot and Fenley [2013, 2015a,b], Béguin, Bonatti, and Yu [2017], and Beguin, Bonatti, and Yu [2016].

Question 9. *If a 3-manifold M with exponential growth of fundamental group admits a (transitive) strongly partially hyperbolic diffeomorphism, does it also (after finite lift) admit a (topological)*[8] *Anosov flow?*

This question is still very far from being understood, but substantial progress has been made in the case where one assumes that f is homotopic to identity. I expect the answer to be yes at least in that case.

When M is Seifert then the answer to Question 9 is affirmative (see Hammerlindl, Potrie, and Shannon [2017]). On the other hand, even in Seifert manifolds, recent examples Bonatti, Gogolev, and Potrie [2016] and Bonatti, Gogolev, Hammerlindl, and Potrie [2017] show that several isotopy classes can be produced, but also one can find some obstructions (see Bonatti, Gogolev, Hammerlindl, and Potrie [2017, Section 3.1]).

When M is toroidal and admits Anosov flows transverse to some tori, also several classes of examples can be constructed, and the question of which isotopy classes admit partially hyperbolic representatives is quite open (see Bonatti, Parwani, and Potrie [2016], Bonatti, Gogolev, and Potrie [2016], Bonatti and Zhang [2017], and Bonatti, Gogolev, Hammerlindl, and Potrie [2017]).

5.2 Integrability. The stable and unstable bundles of f are known to be (uniquely) integrable into f-invariant foliations $\mathcal{W}^s$ and $\mathcal{W}^u$. This is for dynamical reasons (see Hirsch, Pugh, and Shub [1977] and Hammerlindl and Potrie [2018]).

However, the lack of regularity of E^c (it is just Hölder continuous) and the fact that its dynamics is neither contracting or repelling makes its integrability a particular feature. Now, we know several examples where E^c does not integrate into an f-invariant foliation. See F. Hertz, J. Hertz, and Ures [2016] and Bonatti, Gogolev, Hammerlindl, and Potrie [2017].

Rather than asking for integrability of E^c into an invariant foliation, one typically ask whether f is *dynamically coherent* meaning that both $E^{cs} = E^s \oplus E^c$ and $E^{cu} = E^c \oplus E^u$ integrate into f-invariant foliations $\mathcal{W}^{cs}$ and $\mathcal{W}^{cu}$. This implies the existence of an f-invariant foliation $\mathcal{W}^c$. This definition involves several subtleties (it is unknown whether the existence of a f-invariant center foliation implies dynamical coherence, see Burns and Wilkinson [2008] for discussions on this definition).

A fundamental result was proved by Burago and Ivanov [2008] providing the existence of *branching foliations* (a technical object, see Figure 1) for every strong partially hyperbolic diffeomorphism in 3-manifolds. To many effects, these objects (see also Hammerlindl and Potrie [2018, Chapter 4]) replace dynamical coherence quite well and allow

[8]We remark that topological Anosov flows are conjectured to be orbit equivalent to true Anosov flows (Bonatti and Wilkinson [2005]).

one to search for classification results. However, it makes sense to ask when do partially hyperbolic diffeomorphisms are dynamically coherent.

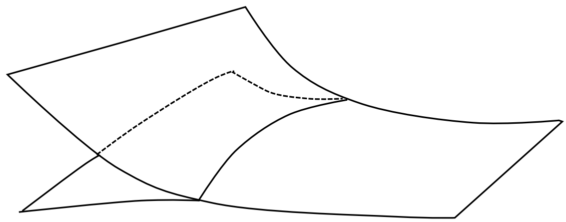


Figure 1: Leaves of a branching foliation may merge.

We first review some results in this setting:

Theorem 5.1 (Brin, Burago, and Ivanov [2009], F. Hertz, J. Hertz, and Ures [2016], and Potrie [2015a])**.** *Let $f : \mathbb{T}^3 \to \mathbb{T}^3$ be a strong partially hyperbolic diffeomorphism. Then, unless there is a torus tangent to E^{cs} or E^{cu} the diffeomorphism f is dynamically coherent.*

This result is proved by obtaining a precise analysis of the structure of the branching foliations provided by Burago and Ivanov [2008] and use this structure to show that branching is not possible (and therefore the branching foliations are indeed foliations). We would like to emphasise that part of the results obtained in Potrie [2015a] hold for general (not necessarily strong) partially hyperbolic diffeomorphisms through the notion of *almost dynamical coherence*, an open and closed property, which has been also exploited in Fisher, Potrie, and M. Sambarino [2014] and Roldán [2016].

The same ideas have been pushed into several new contexts by careful combination of topological analysis of the manifolds in hand, comparison to some 'model' example and giving some structure to the branching foliations. For instance, with A. Hammerlindl, we were able to show the following results:

Theorem 5.2 (Hammerlindl and Potrie [2014])**.** *If N is a non-toral nilmanifold and $f : N \to N$ is partially hyperbolic, then f is dynamically coherent.*

Theorem 5.3 (Hammerlindl and Potrie [2015])**.** *If M is a 3-manifold with (virtually) solvable fundamental group and $f : M \to M$ is partially hyperbolic, then, unless there is a torus tangent E^{cs} or E^{cu} the diffeomorphism f is dynamically coherent.*

These results respond affirmatively to a conjecture by Hertz-Hertz-Ures in such manifolds (see F. Hertz, J. Hertz, and Ures [2016] and Carrasco, F. Hertz, J. Hertz, and Ures [2018]). This conjecture has been disproved recently in Bonatti, Gogolev, Hammerlindl, and Potrie [2017] but there are still some cases where it can be studied:

Theorem 5.4 (Barthelmé, Fenley, Frankel, and Potrie [2017]). *Let $f : M \to M$ be a strong partially hyperbolic diffeomorphism of a Seifert manifold M such that f is homotopic to identity. Then, f is dynamically coherent.*

Seifert manifolds which admit transitive partially hyperbolic diffeomorphisms are, up to finite cover, nilmanifolds or unit tangent bundles of higher genus surfaces (see Hammerlindl, Potrie, and Shannon [2017]). In the case of unit tangent bundles of higher genus surfaces, we have constructed in Bonatti, Gogolev, Hammerlindl, and Potrie [2017] examples which are not dynamically coherent. The configuration responsible for this incoherence is of global nature (see Figure 2), so it seems natural to ask:

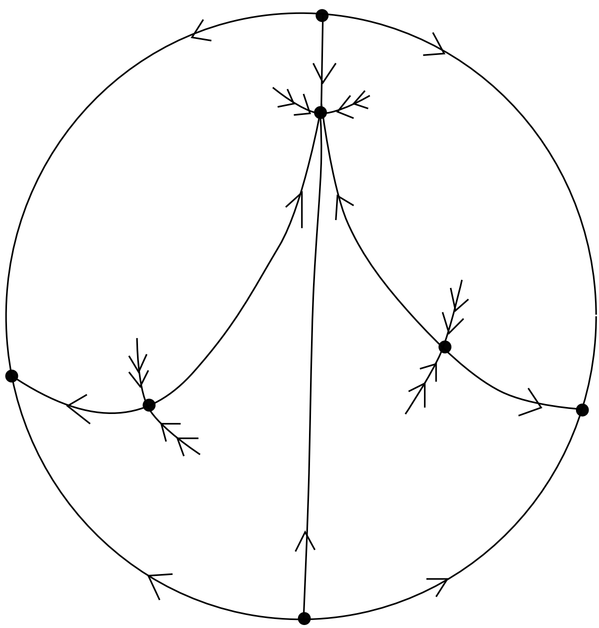


Figure 2: The global index in certain periodic cs-leaves is negative. Unless c-curves merge, there cannot be cancelation so this forces non-dynamical coherence.

Question 10. *For a higher genus surface S, if $f : T^1S \to T^1S$ is a strong partially hyperbolic diffeomorphism which induces*[9] *a pseudo-Anosov mapping class in S then f does not admit an f-invariant foliation tangent to E^c?.*

We expect that the techniques used in Barthelmé, Fenley, Frankel, and Potrie [2017] to classify partially hyperbolic diffeomorphisms of hyperbolic 3-manifolds with f-invariant center foliations will allow to provide a positive answer to the previous question.

When the induced action of f on S is a Dehn-twist (c.f. the examples from Bonatti, Gogolev, and Potrie [2016]) then it seems possible that this will imply that the diffeomorphism can be chosen to be dynamically coherent.

Question 11. *Is there a connected component of partially hyperbolic diffeomorphisms containing both dynamically coherent and non-dynamically coherent diffeomorphisms?*

This question is related with the celebrated *plaque expansivity conjecture* (see Hirsch, Pugh, and Shub [1977] and Berger [2010]) since plaque expansivity ensures stability of the invariant foliation, and in case a connected component as in the question exists, it would be natural to check for plaque expansivity in the boundary of dynamically coherent ones.

5.3 Classification. When $f : M \to M$ is a strong partially hyperbolic diffeomorphism which is dynamically coherent, the right notion of classification is given by *leaf conjugacy*: $f, g : M \to M$ dynamically coherent strong partially hyperbolic diffeomorphisms are said to be *leaf conjugate* if there exists a homeomorphism $h : M \to M$ so that $h(\mathcal{W}^c_f(f(x))) = \mathcal{W}^c_g(g(h(x)))$. This notion goes back to Hirsch, Pugh, and Shub [1977] where a 'local stability result' was shown and was retaken in the thesis of Hammerlindl [2013] to show that some strong partially hyperbolic diffeomorphisms of $\mathbb{T}^3$ are leaf conjugate to linear automorphisms of tori. (Notice that this notion does not really require dynamical coherence but the existence of an f-invariant center foliation.)

In this context, we have shown:

Theorem 5.5 (Hammerlindl and Potrie [2014, 2015]). *If M is a manifold with (virtually) solvable fundamental group and $f : M \to M$ is a strong partially hyperbolic diffeomorphism with an f-invariant center foliation, then (up to finite lift and iterate) it is leaf conjugate to an algebraic example.*

In Hammerlindl and Potrie [2017] we further classify those f in such manifolds which do not have an f-invariant center foliation.

[9] Every diffeomorphism of a Seifert manifold is homotopic to a diffeomorphism preserving the fibers and therefore it has a well defined action on S up to homotopy.

In Seifert manifolds, where (topological) Anosov flows are classified (Ghys [1984], Barbot [1996], and Brunella [1993]) one can classify strong partially hyperbolic diffeomorphisms homotopic to identity.

Theorem 5.6 (Barthelmé, Fenley, Frankel, and Potrie [2017]). *If $f : M \to M$ is a strong partially hyperbolic diffeomorphism homotopic to identity in a Seifert manifold M, then, it is dynamically coherent and leaf conjugate to (up to finite lifts and quotients) the geodesic flow on a surface of negative curvature.*

We remark that R.Ures has announced a similar result in $T^1 S$ which assumes that f is isotopic to the geodesic flow through a path of partially hyperbolic diffeomorphisms (a similar condition to the one studied in Fisher, Potrie, and M. Sambarino [2014]).

The results in Barthelmé, Fenley, Frankel, and Potrie [2017] deal with general partially hyperbolic diffeomorphisms of 3-manifolds which are homotopic to identity. However, there are some points where the precise knowledge of the topology of the manifold under study allows us to give much stronger results.

For example, when M is a hyperbolic 3-manifold we use the existence of transverse pseudo-Anosov flows to uniform foliations (Thurston [1997], Calegari [2000], and Fenley [2002]) to obtain stronger properties and we deduce the following result (which gives a positive answer to a classification conjecture from Carrasco, F. Hertz, J. Hertz, and Ures [2018] for hyperbolic manifolds):

Theorem 5.7 (Barthelmé, Fenley, Frankel, and Potrie [2017]). *Let M be a hyperbolic 3-manifold and $f : M \to M$ a dynamically coherent partially hyperbolic diffeomorphism, then f is leaf conjugate*[10] *to the time one map of a (topological) Anosov flow.*

We make emphasis that this is the first result on classification of partially hyperbolic diffeomorphisms on manifolds where we do not have a model a priori to compare our partially hyperbolic diffeomorphism to.

This result should hold changing dynamical coherence by the existence of an f-invariant foliation tangent to E^c. However, I expect that there will be non-dynamically coherent examples in hyperbolic 3-manifolds since some arguments in this theorem resemble those we use in Bonatti, Gogolev, Hammerlindl, and Potrie [2017] to show that certain examples are not dynamically coherent. These examples would be what we call 'double translations' as they act (in the universal cover[11]) as translation in both the center-stable and center-unstable (branching) foliations.

Indeed, one can think about the following analogy: Let $f : M \to M$ be a strong partially hyperbolic diffeomorphism homotopic to the identity on a hyperbolic 3-manifold

[10]Technically, the time one map of a topological Anosov flow is not a partially hyperbolic diffeomorphism. But the notion still makes sense.

[11]Here, we assume that we take a lift at bounded distance from the identity.

M which is the suspension of a pseudo-Anosov diffeomorphism $\varphi : S \to S$. Then, the fundamental group of M can be written as a semidirect product $\pi_1(S) \rtimes_\varphi \mathbb{Z}$. Moreover, the lift of f at bounded distance from the identity in $\tilde{M}$ commutes with all deck transformations and translates the foliations. Then, we get a group G_f which is the direct product of $\langle f \rangle$ and $\pi_1(M)$.

If one looks at the examples of Bonatti, Gogolev, Hammerlindl, and Potrie [2017] one has that the fundamental group of T^1S enters in a exact sequence $0 \to \mathbb{Z} \to \pi_1(T^1S) \to \pi_1(S) \to 0$ where the inclusion of $\mathbb{Z}$ is the center of the group. A lift $\tilde{g}$ of the diffeomorphism $g : T^1S \to T^1S$ constructed in Bonatti, Gogolev, Hammerlindl, and Potrie [ibid.] (which induces a pseudo-Anosov map in the base) will define a group G_g of diffeomorphisms of $\widetilde{T^1S}$ generated by $\pi_1(T^1S)$ and $\tilde{g}$. Here, the role of $\tilde{f}$ in G_f is played by the center of $\pi_1(T^1S)$ (which translates center stable and center unstable branching leafs) while $\tilde{g}$ plays the role of the 'semidirect product' in the fundamental group of M above.

If such an f existed, the dynamics of these two groups in the 'circle at infinity' would look quite alike. This suggests on the one hand that examples like this may exist in hyperbolic manifolds, and on the other hand, that it should be possible to answer Question 10 using similar ideas to the ones appearing in Theorem 5.7.

5.4 More questions. Several questions remain to be explored in the classification of partially hyperbolic diffeomorphisms in dimension 3. We pose here those we feel are more relevant or that we think might help address the problem of classification. We restrict to the case of 3-manifolds M whose fundamental group is not virtually solvable in view of Theorem 5.5. For simplicity we will assume throughout that everything is orientable (the manifold, the bundles, and that f preserves all the orientations).

Even if a partially hyperbolic diffeomorphism is not dynamically coherent, the one-dimensionality of the center direction allows to integrate the center bundle. However, there may not exist a foliation tangent to it (the strongest integrability that can be ensured beyond the existence of curves tangent to the center is the existence of branching foliations Burago and Ivanov [2008]).

In the non-dynamically coherent examples of Bonatti, Gogolev, Hammerlindl, and Potrie [2017] one sees that the space of center leafs is (naturally) homeomorphic to the space of orbits of a geodesic flow in negative curvature (see Bonatti, Gogolev, Hammerlindl, and Potrie [ibid., Proposition 5.11]) and the dynamics of center leafs is governed by the action of the mapping class of f in the 'boundary at infinity'. This could be a general phenomena:

Question 12. *Let $f : M \to M$ be a (transitive) partially hyperbolic diffeomorphism. Is there (maybe up to finite cover) an Anosov flow ϕ^t on M and continuous degree 1 map $h : M \to M$ sending orbits of ϕ to center leafs of f?*

This question is proposing a way to classify dynamics as one could expect that the isotopy class of f will force some dynamics on the center leafs. A related question (maybe more basic, but probably difficult to approach directly) is:

Question 13. *If $f : M \to M$ is a (transitive) partially hyperbolic diffeomorphism, are all center stable leafs cylinders or planes?*

Proving this in the case isotopic to identity is an important step towards the proof of Theorems 5.6 and 5.7, but as far as I am aware this problem has been never attacked directly in general (see Zhang [2017] for a positive answer assuming that the dynamics is *neutral* in the center direction).

The last two questions are definitely related, but they are independent as there are some subtleties in the notions of dynamical coherence, leaf conjugacy, etc. For example I do not know the answer to the following:

Question 14. *If the center foliation of f is homeomorphic to the orbit foliation of an Anosov flow, are all center stable leaves cylinders or planes?*

I believe that a reasonable way to attack classification would be to try to understand center leaves at infinity (to be able to avoid taking care of how they merge) and for this the tool of *universal circles* has shown to be quite useful in other contexts (see e.g. Thurston [1997], Calegari [2007], Fenley [2002], and Frankel [2015]).

6 Dynamical implications

We will ignore in this section the very important subject of *conservative* partially hyperbolic diffeomorphisms. They have been extensively treated in other recent surveys such as Carrasco, F. Hertz, J. Hertz, and Ures [2018] and Wilkinson [2010] with different points of view but great detail. We shall focus mostly on the subject of *robust transitivity* and finitness or uniqueness of attractors for such systems.

In this direction, one can pose the following question which already appears in Potrie [2014] (see also Bonatti, Gogolev, Hammerlindl, and Potrie [2017]):

Question 15. *Is there an isotopy class of diffeomorphisms of a 3-manifold M such that every strongly partially hyperbolic diffeomorphism in this isotopy class is transitive? Chain-recurrent?*

In several isotopy classes, such as Anosov times identity on $\mathbb{T}^3$ the answer is known to be negative (see e.g. Bonatti and Guelman [2010] and Shi [2014]) for more surprising examples). But one can still wonder about uniqueness of attractors, or minimal u-saturated sets (c.f. Crovisier, Potrie, and M. Sambarino [2017]).

Several isotopy classes of (strongly) partially hyperbolic diffeomorphisms seem to be now ready for studying subtler dynamical properties. There has been quite some progress in this directions, just to mention a few, we refer the reader to Bonatti and Viana [2000], Buzzi, Fisher, M. Sambarino, and Vásquez [2012], Ures [2012], Climenhaga, Fisher, and Thompson [2017], and Viana and J. Yang [2017] and references therein for advances in the DA-case, Viana and J. Yang [2013] and Tahzibi and J. Yang [2016] for the skew-product case and Saghin and J. Yang [2016] for the case of systems leaf conjugate to time one maps of Anosov flows. In all cases it makes sense to try to understand how the entropy behaves[12], how many measures of maximal entropy (or equilibrium states) one may have, physical measures, its statistical properties, its Lyapunov exponents, etc.

I close the paper with a question which I think points towards something we do not really understand yet, and which is more important than the question itself:

Question 16. *Let $f : M \to M$ be in the boundary of robustly transitive diffeomorphisms. Does there exist a center Lyapunov exponent for the maximal entropy measure of f which vanishes?*

It is natural to attack this question in the context of strongly partially hyperbolic diffeomorphisms with one-dimensional center, and it might be that the answer depends on the class. Still, I hope that in the near future we will understand better the transition between the interior of transitive diffeomorphisms and those admitting proper attractors (some progress is in Abdenur, Crovisier, and Potrie [n.d.], see also Crovisier and Potrie [2015, Section 5]).

References

F. Abdenur, S. Crovisier, and R. Potrie (n.d.). "Characterization of robustly transitive partially hyperbolic diffeomorphisms". In preparation (cit. on p. 2098).

A. Avila and J. Bochi (2012). "Nonuniform hyperbolicity, global dominated splittings and generic properties of volume-preserving diffeomorphisms". *Trans. Amer. Math. Soc.* 364.6, pp. 2883–2907. MR: 2888232 (cit. on p. 2085).

A. Avila, A. Kocsard, and X-C. Liu (Nov. 2017). "Livsic theorem for diffeomorphism cocycles". arXiv: 1711.02135 (cit. on p. 2088).

A. Avila and M. Viana (2010). "Extremal Lyapunov exponents: an invariance principle and applications". *Invent. Math.* 181.1, pp. 115–189. MR: 2651382 (cit. on p. 2088).

T. Barbot (1996). "Flots d'Anosov sur les variétés graphées au sens de Waldhausen". *Ann. Inst. Fourier (Grenoble)* 46.5, pp. 1451–1517. MR: 1427133 (cit. on p. 2095).

[12] There is a program in this direction due to J. Buzzi, see Buzzi [2009], but it has evolved since.

T. Barbot and S. Fenley (2013). "Pseudo-Anosov flows in toroidal manifolds". *Geom. Topol.* 17.4, pp. 1877–1954. MR: 3109861 (cit. on p. 2090).

– (2015a). "Classification and rigidity of totally periodic pseudo-Anosov flows in graph manifolds". *Ergodic Theory Dynam. Systems* 35.6, pp. 1681–1722. MR: 3377280 (cit. on p. 2090).

– (Dec. 2015b). "Free Seifert pieces of pseudo-Anosov flows". arXiv: 1512.06341 (cit. on pp. 2083, 2090).

T. Barthelmé, S. Fenley, S. Frankel, and R. Potrie (2017). "Partially hyperbolic diffeomorphisms homotopic to the identity on 3-manifolds". announcement, In preparation. arXiv: 1801.00214 (cit. on pp. 2084, 2087, 2093–2095).

F. Beguin, C. Bonatti, and B. Yu (2016). "A spectral-like decomposition for transitive Anosov flows in dimension three". *Math. Z.* 282.3-4, pp. 889–912. MR: 3473648 (cit. on pp. 2083, 2090).

F. Béguin, C. Bonatti, and B. Yu (2017). "Building Anosov flows on 3-manifolds". *Geom. Topol.* 21.3, pp. 1837–1930. MR: 3650083 (cit. on pp. 2086, 2090).

Y. Benoist (1997). *Sous-groupes discrets des groupes de Lie*. notes from the European Summer School in Group Theory, Luminy (cit. on p. 2089).

P. Berger (2010). "Persistence of laminations". *Bull. Braz. Math. Soc. (N.S.)* 41.2, pp. 259–319. MR: 2738914 (cit. on pp. 2084, 2094).

J. Bochi, B. Fayad, and E. Pujals (2006). "A remark on conservative diffeomorphisms". *C. R. Math. Acad. Sci. Paris* 342.10, pp. 763–766. MR: 2227756 (cit. on p. 2085).

J. Bochi, R. Potrie, and A. Sambarino (2017). "Anosov representations and dominated splittings". To appear in *Journal of the European Math. Soc.* (cit. on p. 2089).

C. Bonatti (2011). "Survey: Towards a global view of dynamical systems, for the C^1-topology". *Ergodic Theory Dynam. Systems* 31.4, pp. 959–993. MR: 2818683 (cit. on p. 2088).

C. Bonatti and L.J. Díaz (1996). "Persistent nonhyperbolic transitive diffeomorphisms". *Ann. of Math. (2)* 143.2, pp. 357–396. MR: 1381990 (cit. on p. 2085).

C. Bonatti, L.J. Díaz, and E. Pujals (2003). "A C^1-generic dichotomy for diffeomorphisms: weak forms of hyperbolicity or infinitely many sinks or sources". *Ann. of Math. (2)* 158.2, pp. 355–418. MR: 2018925 (cit. on p. 2085).

C. Bonatti, L.J. Díaz, and M. Viana (2005). *Dynamics beyond uniform hyperbolicity*. Vol. 102. Encyclopaedia of Mathematical Sciences. A global geometric and probabilistic perspective, Mathematical Physics, III. Springer-Verlag, Berlin, pp. xviii+384. MR: 2105774 (cit. on pp. 2083, 2084).

C. Bonatti, A. Gogolev, A. Hammerlindl, and R. Potrie (June 2017). "Anomalous partially hyperbolic diffeomorphisms III: abundance and incoherence". arXiv: 1706.04962 (cit. on pp. 2084, 2086, 2087, 2091, 2093, 2095–2097).

C. Bonatti, A. Gogolev, and R. Potrie (2016). "Anomalous partially hyperbolic diffeomorphisms II: stably ergodic examples". *Invent. Math.* 206.3, pp. 801–836. MR: 3573973 (cit. on pp. 2084, 2087, 2091, 2094).

C. Bonatti, X. Gómez-Mont, and M. Viana (2003). "Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 20.4, pp. 579–624. MR: 1981401 (cit. on p. 2088).

C. Bonatti and N. Guelman (2010). "Axiom A diffeomorphisms derived from Anosov flows". *J. Mod. Dyn.* 4.1, pp. 1–63. MR: 2643887 (cit. on p. 2097).

C. Bonatti, K. Parwani, and R. Potrie (2016). "Anomalous partially hyperbolic diffeomorphisms I: Dynamically coherent examples". *Ann. Sci. Éc. Norm. Supér. (4)* 49.6, pp. 1387–1402. MR: 3592360 (cit. on pp. 2084, 2087, 2091).

C. Bonatti and M. Viana (2000). "SRB measures for partially hyperbolic systems whose central direction is mostly contracting". *Israel J. Math.* 115, pp. 157–193. MR: 1749677 (cit. on pp. 2085, 2098).

C. Bonatti and A. Wilkinson (2005). "Transitive partially hyperbolic diffeomorphisms on 3-manifolds". *Topology* 44.3, pp. 475–508. MR: 2122214 (cit. on pp. 2083, 2086, 2087, 2090, 2091).

C. Bonatti and J. Zhang (2017). "Transverse foliations on the torus $\mathbb{T}^2$ and partially hyperbolic diffeomorphisms on 3-manifolds". *Comment. Math. Helv.* 92.3, pp. 513–550. arXiv: 1602.04355. MR: 3682779 (cit. on pp. 2087, 2091).

R. Bowen (1975). *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin-New York, pp. i+108. MR: 0442989 (cit. on p. 2081).

M. Bridgeman, R. Canary, F. Labourie, and A. Sambarino (2015). "The pressure metric for Anosov representations". *Geom. Funct. Anal.* 25.4, pp. 1089–1179. MR: 3385630 (cit. on p. 2089).

M. Brin, D. Burago, and S. Ivanov (2004). "On partially hyperbolic diffeomorphisms of 3-manifolds with commutative fundamental group". In: *Modern dynamical systems and applications*. Cambridge Univ. Press, Cambridge, pp. 307–312. MR: 2090777 (cit. on p. 2083).

– (2009). "Dynamical coherence of partially hyperbolic diffeomorphisms of the 3-torus". *J. Mod. Dyn.* 3.1, pp. 1–11. MR: 2481329 (cit. on p. 2092).

M. Brin and Y. Pesin (1974). "Partially hyperbolic dynamical systems". *Izv. Akad. Nauk SSSR Ser. Mat.* 38, pp. 170–212. MR: 0343316 (cit. on p. 2084).

M. Brunella (1993). "Expansive flows on Seifert manifolds and on torus bundles". *Bol. Soc. Brasil. Mat. (N.S.)* 24.1, pp. 89–104. MR: 1224301 (cit. on p. 2095).

D. Burago and S. Ivanov (2008). "Partially hyperbolic diffeomorphisms of 3-manifolds with abelian fundamental groups". *J. Mod. Dyn.* 2.4, pp. 541–580. MR: 2449138 (cit. on pp. 2083, 2085, 2090–2092, 2096).

K. Burns and A. Wilkinson (2008). "Dynamical coherence and center bunching". *Discrete Contin. Dyn. Syst.* 22.1-2, pp. 89–100. MR: 2410949 (cit. on p. 2091).

J. Buzzi (2009). "Dimensional entropies and semi-uniform hyperbolicity". In: *New Trends in Mathematical Physics*. Springer, pp. 95–116 (cit. on p. 2098).

J. Buzzi, S. Crovisier, and O. Sarig (n.d.). "Finite multiplicity for the measures of maximal entropy of surface diffeomorphisms". In Preparation (cit. on p. 2082).

J. Buzzi, T. Fisher, M. Sambarino, and C. Vásquez (2012). "Maximal entropy measures for certain partially hyperbolic, derived from Anosov systems". *Ergodic Theory Dynam. Systems* 32.1, pp. 63–79. MR: 2873158 (cit. on p. 2098).

D. Calegari (2000). "The geometry of $\mathbf{R}$-covered foliations". *Geom. Topol.* 4, pp. 457–515. MR: 1800151 (cit. on p. 2095).

– (2007). *Foliations and the geometry of 3-manifolds*. Oxford Mathematical Monographs. Oxford University Press, Oxford, pp. xiv+363. MR: 2327361 (cit. on pp. 2090, 2097).

P. Carrasco, F. Hertz, J. Hertz, and R. Ures (Jan. 2018). "Partially hyperbolic dynamics in dimension 3". to appear ETDS. arXiv: 1501.00932 (cit. on pp. 2083, 2093, 2095, 2097).

V. Climenhaga, T. Fisher, and D. Thompson (Mar. 2017). "Equilibrium states for Mañé diffeomorphisms". arXiv: 1703.05722 (cit. on p. 2098).

S. Croviser, E. Pujals, and M. Sambarino (n.d.). "Hyperbolicity of extremal bundles". In preparation (cit. on p. 2088).

S. Crovisier (2013). "Perturbation de la dynamique de difféomorphismes en topologie C^1". *Astérisque* 354, pp. x+164. MR: 3136194 (cit. on p. 2088).

– (2014). "Dynamics of C^1-diffeomorphisms: global description and prospects for classification". In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. III*. Kyung Moon Sa, Seoul, pp. 571–595. MR: 3729041 (cit. on p. 2088).

S. Crovisier and R. Potrie (2015). *Introduction to partially hyperbolic dynamics*. Lecture notes for a minicourse at ICTP, available in the webpages of the authors (cit. on pp. 2083–2086, 2088, 2098).

S. Crovisier, R. Potrie, and M. Sambarino (June 2017). "Finiteness of partially hyperbolic attractors with one-dimensional center". arXiv: 1706.08684 (cit. on pp. 2088, 2097).

S. Crovisier and E. Pujals (2015). "Essential hyperbolicity and homoclinic bifurcations: a dichotomy phenomenon/mechanism for diffeomorphisms". *Invent. Math.* 201.2, pp. 385–517. MR: 3370619 (cit. on p. 2088).

S. Crovisier, M. Sambarino, and D. Yang (2015). "Partial hyperbolicity and homoclinic tangencies". *J. Eur. Math. Soc. (JEMS)* 17.1, pp. 1–49. MR: 3312402 (cit. on p. 2088).

P. Dehornoy (2013). "Almost commensurability of 3-dimensional Anosov flows". *C. R. Math. Acad. Sci. Paris* 351.3-4, pp. 127–129. MR: 3038001 (cit. on p. 2087).

L.J. Díaz, E. Pujals, and R. Ures (1999). "Partial hyperbolicity and robust transitivity". *Acta Math.* 183.1, pp. 1–43. MR: 1719547 (cit. on p. 2085).

D. Dolgopyat and Y. Pesin (2002). "Every compact manifold carries a completely hyperbolic diffeomorphism". *Ergodic Theory Dynam. Systems* 22.2, pp. 409–435. MR: 1898798 (cit. on p. 2082).

F.T. Farrell and A. Gogolev (2016). "On bundles that admit fiberwise hyperbolic dynamics". *Math. Ann.* 364.1-2, pp. 401–438. MR: 3451392 (cit. on p. 2086).

S. Fenley (1994). "Anosov flows in 3-manifolds". *Ann. of Math. (2)* 139.1, pp. 79–115. MR: 1259365 (cit. on p. 2086).

– (2002). "Foliations, topology and geometry of 3-manifolds: **R**-covered foliations and transverse pseudo-Anosov flows". *Comment. Math. Helv.* 77.3, pp. 415–490. MR: 1933786 (cit. on pp. 2095, 2097).

T. Fisher, R. Potrie, and M. Sambarino (2014). "Dynamical coherence of partially hyperbolic diffeomorphisms of tori isotopic to Anosov". *Math. Z.* 278.1-2, pp. 149–168. MR: 3267574 (cit. on pp. 2092, 2095).

S. Frankel (July 2015). "Coarse hyperbolicity and closed orbits for quasigeodesic flows". arXiv: 1507.04320 (cit. on p. 2097).

J. Franks (1982). *Homology and dynamical systems*. Vol. 49. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, D.C.; by the American Mathematical Society, Providence, R. I., viii+120 pp. (loose errata). MR: 669378 (cit. on p. 2081).

J. Franks and B. Williams (1980). "Anomalous Anosov flows". In: *Global theory of dynamical systems (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1979)*. Vol. 819. Lecture Notes in Math. Springer, Berlin, pp. 158–174. MR: 591182 (cit. on p. 2086).

D. Fried (1983). "Transitive Anosov flows and pseudo-Anosov maps". *Topology* 22.3, pp. 299–303. MR: 710103 (cit. on pp. 2086, 2087).

E. Ghys (1984). "Flots d'Anosov sur les 3-variétés fibrées en cercles". *Ergodic Theory Dynam. Systems* 4.1, pp. 67–80. MR: 758894 (cit. on p. 2095).

A. Gogolev (2018). "Surgery for partially hyperbolic dynamical systems I. Blow-ups of invariant submanifolds". *Geom. Topol.* 22.4, pp. 2219–2252 (cit. on p. 2086).

A. Gogolev and J-F. Lafont (2016). "Aspherical products which do not support Anosov diffeomorphisms". *Ann. Henri Poincaré* 17.11, pp. 3005–3026. MR: 3556514 (cit. on p. 2085).

A. Gogolev, P. Ontaneda, and F. Hertz (2015). "New partially hyperbolic dynamical systems I". *Acta Math.* 215.2, pp. 363–393. MR: 3455236 (cit. on p. 2086).

S. Goodman (1983). "Dehn surgery on Anosov flows". In: *Geometric dynamics (Rio de Janeiro, 1981)*. Vol. 1007. Lecture Notes in Math. Springer, Berlin, pp. 300–307. MR: 1691596 (cit. on p. 2086).

N. Gourmelon and R. Potrie (n.d.). "Projectively Anosov diffeomorphisms of surfaces". In Preparation (cit. on p. 2082).

F. Guéritaud, O. Guichard, F. Kassel, and A. Wienhard (2017). "Anosov representations and proper actions". *Geom. Topol.* 21.1, pp. 485–584. MR: 3608719 (cit. on p. 2089).

A. Hammerlindl (2013). "Leaf conjugacies on the torus". *Ergodic Theory Dynam. Systems* 33.3, pp. 896–933. MR: 3062906 (cit. on p. 2094).

A. Hammerlindl and R. Potrie (2014). "Pointwise partial hyperbolicity in three-dimensional nilmanifolds". *J. Lond. Math. Soc. (2)* 89.3, pp. 853–875. MR: 3217653 (cit. on pp. 2090, 2092, 2094).

– (2015). "Classification of partially hyperbolic diffeomorphisms in 3-manifolds with solvable fundamental group". *J. Topol.* 8.3, pp. 842–870. MR: 3394318 (cit. on pp. 2090, 2092, 2094).

– (Feb. 2017). "Classification of systems with center-stable tori". arXiv: 1702.06206 (cit. on p. 2094).

– (2018). "Partial hyperbolicity and classification: a survey". *Ergodic Theory Dynam. Systems* 38.2, pp. 401–443 (cit. on pp. 2083, 2084, 2089–2091).

A. Hammerlindl, R. Potrie, and M. Shannon (Jan. 2017). "Seifert manifolds admitting partially hyperbolic diffeomorphisms". to appear in JMD. arXiv: 1701.01196 (cit. on pp. 2091, 2093).

M. Handel and W. Thurston (1980). "Anosov flows on new three manifolds". *Invent. Math.* 59.2, pp. 95–103. MR: 577356 (cit. on p. 2086).

A. Hatcher (n.d.). *Notes on basic 3-manifold topology*. Lecture notes available in the author's webpage (cit. on p. 2090).

F. Hertz, J. Hertz, and R. Ures (2011). "Tori with hyperbolic dynamics in 3-manifolds". *J. Mod. Dyn.* 5.1, pp. 185–202. MR: 2787601.

– (2016). "A non-dynamically coherent example on $\mathbb{T}^3$". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 33.4, pp. 1023–1032. MR: 3519530 (cit. on pp. 2086, 2091–2093).

M. Hirsch, C. Pugh, and M. Shub (1977). *Invariant manifolds*. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, pp. ii+149. MR: 0501173 (cit. on pp. 2084, 2091, 2094).

S. Hurtado (July 2016). "The Burnside problem for $\mathrm{Diff}_{\mathrm{Vol}}(\mathbb{S}^2)$". arXiv: 1607.04603 (cit. on p. 2088).

B. Kalinin (2011). "Livšic theorem for matrix cocycles". *Ann. of Math. (2)* 173.2, pp. 1025–1042. MR: 2776369 (cit. on p. 2088).

F. Kassel and R. Potrie (n.d.). "Eigenvalue gaps for hyperbolic groups". In preparation (cit. on p. 2089).

A. Katok and B. Hasselblatt (1995). *Introduction to the modern theory of dynamical systems*. Vol. 54. Encyclopedia of Mathematics and its Applications. With a supplementary chapter by Katok and Leonardo Mendoza. Cambridge University Press, Cambridge, pp. xviii+802. MR: 1326374 (cit. on p. 2081).

A. Kocsard and R. Potrie (2016). "Livšic theorem for low-dimensional diffeomorphism cocycles". *Comment. Math. Helv.* 91.1, pp. 39–64. MR: 3471936 (cit. on p. 2088).

F. Ledrappier (1986). "Positivity of the exponent for stationary sequences of matrices". In: *Lyapunov exponents (Bremen, 1984)*. Vol. 1186. Lecture Notes in Math. Springer, Berlin, pp. 56–73. MR: 850070 (cit. on p. 2088).

C. Lizana and W. Ranter (Nov. 2017). "Topological obstructions for robustly transitive endomorphisms on surfaces". arXiv: 1711.02218 (cit. on p. 2085).

R. Mañé (1978). "Contributions to the stability conjecture". *Topology* 17.4, pp. 383–396. MR: 516217 (cit. on p. 2086).

– (1982). "An ergodic closing lemma". *Ann. of Math. (2)* 116.3, pp. 503–540. MR: 678479 (cit. on p. 2084).

– (1988). "A proof of the C^1 stability conjecture". *Inst. Hautes Études Sci. Publ. Math.* 66, pp. 161–210. MR: 932138 (cit. on p. 2084).

J. Palis and S. Smale (1970). "Structural stability theorems". In: *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*. Amer. Math. Soc., Providence, R.I., pp. 223–231. MR: 0267603 (cit. on p. 2084).

K. Parwani (2010). "On 3-manifolds that support partially hyperbolic diffeomorphisms". *Nonlinearity* 23.3, pp. 589–606. MR: 2586372 (cit. on p. 2090).

R. Potrie (2014). "A few remarks on partially hyperbolic diffeomorphisms of $\mathbb{T}^3$ isotopic to Anosov". *J. Dynam. Differential Equations* 26.3, pp. 805–815. MR: 3274442 (cit. on p. 2097).

– (2015a). "Partial hyperbolicity and foliations in $\mathbb{T}^3$". *J. Mod. Dyn.* 9, pp. 81–121. MR: 3395262 (cit. on pp. 2090, 2092).

– (2015b). "Partially hyperbolic diffeomorphisms with a trapping property". *Discrete Contin. Dyn. Syst.* 35.10, pp. 5037–5054. MR: 3392660 (cit. on p. 2087).

E. Pujals and M. Sambarino (2000). "Homoclinic tangencies and hyperbolicity for surface diffeomorphisms". *Ann. of Math. (2)* 151.3, pp. 961–1023. MR: 1779562 (cit. on p. 2088).

– (2009). "On the dynamics of dominated splitting". *Ann. of Math. (2)* 169.3, pp. 675–739. MR: 2480616 (cit. on p. 2082).

M. Roldán (2016). "Hyperbolic sets and entropy at the homological level". *Discrete Contin. Dyn. Syst.* 36.6, pp. 3417–3433. MR: 3485443 (cit. on p. 2092).

R. Saghin and J. Yang (2016). "Continuity of topological entropy for perturbation of time-one maps of hyperbolic flows". *Israel J. Math.* 215.2, pp. 857–875. MR: 3552297 (cit. on p. 2098).

O. Sarig (2013). "Symbolic dynamics for surface diffeomorphisms with positive entropy". *J. Amer. Math. Soc.* 26.2, pp. 341–426. MR: 3011417 (cit. on p. 2082).

Y. Shi (2014). "Perturbations of partially hyperbolic automorphisms on Heisenberg nilmanifold". PhD thesis. Dijon and Beijing (cit. on p. 2097).

M. Shub (1987). *Global stability of dynamical systems*. With the collaboration of Albert Fathi and Rémi Langevin, Translated from the French by Joseph Christy. Springer-Verlag, New York, pp. xii+150. MR: 869255 (cit. on p. 2081).

– (2006). "All, most, some differentiable dynamical systems". In: *International Congress of Mathematicians. Vol. III*. Eur. Math. Soc., Zürich, pp. 99–120. MR: 2275672 (cit. on p. 2081).

A. Tahzibi and J. Yang (June 2016). "Invariance principle and rigidity of high entropy measures". arXiv: 1606.09429 (cit. on p. 2098).

W. Thurston (Dec. 1997). "Three-manifolds, Foliations and Circles, I". arXiv: math/9712268 (cit. on pp. 2095, 2097).

R. Ures (2012). "Intrinsic ergodicity of partially hyperbolic diffeomorphisms with a hyperbolic linear part". *Proc. Amer. Math. Soc.* 140.6, pp. 1973–1985. MR: 2888185 (cit. on p. 2098).

M. Viana and J. Yang (2013). "Physical measures and absolute continuity for one-dimensional center direction". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 30.5, pp. 845–877. MR: 3103173 (cit. on p. 2098).

– (2017). "Measure-theoretical properties of center foliations". In: *Modern theory of dynamical systems*. Vol. 692. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 291–320. MR: 3666078 (cit. on p. 2098).

A. Wilkinson (2010). "Conservative partially hyperbolic dynamics". In: *Proceedings of the International Congress of Mathematicians. Volume III*. Hindustan Book Agency, New Delhi, pp. 1816–1836. MR: 2827868 (cit. on pp. 2085, 2097).

J. Zhang (Jan. 2017). "Partially hyperbolic diffeomorphisms with one-dimensional neutral center on 3-manifolds". arXiv: 1701.06176 (cit. on p. 2097).

Received 2017-11-28.

Rafael Potrie
CMAT, Facultad de Ciencias
Universidad de la República
Uruguay
rpotrie@cmat.edu.uy
rafaelpotrie@gmail.com

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2105–2130)

THERMODYNAMIC FORMALISM METHODS IN ONE-DIMENSIONAL REAL AND COMPLEX DYNAMICS

Feliks Przytycki

Abstract

We survey some results on non-uniform hyperbolicity, geometric pressure and equilibrium states in one-dimensional real and complex dynamics. We present some relations with Hausdorff dimension and measures with refined gauge functions of limit sets for geometric coding trees for rational functions on the Riemann sphere. We discuss fluctuations of iterated sums of the potential $-t \log |f'|$ and of radial growth of derivative of univalent functions on the unit disc and the boundaries of range domains preserved by a holomorphic map f repelling towards the domains.

Contents

Partially supported by Polish NCN Grant 2014/13/B/ST1/01033.
MSC2010: primary 37D35; secondary 37E05, 37F10, 37F35, 31A20.
Keywords: one-dimensional dynamics, geometric pressure, thermodynamic formalism, equilibrium states, volume lemma, radial growth, Hausdorff measures, Law of Iterated Logarithm, Lyapunov exponents, harmonic measure.

1 Thermodynamic formalism, introductory notions

Among founders of this theory are Sinaĭ [1972], Bowen [1975] and David Ruelle, who wrote in Ruelle [1978]: "thermodynamic formalism has been developed since G. W. Gibbs to describe [...] physical systems consisting of a large number of subunits". In particular one considers a *configuration space* Ω of functions $\mathbb{Z}^n \to \mathbb{A}$ on the lattice $\mathbb{Z}^n$ with interacting values in $\mathbb{A}$ over its sites, e.g. "spin" values in the Ising model of ferromagnetism. One considers probability distributions on Ω, invariant under translation, called *equilibrium states* for potential functions on Ω.

Given a mapping $f : X \to X$ one considers as a configuration space the set of trajectories $n \mapsto (f^n(x))_{n\in\mathbb{Z}_+}$ or $n \mapsto \Phi(f^n(x))_{n\in\mathbb{Z}_+}$ for a test function $\Phi : X \to Y$.

The following simple fact Bowen [1975, Lemma 1.1] and Ruelle [1978, Introduction], Przytycki and Urbański [2010, Introduction], resulting from Jensen's inequality applied to the function logarithm, stands at the heart of thermodynamic formalism.

Lemma 1.1 (Finite Variational Principle)**.** *For given real numbers $\phi_1, \dots, \phi_d$, the function $F(p_1, \dots p_d) := \sum_{i=1}^n -p_i \log p_i + \sum_{i=1}^d p_i\phi_i$ defined on the simplex $\{(p_1, \dots, p_d) : p_i \geq 0, \sum_{i=1}^d p_i = 1\}$ attains its maximum value $P(\phi_1, \dots, \phi_d) = \log \sum_{i=1}^d e^{\phi_i}$ at and only at $\widehat{p}_j = e^{\phi_j} \left(\sum_{i=1}^d e^{\phi_i}\right)^{-1}$.*

We can read $i \mapsto \phi_i, i = 1, \dots, d$ as a *potential* function and $\widehat{p}_i$ as the equilibrium probability distribution on the finite space $\{1, \dots, d\}$. $P(\phi_1, \dots, \phi_d)$ is called the *pressure* or *free energy*, see Ruelle [1978].

Let $f : X \to X$ be a continuous mapping of a compact metric space X and $\phi : X \to \mathbb{R}$ be a continuous function (the potential). We define the *topological pressure* or free energy by

Definition 1.2.

$$P_{\text{var}}(f, \phi) = \sup_{\mu\in\mathcal{M}(f)} \left(h_\mu(f) + \int_X \phi \, d\mu\right), \tag{1-1}$$

where $\mathcal{M}(f)$ is the set of all f-invariant Borel probability measures on X and $h_\mu(f)$ is measure theoretical entropy. Sometimes we write $\mathcal{M}(f, X)$.

Recall that $h_\mu(f) = \sup_{\mathbb{A}} \lim_{n\to\infty} \frac{1}{n+1} \sum_{A\in\mathbb{A}^n} -\mu(A)\log\mu(A)$, where the supremum is taken over finite partitions $\mathbb{A}$ of X, where $\mathbb{A}^n := \bigvee_{j=0,\dots,n} f^{-j}\mathbb{A}$. Notice that this resembles the sum $\sum_{i=1}^n -p_i \log p_i$ in Lemma 1.1.

Topological pressure can also be defined in other ways, e.g. by (6-2), and then its equality to the one given by (1-1) is called the variational principle. This explains the notation P_{var}. Any $\mu \in \mathcal{M}(f)$ for which the supremum in (1-1) is attained is called *equilibrium*, *equilibrium measure* or *equilibrium state*.

A model case is any map $f : U \to \mathbb{R}^n$ of class C^1, defined on a neighbourhood U of a compact set $X \subset \mathbb{R}^n$, *expanding* (another name: *uniformly expanding* or *hyperbolic* in dimension 1) that is there exist $C > 0, \lambda > 1$ such that for all positive integers n all $x \in X$ and all v tangent to $\mathbb{R}^n$ at x,

$$||Df^n(v)|| \geq C\lambda^n ||v||, \tag{1-2}$$

and *repelling* that is every forward trajectory sufficiently close to X must be entirely in X. Not assuming the differentiability of f one uses the notion of *distance expanding* meaning the increase of distances under the action of f by a factor at least $\lambda > 1$ for pairs of distinct points sufficiently close to each other. Repelling happens to be equivalent to the internal condition: $f|_X$ being an open map, provided f is open on a neighbourhood of X, see Przytycki and Urbański [2010, Lemma 6.1.2]. Then the classical theorem holds, here in the version from Przytycki and Urbański [ibid., Section 5.1]:

Theorem 1.3. *Let $f : X \to X$ be a distance expanding, topologically transitive continuous open map of a compact metric space X and $\phi : X \to \mathbb{R}$ be a Hölder continuous potential. Then, there exists exactly one measure $\mu_\phi \in \mathcal{M}(f, X)$, called the Gibbs measure, satisfying*

$$C < \frac{\mu_\phi(f_x^{-n}(B(f^n(x), r_0))}{\exp(S_n\phi(x) - nP(\phi))} < C^{-1} \tag{1-3}$$

where f_x^{-n} is the branch of f^{-n} mapping $f^n(x)$ to x (locally making sense, since f is a local homeomorphism) and $S_n\phi(x) := \sum_{j=0}^{n-1}\phi(f^j(x))$.

The measure μ_ϕ is the only equilibrium state for ϕ. It is equivalent to the unique ϕ-conformal measure m_ϕ, that is a forward quasi-invariant Borel probability measure m_ϕ with Jacobian $\exp -(\phi - P(\phi))$. *Moreover, the limit $P(\phi) = P(f,\phi) :=$* $\lim_{n\to\infty} \frac{1}{n} \log \sum_{x\in f^{-n}(x_0)} \exp S_n\phi(x)$ *exists and is equal to $P_{\text{var}}(f,\phi)$ for every $x \in X$.*

This $P(\phi)$ is a normalizing quantity corresponding to $P(\phi_1, \dots, \phi_d)$ in Lemma 1.1 and the sum in the definition of $P(\phi)$ corresponds to the so called *statistical sum* over the space Ω_n of all admissible configurations over $\{0, 1, \dots, n-1\}$, as in the Ising model. Compare to the *tree pressure* defined in Definition 6.2.

So $\varsigma : \Sigma^d \to \Sigma^d$, the shift to the left on the space $\Sigma^d = \{(\alpha_0, \alpha_1, \dots) : \alpha_j \in \{1, \dots, d\}\}$, defined by $\varsigma((\alpha_n)) = (\alpha_{n+1})$, is an example where Theorem 1.3 holds. The sets $f_x^{-n}(B(f^n(x), r_0)$ correspond to *cylinders* of fixed $\{\alpha_j \in \{1, \dots, d\}, j = 0, \dots, n-1\}$. One can impose an admissibility condition: $\alpha_i \alpha_{i+1}$ admissible if the pair has the digit 1 attributed in a defining 0,1 $d \times d$ matrix. Then one calls the system a *one-sided topological Markov chain*.

The condition of openness of f can be replaced by a weaker one: the existence of a finite Markov partition, see Przytycki and Urbański [2010].

The existence of a conformal measure follows from the existence of a fixed point in the convex weakly*-compact set of probability measures for the dual operator to the transfer (Perron-Frobenius-Ruelle) operator $\mathcal{L}$ divided by the norm, where for $u : X \to \mathbb{R}$ continuous one defines

$$\mathcal{L}(u)(x) := \sum_{y \in f^{-1}(x)} u(y) \exp \phi(y). \tag{1-4}$$

Indeed, for every Borel set $Y \subset X$ on which f is injective, denoting by I_Y indicator function: 1 on Y, 0 outside Y, due to an approximation by continuous functions, one has for every finite Borel measure ν on X

$$(\mathcal{L}^*(\nu))(Y)) = \int_X \mathcal{L}(I_Y)\, d\nu = \int_{f(Y)} \exp \phi \circ f|_Y^{-1}\, d\nu. \tag{1-5}$$

Hence the (positive) eigen-measure m_ϕ has Jacobian for $(f|_Y)^{-1}$ equal to $\exp(\phi \circ f|_Y^{-1})/\lambda$, hence f has Jacobian $\exp(-\phi)$ multiplied by an eigenvalue $\lambda := \exp P(\phi)$.

The proof of the existence of an invariant Gibbs measure equivalent to m_ϕ is harder. One first proves the existence of a positive eigenfunction u_ϕ for $\mathcal{L}$ and then defines $\mu_\phi = u_\phi m_\phi$. For a more complete introduction to this theory, see e.g. Przytycki and Urbański [ibid.].

2 Introduction to dimension 1

Thermodynamic formalism is useful for studying properties of the underlying space X. In dimension 1, for f real of class $C^{1+\varepsilon}$ or f holomorphic, for an expanding repeller X, considering $\phi = \phi_t := -t \log |f'|$ for $t \in \mathbb{R}$, (1-3) gives

$$\begin{aligned} \mu_{\phi_t}(f_x^{-n}(B(f^n(x), r_0))) \approx \exp(S_n\phi(x) - nP(\phi_t)) \approx \\ \operatorname{diam} f_x^{-n}(B(f^n(x), r_0))^t \exp -nP(\phi_t). \end{aligned} \tag{2-1}$$

The latter follows from a comparison of the diameter with the inverse of the absolute value of the derivative of f^n at x, due to *bounded distortion*. Here, the symbol "$\approx$" denotes that the mutual ratios are bounded by a constant.

When $t = t_0$ is a zero of the function $t \mapsto P(\phi_t)$, this gives

$$\mu_{\phi_{t_0}}(B) \approx (\operatorname{diam} B)^{t_0} \tag{2-2}$$

for all small balls B (the t_0-Ahlfors measure property). We obtain the so-called Bowen's formula for Hausdorff dimension:

$$\mathrm{HD}(X) = t_0. \tag{2-3}$$

Moreover, the Hausdorff measure of X in this dimension is finite and nonzero.

A model example of application is the proof of

Theorem 2.1. *For $f_c(z) := z^2 + c$ for an arbitrary complex number $c \neq 0$ sufficiently close to 0, the invariant Jordan curve J (Julia set for f_c) is a fractal, i.e. has Hausdorff dimension bigger than 1.*

Sketch of Proof. $t_0 > 1$ yields $\mathrm{HD}(J) = t_0 > 1$ by (2-2) (one does not need to use the invariance of $\mu_{\phi_{t_0}}$).

The case $t_0 = 1$ yields by (2-2) finite Hausdorff measure in dimension 1, i.e. the rectifiability of J. To conclude that J is a circle and $c = 0$, one can use ergodic invariant measures in the classes of harmonic ones on J from inside and outside. They must coincide. This relies on Birkhoff's Ergodic Theorem, the heart of ergodic theory. This is an "echo" of the celebrated Mostov Rigidity Theorem. See Sullivan [1982] and Przytycki and Urbański [2010, Theorem 9.5.5]. □

In dimension 1 (real or complex), we call c a critical point if the derivative $f'(c) = 0$. The set of critical points will be denoted by $\mathrm{Crit}(f)$.

In this survey, we allow for the presence of critical points and concentrate mainly on two cases:

1. (Complex case) f is a rational mapping of degree at least 2 of the Riemann sphere $\overline{\mathbb{C}}$. We consider f acting on its Julia set $K = J(f)$.

For entire or meromorphic maps see e.g. Barański, Karpińska, and Zdunik [2009, 2012], compare Proposition 5.2.

2. (Real case) f is a *generalized multimodal map* defined on a neighbourhood $U_K \subset \mathbb{R}$ of its compact invariant subset K. We assume that $f \in C^2$, is non-flat at all of its turning and inflection critical points, satisfies the *bounded distortion* property for iterates, abbr.

BD, see Przytycki and Rivera-Letelier [2014], is topologically transitive and has positive topological entropy on K.

We assume that K is a maximal invariant subset of a finite union of pairwise disjoint closed intervals $\widehat{I} = I^1 \cup \cdots \cup I^k \subset U_K$ whose endpoints are in K. (This maximality corresponds to the Darboux property, compare Przytycki and Rivera-Letelier [ibid., Appendix A] and Misiurewicz and Szlenk [1980, page 49].) We write $(f, K) \in \mathcal{A}_+^{\mathrm{BD}}$, with the subscript + to mark positive entropy. In place of BD one can assume C^3 (and write $(f, K) \in \mathcal{A}_+^3$), and assume that all periodic orbits in K are hyperbolic repelling. Indeed, changing f outside K if necessary, one can get the corrected (f, K) in $\mathcal{A}_+^{\mathrm{BD}}$.

Recall the notions concerning periodic orbits: *Parabolic* means $f^n(p) = p$ with $(f^n)'(p)$ being a root of unity. For $|(f^n)'(p)| = 1$ the term *indifferent periodic* is used and for $|(f^n)'(p)| > 1$ the term *hyperbolic repelling*. If $|(f^n)'(p)| < 1$ the orbit is called *hyperbolic attracting*.

For the real setting, see Przytycki and Rivera-Letelier [2014], Gelfert, Przytycki, and Rams [2016] and Przytycki [2018]. Examples are provided by basic sets in the spectral decomposition de Melo and van Strien [1993].

Question. Are there any other examples?

Problem. Generalize the real case theory, see further sections, to the piecewise continuous maps, that is allow the intervals I^j to have common ends (see Hofbauer and Urbański [1994] for some results in this direction).

In this survey, we compare equilibrium states to (refined) Hausdorff measures in the complex case. For the real case, we refer the reader to Hofbauer and Keller [1993] and the references therein.

3 Hyperbolic potentials

For general $f : X \to X$ and $\phi : X \to \mathbb{R}$ as in Definition 1.2 the following conditions are of special interest Inoquio-Renteria and Rivera-Letelier [2012],

1) $P(f, \phi) > \sup \phi$,
2) $P(f^n, S_n\phi) > \sup_X S_n\phi$ for an integer n,
3) $P(f, \phi) > \sup_{\nu \in \mathcal{M}(f)} \int \phi \, d\nu$,
4) For each equilibrium state μ for the potential ϕ, the entropy $h_\mu(f)$ is positive.

The conditions 2) – 4) are equivalent, see Inoquio-Renteria and Rivera-Letelier [ibid., Proposition 3.1]. Potentials ϕ satisfying them have been called in Inoquio-Renteria and Rivera-Letelier [ibid.] *hyperbolic*. The condition 1) has longer traditions, see Denker and Urbański [1991]. The intuitive meaning is that no minority of trajectories carries the full pressure.

For every $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ rational of degree at least 2 and $\phi : J(f) \to \mathbb{R}$ Hölder continuous, the following condition is also equivalent to 2) – 4), see Inoquio-Renteria and Rivera-Letelier [2012]:

5) For each ergodic equilibrium state μ for ϕ, the Lyapunov exponent $\chi(\mu) := \int \log |f'| \, d\mu$ is positive, that is for μ-a.e. x,

$$\chi(\mu) = \chi(x) := \lim_{n \to \infty} \frac{1}{n} \log |(f^n)'(x)| > 0.$$

The conditions 2)-5) are also equivalent in the real case for $(f, K) \in \mathcal{A}_+^{\mathrm{BD}}$ or $(f, K) \in \mathcal{A}_+^3$ and all periodic orbits hyperbolic repelling. The arguments in Inoquio-Renteria and Rivera-Letelier [ibid.] work. See also Li and Rivera-Letelier [2014b].

Theorem 3.1. *Let $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ be a rational mapping as above. If ϕ is a Hölder continuous hyperbolic potential on $J(f)$, then there exists a unique equilibrium state μ_ϕ. For every Hölder $u : J(f) \to \mathbb{R}$, the Central Limit Theorem (abbr. CLT) for the sequence of random variables $u \circ f^n$ and μ_ϕ holds.*

For a proof, see Przytycki [1990] and preceding Denker and Urbański [1991]. To find this equilibrium one can iterate the transfer operator proving $\mathcal{L}^n(\mathbb{1})/\exp nP(f, \phi) \to u_\phi$. The convergence is uniformly $\exp -\sqrt{n}$ fast and the limit is Hölder continuous, Denker, Przytycki, and Urbański [1996]. Finally, define $\mu_\phi := u_\phi \cdot m_\phi$, as at the end of Section 1.

Remark 3.2. Given μ_ϕ a priori, an efficient way to study it is an inducing method, see Szostakiewicz, Urbański, and Zdunik [2015], i.e. the use of a return map $A \ni x \mapsto f^{n(x)}(x) \in A$ for A and $n(x)$ adequate to μ_ϕ. Then one proves even an exponential convergence (with any u Hölder in place of $\mathbb{1}$), which yields exponential mixing, hence stochastic laws for $u \circ f^n$ for Hölder u, e.g. CLT, LIL, compare Sections 9 and 10. See also Remark 5.4. The key feature is the exponential decay of $\mu_\phi(A_n)$, where $A_n := \{x \in A : n(x) \geq n\}$.

See also Bruin and Todd [2008] for the real case, and stronger Li and Rivera-Letelier [2014a] and Li and Rivera-Letelier [2014b] including also the complex case proving the exponential convergence to u_ϕ, hence CLT and LIL. See also Szostakiewicz, Urbański, and Zdunik [2014] for endomorphisms f of higher dimensional complex projective space, where 1) is replaced by a stronger "gap" assumption.

4 Non-uniform hyperbolicity in real and complex dimension 1

Here we discuss a set of conditions, valid in both the real and complex situations. Below we concentrate on the case of complex rational maps with $K = J(f)$, only remarking differences in the real case.

(a) CE. *Collet-Eckmann condition.* There exist $\lambda_{CE} > 1$ and $C > 0$ such that for every critical point c in $J(f)$, whose forward orbit does not meet other critical points, for every $n \geq 0$ we have

$$|(f^n)'(f(c))| \geq C\lambda_{CE}^n.$$

Moreover, there are no parabolic (indifferent) periodic orbits.

(b) CE2(z_0). *Backward* or *second Collet-Eckmann condition at* $z_0 \in J(f)$. There exist $\lambda_{CE2} > 1$ and $C > 0$ such that for every $n \geq 1$ and every $w \in f^{-n}(z_0)$ (in a neighbourhood of K in the real case)

$$|(f^n)'(w)| \geq C\lambda_{CE2}^n.$$

(b') CE2. *The second Collet-Eckmann condition.* CE2(c) holds for all critical points c not in the forward orbit of any other critical point.

(c) TCE. *Topological Collet-Eckmann condition.* There exist $M \geq 0, P \geq 1, r > 0$ such that for every $x \in K$ there exists a strictly increasing sequence of positive integers n_j, $j = 1, 2, \dots$, such that $n_j \leq P \cdot j$ and for each j (and discs $B(\cdot)$ below understood in $\overline{\mathbb{C}}$ or $\mathbb{R}$)

$$\#\{0 \leq i < n_j : \mathrm{Comp}_{f^i(x)} f^{-(n_j-i)} B(f^{n_j}(x), r)) \cap \mathrm{Crit}(f) \neq \varnothing\} \leq M, \tag{4-1}$$

where in general $\mathrm{Comp}_z V$ means for $z \in V$ the component of V containing z.

In the real case, one adds the condition that there are no parabolic periodic orbits, which is automatically true in the case of complex rational maps.

(d) ExpShrink. *Exponential shrinking of components.* There exist $\lambda_{\mathrm{Exp}} > 1$ and $r > 0$ such that for every $x \in K$, every $n > 0$ and every connected component W_n of $f^{-n}(B(x,r))$ for the disc (interval) $B(x,r)$ in $\overline{\mathbb{C}}$ (or $\mathbb{R}$), intersecting K

$$\mathrm{diam}(W_n) \leq \lambda_{\mathrm{Exp}}^{-n}. \tag{4-2}$$

(e) LyapHyp (*Lyapunov hyperbolicity*). There is a constant $\lambda_{\mathrm{Lyap}} > 1$ such that the Lyapunov exponent $\chi(\mu)$ of any ergodic measure $\mu \in \mathcal{M}(f, K)$ satisfies $\chi(\mu) \geq \log \lambda_{\mathrm{Lyap}}$.

(f) UHP. *Uniform Hyperbolicity on periodic orbits.* There exists $\lambda_{\mathrm{Per}} > 1$ such that every periodic point $p \in K$ of period $k \geq 1$ satisfies

$$|(f^k)'(p)| \geq \lambda_{\mathrm{Per}}^k.$$

We distinguish LyapHyp as the most adequate among these conditions to carry the name (strong) non-uniform hyperbolicity.[1]

Theorem 4.1. *1. The conditions (c)–(f) and else (b) for some z_0 are equivalent in the complex case. In the real case, the equivalence also holds under the assumption of weak isolation (see the definition below).*

2. In the complex case, the suprema over all possible constants λ_{Exp}, λ_{CE2} (supremum over all z_0), λ_{Per} and λ_{Lyap} coincide.

3. Both CE and CE2 imply (c)–(f).

4. If there is only one critical point in the Julia set in the complex case or if f is S-unimodal on $K = I$ in the real case, i.e. has just one turning critical point c and negative Schwarzian derivative on $I \setminus \{c\}$, then all conditions above are equivalent to each other.

For more details, see Przytycki, Rivera-Letelier, and Smirnov [2003], Rivera-Letelier [2012] and Przytycki and Rivera-Letelier [2014].

Definition 4.2. $(f, K) \in \mathcal{A}$ is said to be *weakly isolated* if there exists an open neighbourhood U of K in the domain of f such that for every f-periodic orbit $O(p) \subset U$ is contained in K.

In the complex case, we can replace (4-1) by

$$\deg\Big(f^{n_j}|_{\mathrm{Comp}_x f^{-n_j}(B(f^{n_j}(x),r))}\Big) \le M'$$

for a constant M'. In the real case, this condition is weaker than (4-1) since f mapping W_{n+1} into W_n may happen not surjective. It can have folds, thus truncating backward trajectories of critical points acquired before when pulling back.

In the real case, the proof of CE⇒TCE can be found in Nowicki and Przytycki [1998]. For the complex case, we refer the reader to Przytycki and Rohde [1998].

The implication TCE⇒CE was proved in the complex case in Przytycki [2000, Theorem 4.1]. The proof used the idea of the "reversed telescope" by Graczyk and Smirnov [1998]. In the real case, this implication was proved for S-unimodal maps in Nowicki and Sands [1998]. In presence of more than one critical point this implication may be false, see Przytycki, Rivera-Letelier, and Smirnov [2003, Appendix C].

Question. Is this implication true for every $(f, K) \in \mathcal{A}_+^{\mathrm{BD}}$ with one critical point, provided it is weakly isolated? See Definition 4.2. It seems that the answer is yes.

[1]Then all Hölder continuous potentials are hyperbolic, see Condition 5) in Section 3 and Inoquio-Renteria and Rivera-Letelier [2012].

Since the condition TCE is stated in purely topological terms (in the class of maps without indifferent periodic orbits), it is invariant under topological conjugacy. So we obtain the following immediate corollary.

Corollary 4.3. *All equivalent conditions listed above are invariant under topological conjugacies between (f, K)'s).*

Another proof of the topological invariance of CE in the complex case was provided in Przytycki and Rohde [1999] with the use of Heinonen and Koskela criterion for quasiconformality, Heinonen and Koskela [1995].

Note that this topological invariance is surprising, as all the conditions except TCE are expressed in geometric-differential terms. I do not know how to express CE for unimodal maps of interval in the (topological-combinatorial) kneading sequence terms.

An important lemma used here has been an estimate of an average distance in the logarithmic scale of every orbit from $\mathrm{Crit}(f)$, see Denker, Przytycki, and Urbański [1996]. Namely

Lemma 4.4.

$$\sum_{j=0}^{n}{}' -\log|f^j(x)-c| \leq Qn \tag{4-3}$$

for a constant $Q > 0$ every $c \in \mathrm{Crit}(f)$, every $x \in K$ and every integer $n > 0$. Σ' means that we omit in the sum an index j of smallest distance $|f^j(x)-c|$.

An order of proving the equivalences in Theorem 4.1 is
$\mathrm{CE2}(z_0)$ ⇒ExpShrink⇒LyapHyp⇒UHP⇒$\mathrm{CE2}(z_0)$ and separately $\mathrm{CE2}(z_0)$ ⇔TCE.
E.g. assumed UHP one proves $\mathrm{CE2}(z_0)$ by "shadowing", compare the beginning of Section 6.

5 Geometric pressure and equilibrium states

We go back to topological pressure, Definition 1.2, but for $\phi = -t\log|f'|$, $t \in \mathbb{R}$ in the complex $K = J(f)$ or real cases, where ϕ can attain the values $\pm\infty$ at the critical points of f. See the beginning of Section 2. We call it the geometric pressure, because it is useful in studying of geometry of the underlying space, e.g. as in (2-3) via equilibrium states for all t.

The definition of $P_{\mathrm{var}}(f, -t\log|f'|)$ in Definition 1.2 makes sense due to $\chi(\mu) \geq 0$ for all $\mu \in \mathcal{M}(f)$, in particular due to the integrability of $\log|f'|$, see Przytycki [1993] and Rivera-Letelier [2012, Appendix A] for a simpler proof. We conclude that it is convex

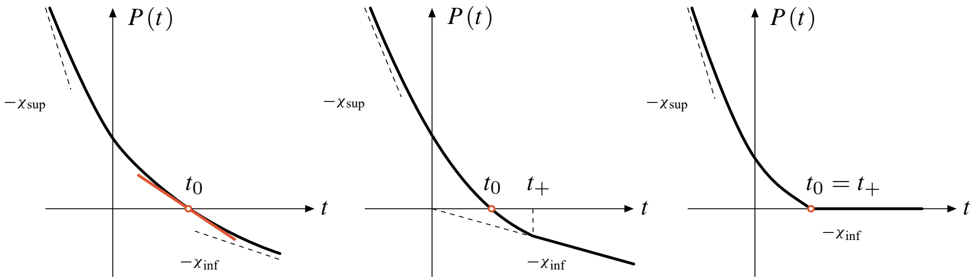


Figure 1: The geometric pressure: LyapHyp with $t_+ = \infty$, LyapHyp with $t_+ < \infty$, and non-LyapHyp. This Figure is taken from Gelfert, Przytycki, and Rams [2016], see notation in Remarks below.

and monotone decreasing. We start by defining a quantity occurring equal to $P(t) = P_{\mathrm{var}}(t) := P_{\mathrm{var}}(f, -t \log|f'|)$, to explain its geometric meaning, compare with Section 2.

Definition 5.1 (Hyperbolic pressure)**.**

$$P_{\mathrm{hyp}}(t) := \sup_{X \in \mathcal{H}(f,K)} P(f|_X, -t \log|f'|),$$

where $\mathcal{H}(f, K)$ is defined as the space of all compact forward f-invariant (that is $f(X) \subset X$) hyperbolic subsets of K, repellers in $\mathbb{R}$.

From this definition, it immediately follows that:

Proposition 5.2. (Generalized Bowen's formula, compare (2-3)) *The first zero t_0 of $t \mapsto P_{\mathrm{hyp}}(K,t)$ is equal to the hyperbolic dimension* $\mathrm{HD}_{\mathrm{hyp}}(K)$ *of K, defined by* $\mathrm{HD}_{\mathrm{hyp}}(K) := \sup_{X \in \mathcal{H}(f,K)} \mathrm{HD}(X)$.

For the discussion $\mathrm{HD}_{\mathrm{hyp}}(J(f))$ vs $\mathrm{HD}(J(f))$, see Lyubich [2014, Section 2.13].

Below we state Theorem 5.3 proved in Przytycki and Rivera-Letelier [2011] in the complex setting and in Przytycki and Rivera-Letelier [2014] in the real setting. It extends Bruin and Todd [2009] and Pesin and Senti [2008] and Iommi and Todd [2010]. See also impressive Dobbs and Todd [2015].

Theorem 5.3. *1. Real case, Przytycki and Rivera-Letelier [2014]. Let $(f, K) \in \mathcal{A}_+^3$ and let all f-periodic orbits in K be hyperbolic repelling. Then $P(t)$ is real analytic on the open interval bounded by the "phase transition parameters" t_- and t_+. For every $t \in (t_-, t_+)$, the domain where*

$$P(t) > \sup_{\nu \in \mathcal{M}(f)} -t \int \log|f'|\, d\nu, \tag{5-1}$$

there is a unique invariant equilibrium state. It is ergodic and absolutely continuous with respect to an adequate conformal measure m_{ϕ_t} with the density bounded from below by a positive constant almost everywhere. If furthermore f is topologically exact on K (that is for every V an open subset of K there exists $n \geq 0$ such that $f^n(V) = K$), then this measure is mixing, has exponential decay of correlations and it satisfies the Central Limit Theorem for Lipschitz gauge functions.

2. Complex case, Przytycki and Rivera-Letelier [2011]. The assertion is the same. One assumes a very weak expansion: the existence of arbitrarily small nice, or pleasant, couples and hyperbolicity away from critical points.

Remarks. 1) t_- and t_+ are called the phase transition parameters. Since $P(0) = h_{\text{top}}(f) > 0$, $t_- < 0 < t_+$, they need not exist; we say then they are equal to $-\infty$ and/or $+\infty$ respectively. $P(t)$ is linear to the left of t_- and to the right of t_+, equal to $t \mapsto -t\chi_{\sup}$ where $\chi_{\sup} := \sup_\nu \chi(\nu)$ and $t \mapsto -t\chi_{\inf}$, where $\chi_{\inf} := \inf_\nu \chi(\nu)$, respectively. Of course, $P(t)$ is not real-analytic at finite t_- and t_+.

2) For $f(z) = z^2 - 2$, $f : [-2, 2] \to [-2, 2]$ (the Tchebyshev polynomial), we have $f(2) = 2$, $f'(2) = 4$, $\chi(l) = \log 2$, where l is the normalized length measure. We have $P(t) = \log 2 - t \log 2$ for $t \geq -1$ and $P(t) = -t \log 4$ for $t \leq -1$, so $t_- = -1$, $P(t)$ is non-differentiable at t_- and for $t = -1$ there are two ergodic equilibrium states: Dirac at $z = 2$ and l.

3) For any f non-LyapHyp, $t_+ = t_0 < \infty$. However $t_+ < \infty$ can happen even for f LyapHyp, see N. Makarov and Smirnov [2003] and Coronel and Rivera-Letelier [2013, 2015].

4) Notice that the condition (5-1) is similar to the condition 3) from Section 3. For f LyapHyp and $t > t_+$, no equilibrium state can exist, see Inoquio-Renteria and Rivera-Letelier [2012].

5) For real f as in Theorem 5.3 satisfying LyapHyp and $K = \widehat{I}$, we have $t_0 = 1$ and for $-\log|f'|$ we conclude that a unique equilibrium state exists which is a.c.i.m.(that is: invariant absolutely continuous with respect to Lebesgue measure). In fact this assertions hold even for $t = t_0 = t_+ = 1$ with very weak hyperbolicity properties e.g. $|(f^n)'(f(c))| \to \infty$ for all $c \in \text{Crit}(f)$, see Bruin, Rivera-Letelier, Shen, and van Strien [2008] and Shen and van Strien [2014]. For the complex case, see Graczyk and Smirnov [2009] and stronger Rivera-Letelier and Shen [2014].

Remark 5.4. In the proof of Theorem 5.3, we use (compare with the Remark 3.2) a return map $F(x) = f^{n(x)}$ to a "nice" (Markov) domain. However unlike in Szostakiewicz, Urbański, and Zdunik [2015], we do not use in the construction of this set the equilibrium measure μ_ϕ because we do not know a priori that it exists. The construction is geometric. F is an infinite Iterated Function System, more precisely the family of all branches of F^{-1}

is, see Mauldin and Urbański [2003] and Pesin [2014] and references therein, expanding due to the "acceleration" from f to F. Then we consider an equilibrium state $\mathtt{P}$ for (F, Φ) where $\Phi(x) := \sum_{j=0}^{n(x)-1} \phi_t(f^j(x))$, and consider an equivalent conformal measure. We propagate these measures to the Lai-Sang Young tower $\{(x, j) : 0 \le j < n(x)\}$ and project by $(x, j) \mapsto f^j(x)$ to K.[2]

Stochastic properties of $\mathtt{P}$ stay preserved along the construction to μ_ϕ. The analyticity of $P(t)$ follows from expressing $P(t)$ as zero of a pressure for F with potential depending on two parameters and Implicit Function Theorem. The latter idea came from Stratmann and Urbanski [2003].

Remark 5.5. For probability measures μ_n weakly* convergent to some $\widehat{\mu}$, in presence of critical points $\int \log|f'|\,d\mu_n$ need not converge to $\int \log|f'|\,d\widehat{\mu}$. Only upper semicontinuity holds. Therefore, for $t > 0$, the equilibrium states for $t_n \to t$ need not converge to an equilibrium state for t. A priori, the free energy in the Definition 1.2 can jump down. However, a modification of this method to prove existence of equilibria works, see Dobbs and Todd [2015].

Notice also that passing to a weak*-limit with averages of Dirac measures on $\{x, \dots, f^n(x)\}$ proves $\limsup_{n\to\infty} \sup_{x\in K} \frac{1}{n} S_n(\log|f'|)(x) \le \chi_{\max}$. However an analogous inequality $\liminf \cdots \ge \chi_{\inf}$ is obviously false. These observations contribute to the understanding of Lyapunov spectrum.

Remarks on the Lyapunov spectrum. Theorem 5.3 allows us to express the so-called dimension spectrum for Lyapunov exponents with the use of Legendre transform, that is for all $\alpha > 0$ and $\mathcal{L}(\alpha) := \{x \in K : \chi(x) = \alpha\}$

$$\mathrm{HD}(\mathcal{L}(\alpha)) = \frac{1}{|\alpha|} \inf_{t\in\mathbb{R}} (P(t) + \alpha t). \tag{5-2}$$

An ingredient is **Mañé's equality**

$$\mathrm{HD}(\mu) = h_\mu(f)/\chi(\mu) \tag{5-3}$$

provided $\chi(\mu) > 0$, Przytycki and Urbański [2010], where $\mathrm{HD}(\mu) := \sup\{\mathrm{HD}(X) : \mu(X) = 1\}$, applied to μ_{ϕ_t}.

The equality (5-2) concerns regular x's, where $\chi(x) = \lim_{n\to\infty} \frac{1}{n} \log|(f^n)'(x)|$ exists. It is also possible to provide formulas or at least estimates for Hausdorff dimension of the sets of irregular points $\mathcal{L}(\alpha,\beta) := \{x \in K : \underline{\chi}(x) = \alpha, \overline{\chi}(x) = \beta\}$ for lower and upper Lyapunov exponents where we replace lim by lim inf and lim sup respectively. See

[2] For applications to decide the existence or nonexistence of a finite a.c.i.m. for maps of interval with flat critical points or for entire or meromorphic maps depending on the $\mathtt{P}$-integrability of the first return time, see papers by N. Dobbs, B. Skorulski, J. Kotus, G. Świątek.

Gelfert, Przytycki, and Rams [2010] and Gelfert, Przytycki, and Rams [2016] for this theory in complex and real settings.

However, these papers give no information about the size of sets with zero (upper) Lyapunov exponent. Note at least that if $J(f) \neq \overline{\mathbb{C}}$ then $\mathrm{Leb}_2\{x \in J(f) : \overline{\chi}(x) > 0\} = 0$. This is so because $\overline{\chi}(x) > 0$ implies there exists $\mathfrak{N} \subset \mathbb{Z}_+$ of positive upper density, such that for $n \in \mathfrak{N}$, (4-2) and (4-1) hold, see Levin, Przytycki, and Shen [2016, Section 3].

We do not know whether $\chi(x) = -\infty$ can happen for x not pre-critical, except there is only one critical point in K, where $\chi(x) > -\infty$ follows from (4-3), see Gelfert, Przytycki, and Rams [2010, Lemma 6].

For x being a critical value we can prove (in analogy to $\chi(\mu) \geq 0$):

Theorem 5.6 (Levin, Przytycki, and Shen [2016])**.** *If for a rational function $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ there is only one critical point c in $J(f)$ and no parabolic periodic orbits, then $\underline{\chi}(f(c)) \geq 0$.*

For S-unimodal maps of interval this was proved by Nowicki and Sands [1998].

6 Other definitions of geometric pressure

Definition 6.1 (safe)**.** See Przytycki and Urbański [2010, Definition 12.5.7]. We call $z \in K$ *safe* if $z \notin \bigcup_{j=1}^{\infty}(f^j(\mathrm{Crit}(f)))$ and for every $\delta > 0$ and all n large enough $B(z, \exp(-\delta n)) \cap \bigcup_{j=1}^{n}(f^j(\mathrm{Crit}(f))) = \varnothing$.

Notice that this definition implies that all points except at most a set of Hausdorff dimension 0, are safe.

Definition 6.2 (Tree pressure)**.** For every $z \in K$ and $t \in \mathbb{R}$ define

$$P_{\mathrm{tree}}(z,t) = \limsup_{n\to\infty} \frac{1}{n} \log \sum_{f^n(x)=z,\, x\in K} |(f^n)'(x)|^{-t}. \tag{6-1}$$

Compare with $P(f, \phi)$ from Theorem 1.3. Under suitable conditions, e.g. for z "safe" the limit exists, it is independent of z and equal to $P(t)$. See Przytycki [1999], Przytycki, Rivera-Letelier, and Smirnov [2003] and Przytycki and Urbański [2010] for the complex case and Przytycki and Rivera-Letelier [2014] and Przytycki [2018] for the real case.

A key is to extend all trajectories $T_n(x) = \{x, \ldots, z\}$ backward and forward by time $m \ll n$ to get an Iterated Function System for f^{n+m} and to consider its limit set. Its trajectories for time n "shadow" $T_n(x)$. This proves $P_{\mathrm{tree}}(z,t) \leq P_{\mathrm{hyp}}(t)$. The opposite inequality is immediate.

(Similarly one proves $P_{\mathrm{var}}(t) \leq P_{\mathrm{hyp}}(t)$. Given μ with $\chi(\mu) > 0$ one captures a hyperbolic X by Pesin-Katok method.)

For a continuous potential $\phi : X \to \mathbb{R}$, consider

$$P_{\mathrm{sep}}(f, \phi) := \lim_{\varepsilon \to 0} \limsup_{n\to\infty} \frac{1}{n} \log \big(\sup_Y \sum_{y \in Y} \exp S_n\phi(y) \big), \tag{6-2}$$

where the supremum is taken over all (n, ε)-separated sets $Y \subset X$, that is such Y that for every distinct $y_1, y_2 \in Y$, $\rho_n(y_1, y_2) \geq \varepsilon$, where ρ_n is the metric defined by $\rho_n(x, y) = \max\{\rho(f^j(x), f^j(y)) : j = 0, \dots, n\}$.

For $\phi = -t \log |f'|$ for positive t, in presence of critical points for f, P_{sep} is always equal to ∞ by putting a point of a separated set at a critical point. So we replace it by the tree pressure. One can however use infimum over (n, ε)-spanning sets, thus defining $\underline{P_{\mathrm{spanning}}(f, \phi)}$. This is a valuable notion, often coinciding with other pressures. See Przytycki [ibid.] for an outline of a respective theory. Let me mention only that this is equal to $P(f, -t \log |f'|)$ for $t > 0$ in the complex case if

Definition 6.3. f is *weakly backward Lyapunov stable* which means that for every $\delta > 0$ and $\varepsilon > 0$ for all n large enough and every disc $B = B(x, \exp -\delta n)$ centered at $x \in K$, for every $0 \leq j \leq n$ and every component V of $f^{-j}(B)$ intersecting K, it holds that $\operatorname{diam} V \leq \varepsilon$.

This holds for all rational maps with at most one critical point whose forward trajectory is in $J(f)$ or is attracted to $J(f)$, due to Theorem 5.6.

Question. Does backward weak Lyapunov stability hold for all rational maps?

Finally, *periodic pressure* P_{Per} is defined as P_{tree} with $x \in \mathrm{Per}_n$ (periodic of period n) rather than f^n-preimages of z. In Przytycki, Rivera-Letelier, and Smirnov [2004], this was proved for rational f (see also Binder, N. Makarov, and Smirnov [2003] for a class of polynomials) on $K = J(f)$ that $P_{\mathrm{Per}}(t) = P(t)$ provided

Hypothesis H. For every $\delta > 0$ and all n large enough, if for a set $\mathcal{P} \subset \mathrm{Per}_n$ for all $p, q \in P$ and all $i : 0 \leq i < n$ $\operatorname{dist}(f^i(p), f^i(q)) < \exp -\delta n$, then $\#\mathcal{P} \leq \exp \delta n$.

Question. Does this condition always hold? In particular, can large bunches of periodic orbits exist with orbits exponentially close to a Cremer fixed point?

7 Geometric coding trees, limit sets, Gibbs meets Hausdorff

The notion of geometric coding tree, g.c.t., already appeared in the work Jakobson [1978], where in the expanding case the finite-to-one property of the resulting coding was proved.

It was used later in Przytycki [1985, 1986] and in a full strength in Przytycki, Urbański, and Zdunik [1989, 1991] and papers following them. Similar graphs have since been constructed to analyse the topological aspects of non-invertible dynamics, see for instance Nekrashevych [2011] and Haïssinsky and Pilgrim [2009].

Definition 7.1. Let U be an open connected subset of the Riemann sphere $\overline{\mathbb{C}}$. Consider a holomorphic mapping $f : U \to \overline{\mathbb{C}}$ such that $f(U) \supset U$ and $f : U \to f(U)$ is a proper map. Suppose that $\mathrm{Crit}(f)$ is finite. Consider an arbitrary $z \in f(U)$. Let $z^1, z^2, \dots, z^d$ be some of the f-preimages of z in U with $d \geq 2$. Consider smooth curves $\gamma^j : [0,1] \to f(U)$, $j = 1, \dots, d$, joining z to z^j respectively (i.e. $\gamma^j(0) = z, \gamma^j(1) = z^j$), intersections allowed, such that $\gamma^j \cap f^n(\mathrm{Crit}(f)) = \varnothing$ for every j and $n > 0$.

For every sequence $\alpha = (\alpha_n)_{n=0}^{\infty} \in \Sigma^d$ (shift space with left shift map ς defined in Section 1) define $\gamma_0(\alpha) := \gamma^{\alpha_0}$. Suppose that for some $n \geq 0$, for every $0 \leq m \leq n$, and all $\alpha \in \Sigma^d$, curves $\gamma_m(\alpha) : [0.1] \to U$ are already defined. Suppose that for $1 \leq m \leq n$ we have $f \circ \gamma_m(\alpha) = \gamma_{m-1}(\varsigma(\alpha))$, and $\gamma_m(\alpha)(0) = \gamma_{m-1}(\alpha)(1)$. Define the curves $\gamma_{n+1}(\alpha)$ so that the previous equalities hold by taking respective f-preimages of curves γ_n. For every $\alpha \in \Sigma^d$ and $n \geq 0$ denote $z_n(\alpha) := \gamma_n(\alpha)(1)$.

The graph $\mathcal{T} = \mathcal{T}(z, \gamma^1, \dots, \gamma^d)$ with the vertices z and $z_n(\alpha)$ and edges $\gamma_n(\alpha)$ is called a *geometric coding tree* with the root at z. For every $\alpha \in \Sigma^d$ the subgraph composed of $z, z_n(\alpha)$ and $\gamma_n(\alpha)$ for all $n \geq 0$ is called an *infinite geometric branch* and denoted by $b(\alpha)$. It is called *convergent* if the sequence $\gamma_n(\alpha)$ is convergent to a point in $\mathrm{cl}\, U$. We define the *coding map* $z_\infty : \mathcal{D}(z_\infty) \to \mathrm{cl}\, U$ by $z_\infty(\alpha) := \lim_{n\to\infty} z_n(\alpha)$ on the domain $\mathcal{D} = \mathcal{D}(z_\infty)$ of all such α's for which $b(\alpha)$ is convergent.

Denote $\Lambda := z_\infty(\mathcal{D}(z_\infty))$. If the map f extends holomorphically to a neighbourhood of its closure $\mathrm{cl}\,\Lambda$ in $\overline{\mathbb{C}}$, then Λ is called a *quasi-repeller*, see Przytycki, Urbański, and Zdunik [1989].

A set formally larger than $\mathrm{cl}\,\Lambda$ is of interest, namely $\widehat{\Lambda}$ being the set of all accumulation points of $\{z_n(\alpha) : \alpha \in \Sigma^d\}$ as $n \to \infty$. If our g.c.t. is in Ω being an RB-domain, see Section 8, or f is just $R \circ g \circ R^{-1}$ defined only on Ω, see Remarks below, then it is easy to see that $\mathrm{cl}\,\Lambda = \widehat{\Lambda}$. I do not know how general this equality is.

Remarks. Given a Riemann map $R : \mathbb{D} \to \Omega$ to a connected simply connected domain $\Omega \subset \mathbb{C}$, (i.e. holomorphic bijection) we can consider a branched covering map, say $g(z) = z^d$ on $\mathbb{D}$, and $f = R \circ g \circ R^{-1}$. Then, chosen $z \in \Omega$ and γ^j joining it with its preimages in Ω (close to $\mathrm{Fr}\,\Omega$) we can consider the respective tree $\mathcal{T}$. Then instead of considering R and its radial limit $\overline{R}$, we can consider the limit (along branches) $z_\infty : \Sigma^d \to \mathrm{Fr}\,\Omega$. This provides a structure of symbolic dynamics useful to verify stochastic laws.

This is especially useful if considered measures come from $\partial\mathbb{D}$ via $\overline{R}$, rather than being some equilibrium states for potentials living directly on $\mathrm{Fr}\,\Omega$. This is the case of harmonic measure ω which is the $\overline{R}_*$-image of a length measure l. We can consider the lift of l to Σ^d via coding by the tree $\mathcal{T}' = R^{-1}(\mathcal{T})$ and next its projection by $(z_\infty)_*$ to $\mathrm{Fr}\,\Omega$.

Our g.c.t.'s are always available in presence of adequate holomorphic f, even in the absence of Ω, i.e. in the absence of a Riemann map. The tree with the coding it induces yields a discrete generalization/replacement of a Riemann map.

It was proved in Przytycki and Skrzypczak [1991] that $\mathcal{D}$ is the whole Σ^d except a "thin" set. In particular, for a Gibbs measure ν for a Hölder potential, $z_\infty(\alpha)$ exists for ν-a.e. α, hence the push forward measure $(z_\infty)_*(\nu)$ makes sense. Moreover, our codings ζ_∞ are always "thin"-to-one. This is a discrete generalization of Beurling's Theorem concerning the boundary behaviour of Riemann maps. "Thin" means of zero logarithmic capacity type, depending on the properties of the tree (the speed of the accumulation of γ^j by critical trajectories; the speed does not matter if we replace "thin" by zero Hausdorff dimension). In particular this coding preserves the entropies.

For appropriate $\nu \in \mathcal{M}(\varsigma, \Sigma^d)$ and $\psi : \Sigma^d \to \mathbb{R}$ with $\int \psi \, d\nu = 0$, consider the *asymptotic variance* (of course one can consider spaces more general than Σ^d)

$$\sigma^2 = \sigma_\nu^2(\psi) := \lim_{n\to\infty} \frac{1}{n} \int (S_n\psi)^2 \, d\nu. \tag{7-1}$$

Theorem 7.2. *Let Λ be a quasi-repeller for a geometric coding tree for a holomorphic map $f : U \to \overline{\mathbb{C}}$. Let ν be a ς-invariant Gibbs measure on Σ^d for a Hölder continuous real-valued function ϕ on Σ^d. Assume $P(\varsigma, \phi) = 0$. Consider $\mu := (z_\infty)_*(\nu)$.*

Then, for $\psi := -\mathrm{HD}(\mu)(\log|f'|\circ z_\infty)) - \phi$, we have $\int \psi \, d\nu = 0$.

If the asymptotic variance $\sigma^2 = \sigma_\nu^2(\psi)$ is positive, then there exists a compact f-invariant hyperbolic repeller X being a subset of Λ such that $\mathrm{HD}(X) > \mathrm{HD}(\mu)$. In consequence $\mathrm{HD}_{\mathrm{hyp}}(\Lambda) > \mathrm{HD}(\mu)$ (defined after (5-2)).

If $\sigma^2 = 0$ then ψ is cohomologous to 0. Then for each $x, y \in \mathrm{cl}\,\Lambda$ not postcritical, if $z = f^n(x) = f^m(y)$ for some positive integers n, m, the orders of criticality of f^n at x and f^m at y coincide. In particular all critical points in $\mathrm{cl}\,\Lambda$ are pre-periodic.

The latter condition happens only in special situations, see e.g. Theorem 7.3 below. See Szostakiewicz, Urbański, and Zdunik [2015] for more details; ϕ lives there directly on $J(f)$, but it does not make substantial difference. See also Section 10.

Given a mapping $f : X \to X$, given two functions $u, v : X \to \mathbb{R}$ we call u *cohomologous* to v in class $\mathcal{C}$ if there exists $h : X \to \mathbb{R}$ belonging to $\mathcal{C}$ such that $u - v = h \circ f - h$. An important Przytycki, Urbański, and Zdunik [1989, Lemma 1] says that $\sigma^2 = 0$ above implies ψ cohomologous to 0 in $L^2(\mu)$ and often in a smaller class depending on ψ (Livšic type rigidity).

Notice that $\int \psi\, d\nu = -\mathrm{HD}(\mu)\chi(\mu) - \int \phi\, d\nu = -h_\mu(f) - \int \phi\, d\nu = -h_\nu(\varsigma) - \int \phi\, d\nu = P(\varsigma, \phi) = 0$. Now, to prove Theorem 7.2 note $2\chi(\mu) \geq h_\mu(f) = h_\nu(\varsigma) > 0$, see Przytycki and Urbański [2010, Ruelle's inequality] (used also to 3)⇒5) in Section 3) and Przytycki [1985]. So considering the natural extension of $(\Sigma^d, \nu, \varsigma)$ (here two-sided shift space) and Katok-Pesin theory, we find hyperbolic X with $\mathrm{HD}(X) \geq \mathrm{HD}(\mu) - \varepsilon$ for an arbitrary $\varepsilon > 0$. Compare comments on shadowing in Section 6.

- The positive σ^2 yields by Central Limit Theorem large fluctuations of the sums $\sum_{j=0}^{n-1} \psi \circ \varsigma^j$ from $n \int \psi\, d\nu$ (here 0), allowing to find X with $\mathrm{HD}(X) > \mathrm{HD}(\mu)$.

A special care is needed to get $X \subset \Lambda$, see Przytycki [2005] (originated in Przytycki and Zdunik [1994]).

The above fluctuations were used by A. Zdunik to prove for constant ϕ

Theorem 7.3 (Zdunik [1990]). *Let $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ be a rational mapping of degree $d \geq 2$. If $\sigma^2 > 0$, then for $\mu_{\max}(f)$ the measure of maximal entropy (equal* $\log d$*),* $\mathrm{HD}(J(f)) > \mathrm{HD}(\mu_{\max}(f))$. *Otherwise, f is postcritically finite with a parabolic orbifold, Milnor [2006].*

She proved in fact the existence of a hyperbolic $X \subset J(f)$ satisfying $\mathrm{HD}(X) > \mathrm{HD}(\mu_{\max}(f))$, hence $\mathrm{HD}_{\mathrm{hyp}}(J(f)) > \mathrm{HD}(\mu_{\max}(f))$.

- In the $\sigma^2 = 0$ case, $v : J(f) \to \mathbb{R}$ satisfying the cohomology equation $\log |f'| = v \circ f - v + \mathrm{Const}$ on $J(f)$ extends to a harmonic function beyond $J(f)$ (Livšic rigidity) giving this equality on the union of real analytic curves containing $J(f)$ (called *real case*) or to $\overline{\mathbb{C}}$. In Theorem 7.2 on Λ and for the extension beyond, in Theorem 7.3, the "orders" of growth of $-\log|(f^n)'|$ at x and of $-\log|(f^m)'|$ at y must by cohomology equation be equal to the "order" of growth of v at z, so they must coincide (a phenomenon "conjugated" to the presence of an invariant line field). This implies parabolic orbifold for Theorem 7.3.

Theorem 7.3 applied to a polynomial f with connected Julia set, by $\mathrm{HD}(\mu_{\max}(f)) = 1$ Manning [1984], implies the following Zdunik's celebrated result:

Theorem 7.4 (Zdunik [1990]). *For every polynomial f of degree at least 2, with connected Julia set, either $J(f)$ is a circle or an interval or else it is fractal, namely* $\mathrm{HD}(J(f)) > 1$.

8 Boundaries, radial growth, harmonic vs Hausdorff

For polynomials with connected Julia set the measure $\mu_{\max}(f)$ coincides with harmonic measure ω (viewed from ∞). This led to another proof of Theorem 7.4, especially the $\sigma^2 = 0$ part, see Zdunik [1991], in the language of boundary behaviour of Riemann map and harmonic measure (compare also model Theorem 2.1).

Theorem 7.4 has been strengthened from this point of view in Przytycki [2006], preceded by Przytycki and Zdunik [1994], as follows.

Theorem 8.1. *Let $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ be a rational map of degree at least 2 and Ω be a simply connected immediate basin of attraction to an attracting periodic orbit (that is a connected component of the set attracted to the orbit, intersecting it). Then, provided f is not a finite Blaschke product in some holomorphic coordinates, or a two-to-one holomorphic factor of a Blaschke product,* $\mathrm{HD}_{\mathrm{hyp}}(\mathrm{Fr}\,\Omega) > 1$.

The novelty was to show how to "capture" a large hyperbolic X in $\mathrm{Fr}\,\Omega$ in the case it was not the whole $J(f)$.

In fact the following "local" version of this theorem was proved in Przytycki [2006]

Theorem 8.2. *Assume that f is defined and holomorphic on a neighbourhood W of* $\mathrm{Fr}\,\Omega$, *where Ω is a connected simply connected domain in $\overline{\mathbb{C}}$ whose boundary has at least 2 points. We assume that $f(W \cap \Omega) \subset \Omega$, $f(\mathrm{Fr}\,\Omega) \subset \mathrm{Fr}\,\Omega$ and $\mathrm{Fr}\,\Omega$ repels to the side of Ω, that is $\bigcap_{n=0}^{\infty} f^{-n}(W \cap \mathrm{cl}\,\Omega) = \mathrm{Fr}\,\Omega$. Then either $\mathrm{HD}_{\mathrm{hyp}}(\mathrm{Fr}(\Omega)) > 1$ or $\mathrm{Fr}\,\Omega$ is a real-analytic Jordan curve or arc.*

Ω with f as above has been called an *RB-domain* (repelling boundary), introduced in Przytycki [1986] and Przytycki, Urbański, and Zdunik [1989]. Theorem 8.2 (at least the $\sigma^2 > 0$ part) follows directly from Theorem 7.2. Let $R : \mathbb{D} \to \Omega$ be a Riemann map and $g : W' \to \mathbb{D}$ be defined by $g := R^{-1} \circ f \circ R$ on $W' = R^{-1}(W \cap \Omega)$. We consider a g.c.t. $\mathcal{T} = \mathcal{T}(z, \gamma^1, \dots, \gamma^d)$ with z and γ^j in $W \cap \Omega$, sufficiently close to $\mathrm{Fr}\,\Omega$ that the definition makes sense, and with $d = \deg f|_{W \cap \Omega}$, (the situation is the same as in Remarks in Section 7 above, but the order of defining f and g is different). Consider the g.c.t. $\mathcal{T}' = R^{-1}(\mathcal{T})$. The function g extends holomorphically beyond the circle $\partial\mathbb{D}$ and it is expanding. Hence $\phi : \Sigma^d \to \mathbb{R}$ defined by $\phi(\alpha) = -\log|g'| \circ (R^{-1}(z))_\infty(\alpha)$ for the tree $\mathcal{T}'$ is Hölder continuous. Let $\nu = \nu_\phi$.

Note that here $P(\phi) = 0$, e.g. since by expanding property of g on $\partial\mathbb{D}$ there exists $\widehat{l} \in \mathcal{M}(g)$, equivalent to length measure l (a.c.i.m.). Then ν is the lift of $\widehat{l}$ to Σ^d with the use of $\mathcal{T}'$. Note that our $\mu = z_\infty(\nu)$ is equal to $\widehat{\omega} = \overline{R}_*(\widehat{l})$ which is f-invariant, equivalent to harmonic measures ω on $\mathrm{Fr}\,\Omega$ viewed from Ω.

Note that $\mathrm{HD}(\widehat{\omega}) = 1$ due to Mañé's equality, (5-3), $h_{\widehat{\omega}}(f) = h_{\widehat{l}}(g)$, see Przytycki [1985, 1986], and the equality of Lyapunov exponents $\int \log|f'|\, d\widehat{\omega} = \int \log|g'|\, d\widehat{l} > 0$. The latter equality holds due to the equality for almost every $\zeta \in \partial\mathbb{D}$:

$$\lim_{r \to 1} \frac{\log|(f^n)'(R(r\zeta))| - \log|(g^n)'(r\zeta))}{\log(1-r)} = \lim_{r \to 1} \frac{-\log|R'(r\zeta)|}{\log(1-r)} = 0. \tag{8-1}$$

The first equality is proved using $f \circ R = R \circ g$ in $\mathbb{D}$, first applying R close to $\partial\mathbb{D}$, next by iterating f applying R^{-1} well inside Ω, finally iterating g back. The latter equality relies on the harmonicity of $\log|R'|$ allowing to replace its integral on circles by its value at the origin. For details see Przytycki [1986]. Remind however that in fact $\mathrm{HD}(\omega) = 1$ holds in general, see N. G. Makarov [1985].

The sketch of Proof of Theorem 8.2 for $\sigma^2 > 0$ is over. That $\sigma^2 = 0$ implies the analyticity of $\mathrm{Fr}\,\Omega$ was already commented at the beginning of this Section.

9 Law of Iterated Logarithm refined versions

Applying Law of Iterated Logarithm (abbr. LIL) to $\psi : \Sigma^d \to \mathbb{R}$ the fluctuations of $S_n\psi$ from 0 which follow lead to, see Przytycki, Urbański, and Zdunik [1989] and Przytycki and Urbański [2010],

Theorem 9.1. *In the setting of Theorem 7.2 if $\sigma^2 = \sigma_\nu^2(\psi) > 0$, for $c(\mu) := \sqrt{2\sigma^2/\chi(\mu)}$, $\kappa := \mathrm{HD}(\mu)$ and $\alpha_c(r) := r^\kappa \exp(c\sqrt{\log 1/r \log\log\log 1/r})$*

1) $\mu \perp H_{\alpha_c}$, that is singular with respect to the refined Hausdorff measure, Przytycki and Urbański [ibid., Section 8.2] for the gauge function α_c), for all $0 < c < c(\mu)$;

2) $\mu \ll H_{\alpha_c}$, that is absolutely continuous, for all $c > c(\mu)$.

Indeed, substituting in LIL $n \sim (\log 1/r_n)/\chi(\mu)$ for $r_n = |(f^n)'(z)|^{-n}$, we get for μ-a.e. z

$$\limsup_{n\to\infty} \frac{\mu(B(z,r_n))}{\alpha_c(r_n)} = \infty \text{ for } 0 < c < c(\mu) \quad \text{and} \quad \cdots = 0 \text{ for } c > c(\mu). \tag{9-1}$$

This is called the Refined Volume Lemma, Przytycki, Urbański, and Zdunik [1989, Section 4] and, the harder case: $c > c(\mu)$, Przytycki, Urbański, and Zdunik [1991, Section 5].

We can apply the assertion of Theorem 9.1 for $\mu = \widehat{\omega} \in \mathcal{M}(f, \mathrm{Fr}\,\Omega)$ equivalent to a harmonic measure ω as Section 8.

This yields refined information about the radial growth of the derivative of Riemann maps, following the proof of (8-1):

Theorem 9.2. *Let Ω be a simply connected RB-domain in $\overline{\mathbb{C}}$ with non-analytic boundary and $R : \mathbb{D} \to \Omega$ be a Riemann map. Then there exists $c(\Omega) > 0$ such that for Lebesgue a.e. $\zeta \in \partial\mathbb{D}$*

$$G^+(\zeta) := \limsup_{r\to 1} \frac{\log|R'(r\zeta)|}{\sqrt{\log(1/1-r)\log\log\log(1/1-r)}} = c(\Omega). \tag{9-2}$$

Similarly $G^-(\zeta) := \liminf \cdots = -c(\Omega)$. Finally $c(\Omega) = c(\widehat{\omega})$ in Theorem 9.1.

In fact Theorem 9.1 for $\mu = \widehat{\omega}$ and Theorem 9.2 hold for every connected, simply connected open $\Omega \subset \mathbb{C}$, together with $c(\Omega) = c(\widehat{\omega})$. No dynamics is needed. Of course one should add to both definitions ess sup over $\zeta \in \partial\mathbb{D}$ and over $z \in \operatorname{Fr}\Omega$ (for $c(z) = c(\omega)$ calculated from (9-1), see Przytycki and Urbański [2010, Th. 8.6.1]) respectively, since in the absence of ergodicity these functions need not be constant. See Garnett and Marshall [2005, Th. VIII.2.1 (a)] and references to Makarov's breakthrough papers therein, in particular N. G. Makarov [1985].

There is a universal Makarov's upper bound $C_{\mathrm{M}} < \infty$ for all $c(\Omega), c(\widehat{\omega})$'s in (9-2). The best upper estimate I found in literature is $C_{\mathrm{M}} \leq 1.2326$, Hedenmalm and Kayumov [2007]. I proved in Przytycki [1989] a much weaker estimate, using a natural method of representing $\log|R'|$ by a series of weakly dependent random variables leading to a martingale on $\partial\mathbb{D}$, thus satisfying LIL. Unfortunately consecutive approximations resulted with looses in the final estimate.

For a holomorphic expanding repeller $f : X \to X$ and a Hölder continuous potential $\phi : X \to X$, the asymptotic variance for the equilibrium state $\mu = \mu_{t_0\phi}$ for every $t_0 \in \mathbb{R}$ satisfies Ruelle's formula (see Przytycki and Urbański [2010]):

$$\sigma_\mu^2(\phi - \int \phi \, d\mu) = \frac{d^2 P(t\phi)}{dt^2}\bigg|_{t=t_0}. \tag{9-3}$$

Question. Does (9-3) hold for all rational maps and hyperbolic potentials on Julia sets? For all simply connected RB-domains, $f : \operatorname{Fr}\Omega \to \operatorname{Fr}\Omega$ and $\mu = \widehat{\omega}$?

For a simply connected RB-domain Ω for f and for $\phi = -\log|f'|$, if $g(z) = z^d$ (e.g. Ω being the basin of ∞ for a polynomial f), one considers the *integral means spectrum* depending only on Ω,

$$\beta_\Omega(t) := \limsup_{r \to 1} \frac{1}{|\log(1-r)|} \log \int_{\zeta \in \partial\mathbb{D}} |R'(r\zeta)|^t \, |d\zeta| \tag{9-4}$$

which happens to satisfy $\beta_\Omega(t) = t - 1 + \frac{P(t\phi)}{\log d}$, see e.g. Przytycki and Urbański [ibid., Eq. (9.6.2.)].

For $t_0 = 0$ we have $\mu = \widehat{\omega}$ and the left hand side of (9-3) can be written as $(\frac{1}{2}\log d)\sigma^2(\log R')$, see (7-1) and (8-1), where

$$\sigma^2(\log R') := \limsup_{r \to 1} \frac{\int_{\partial\mathbb{D}} |\log R'(t\zeta)|^2 \, |d\zeta|}{-2\pi \log(1-r)|}.$$

So (9-3) changes to $\sigma^2(\log R') = 2\frac{d^2\beta_\Omega(t))}{dt^2}|_{t=0}$, compare Ivrii [2016]. It has an analytic, non-dynamical, meaning. It is also related to the Weil-Petersson metric, see McMullen [2008].

10 Accessibility

Let us recall the following theorem from Przytycki [1994].

Theorem 10.1. *Let Λ be a quasi-repeller for a geometric coding tree for a holomorphic map $f : U \to \overline{\mathbb{C}}$. Suppose that*

$$\operatorname{diam}(\gamma_n(\alpha)) \to 0, \text{ as } n \to \infty \tag{10-1}$$

uniformly with respect to $\alpha \in \Sigma^d$. Then every good $q \in \widehat{\Lambda}$ (defined in Section 7) is a limit of a convergent branch $b(\alpha)$. So $q \in \Lambda$. In particular, this holds for every q with $\underline{\chi}(q) > 0$ and the local backward inviariance (explained below).

For the definition of "good", see Przytycki [ibid., Definition 2.5]. It roughly says that there are many integers n (positive lower density) for which f^n properly map small domains $D_{n,0}$ in U close to q onto large $D_n \subset U$, giving "telescopes" Tel_k with "traces" $D_{n_k,0} \subset D_{n_{k-1},0} \subset \cdots \subset D_{n_1,0} \subset D_0$; for each k the choices may be different. A part of this condition that $D_{n,0} \subset U$ can be called a "local backward invariance" of U along the forward trajectory of q.

When U is an immediate basin of attraction of an attracting fixed point for a rational map f or just an RB-domain then this theorem asserts that q is an endpoint of a continuous curve in U. This is a generalization of the Douady-Eremenko-Levin-Petersen theorem where q is a repelling periodic point and the domain is completely invariant, e.g. basin of attraction to ∞ for f a polynomial.

Due to this theorem we can prove that invariant measures of positive Lyapunov exponents lift to Σ^d. More precisely, the following holds:

Corollary 10.2. *Every non-atomic hyperbolic probability measure μ (i.e. $\chi(\mu) > 0$), on $\widehat{\Lambda}$, is the $(z_\infty)_*$ image of a probability ς-invariant measure ν on Σ^d, assumed (10-1), $\mathcal{T}$ has no self-intersections and else μ-a.e. local backward invariance of U,. In particular, ν exists for every RB-domain which is completely (i.e. backward) invariant.*

Proof. (the lifting part missing in Przytycki [ibid.] and Przytycki [2006]). By Theorem 10.1 μ is supported on Λ i.e. on $z_\infty(\mathcal{D}(z_\infty))$. The lift of μ to μ' on the pre-image $\mathcal{B}'$ under z_∞ of the Borel σ-algebra of subsets of Λ can be extended to a ς-invariant ν on $\mathcal{B}$ the Borel σ-algebra of the subsets of Σ^d by using the fact that the set of at least triple points (limit points of at least three infinite branches of $\mathcal{T}$) is countable, hence $z_\infty^{-1}(x)$ of μ-a.e x contains at most 2 points. More precisely, let A_1 be the set of points having one z_∞-preimage, A_2 two preimages. They are both f-invariant (except measure 0), so are their z_∞-preimages A_1' and A_2' under ς. We extend μ' by distributing conditional measures on the two points preimages of points in A_2 half-half and Dirac on one point preimages.

∎

This allows to conclude Theorem 9.1 (a part relying on CLT) and Theorem 7.2 for equilibrium states for rational maps and Hölder potentials on $J(f)$ by lifting μ_ϕ to Σ^d as in Przytycki [ibid.]. However, this seems useless since the proof of CLT in Przytycki [ibid.] is done directly on $J(f)$ (seemingly also for LIL, for which one should however refer to the proofs in Przytycki, Urbański, and Zdunik [1989]) and there are direct proofs of LIL in Li and Rivera-Letelier [2014b] and Szostakiewicz, Urbański, and Zdunik [2015].

Acknowledgments. Thanks to H. Hedenmalm, O. Ivrii, J. Rivera-Letelier, M. Sabok, M. Urbański, and A. Zdunik for comments and corrections.

References

Krzysztof Barański, Bogusława Karpińska, and Anna Zdunik (2009). "Hyperbolic dimension of Julia sets of meromorphic maps with logarithmic tracts". *Int. Math. Res. Not. IMRN* 4, pp. 615–624. MR: 2480096 (cit. on p. 2109).

– (2012). "Bowen's formula for meromorphic functions". *Ergodic Theory Dynam. Systems* 32.4, pp. 1165–1189. MR: 2955309 (cit. on p. 2109).

I. Binder, N. Makarov, and S. Smirnov (2003). "Harmonic measure and polynomial Julia sets". *Duke Math. J.* 117.2, pp. 343–365. MR: 1971297 (cit. on p. 2119).

Rufus Bowen (1975). *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin-New York, pp. i+108. MR: 0442989 (cit. on p. 2106).

H. Bruin, J. Rivera-Letelier, W. Shen, and S. van Strien (2008). "Large derivatives, backward contraction and invariant densities for interval maps". *Invent. Math.* 172.3, pp. 509–533. MR: 2393079 (cit. on p. 2116).

Henk Bruin and Mike Todd (2008). "Equilibrium states for interval maps: potentials with $\sup\phi - \inf\phi < h_{\mathrm{top}}(f)$". *Comm. Math. Phys.* 283.3, pp. 579–611. MR: 2434739 (cit. on p. 2111).

– (2009). "Equilibrium states for interval maps: the potential $-t\log|Df|$". *Ann. Sci. Éc. Norm. Supér. (4)* 42.4, pp. 559–600. MR: 2568876 (cit. on p. 2115).

Daniel Coronel and Juan Rivera-Letelier (2013). "Low-temperature phase transitions in the quadratic family". *Adv. Math.* 248, pp. 453–494. arXiv: 1205.1833. MR: 3107518 (cit. on p. 2116).

– (2015). "High-order phase transitions in the quadratic family". *J. Eur. Math. Soc. (JEMS)* 17.11, pp. 2725–2761. arXiv: 1305.4971. MR: 3420520 (cit. on p. 2116).

Manfred Denker, Feliks Przytycki, and Mariusz Urbański (1996). "On the transfer operator for rational functions on the Riemann sphere". *Ergodic Theory Dynam. Systems* 16.2, pp. 255–266. MR: 1389624 (cit. on pp. 2111, 2114).

Manfred Denker and Mariusz Urbański (1991). "Ergodic theory of equilibrium states for rational maps". *Nonlinearity* 4.1, pp. 103–134. MR: 1092887 (cit. on pp. 2110, 2111).

Neil Dobbs and Mike Todd (Dec. 2015). "Free energy jumps up". arXiv: 1512.09245 (cit. on pp. 2115, 2117).

John B. Garnett and Donald E. Marshall (2005). *Harmonic measure*. Vol. 2. New Mathematical Monographs. Cambridge University Press, Cambridge, pp. xvi+571. MR: 2150803 (cit. on p. 2125).

Katrin Gelfert, Feliks Przytycki, and Michał Rams (2010). "On the Lyapunov spectrum for rational maps". *Math. Ann.* 348.4, pp. 965–1004. MR: 2721648 (cit. on p. 2118).

– (2016). "Lyapunov spectrum for multimodal maps". *Ergodic Theory Dynam. Systems* 36.5, pp. 1441–1493. MR: 3519419 (cit. on pp. 2110, 2115, 2118).

Jacek Graczyk and Stanislav Smirnov (2009). "Non-uniform hyperbolicity in complex dynamics". *Invent. Math.* 175.2, pp. 335–415. MR: 2470110 (cit. on p. 2116).

Jacek Graczyk and Stas Smirnov (1998). "Collet, Eckmann and Hölder". *Invent. Math.* 133.1, pp. 69–96. MR: 1626469 (cit. on p. 2113).

Peter Haïssinsky and Kevin M. Pilgrim (2009). "Coarse expanding conformal dynamics". *Astérisque* 325, viii+139 pp. (2010). MR: 2662902 (cit. on p. 2120).

Håkan Hedenmalm and Ilgiz Kayumov (2007). "On the Makarov law of the iterated logarithm". *Proc. Amer. Math. Soc.* 135.7, pp. 2235–2248. MR: 2299501 (cit. on p. 2125).

Juha Heinonen and Pekka Koskela (1995). "Definitions of quasiconformality". *Invent. Math.* 120.1, pp. 61–79. MR: 1323982 (cit. on p. 2114).

Franz Hofbauer and Gerhard Keller (1993). "Equilibrium states and Hausdorff measures for interval maps". *Math. Nachr.* 164, pp. 239–257. MR: 1251466 (cit. on p. 2110).

Franz Hofbauer and Mariusz Urbański (1994). "Fractal properties of invariant subsets for piecewise monotonic maps on the interval". *Trans. Amer. Math. Soc.* 343.2, pp. 659–673. MR: 1232188 (cit. on p. 2110).

Irene Inoquio-Renteria and Juan Rivera-Letelier (2012). "A characterization of hyperbolic potentials of rational maps". *Bull. Braz. Math. Soc. (N.S.)* 43.1, pp. 99–127. MR: 2909925 (cit. on pp. 2110, 2111, 2113, 2116).

Godofredo Iommi and Mike Todd (2010). "Natural equilibrium states for multimodal maps". *Comm. Math. Phys.* 300.1, pp. 65–94. MR: 2725183 (cit. on p. 2115).

Oleg Ivrii (Apr. 2016). "On Makarov's principle in conformal mapping". To appear in Int. Math. Res. Not. arXiv: 1604.05619 (cit. on p. 2125).

M. V. Jakobson (1978). "Markov partitions for rational endomorphisms of the Riemann sphere. Mnogokomponientnyje slučajnyje sistemy". *Moscow: Nauka (In Russian)* 440, pp. 303–319 (cit. on p. 2119).

Genadi Levin, Feliks Przytycki, and Weixiao Shen (2016). "The Lyapunov exponent of holomorphic maps". *Invent. Math.* 205.2, pp. 363–382. MR: 3529117 (cit. on p. 2118).

Huaibin Li and Juan Rivera-Letelier (2014a). "Equilibrium states of interval maps for hyperbolic potentials". *Nonlinearity* 27.8, pp. 1779–1804. MR: 3246155 (cit. on p. 2111).
– (2014b). "Equilibrium states of weakly hyperbolic one-dimensional maps for Hölder potentials". *Comm. Math. Phys.* 328.1, pp. 397–419. MR: 3196990 (cit. on pp. 2111, 2127).
Mikhail Lyubich (2014). "Analytic low-dimensional dynamics: from dimension one to two". In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. 1*. Kyung Moon Sa, Seoul, pp. 443–474. MR: 3728479 (cit. on p. 2115).
N. G. Makarov (1985). "On the distortion of boundary sets under conformal mappings". *Proc. London Math. Soc. (3)* 51.2, pp. 369–384. MR: 794117 (cit. on pp. 2124, 2125).
N. Makarov and S. Smirnov (2003). "On thermodynamics of rational maps. II. Non-recurrent maps". *J. London Math. Soc. (2)* 67.2, pp. 417–432. MR: 1956144 (cit. on p. 2116).
Anthony Manning (1984). "The dimension of the maximal measure for a polynomial map". *Ann. of Math. (2)* 119.2, pp. 425–430. MR: 740898 (cit. on p. 2122).
R. Daniel Mauldin and Mariusz Urbański (2003). *Graph directed Markov systems*. Vol. 148. Cambridge Tracts in Mathematics. Geometry and dynamics of limit sets. Cambridge University Press, Cambridge, pp. xii+281. MR: 2003772 (cit. on p. 2117).
Curtis T. McMullen (2008). "Thermodynamics, dimension and the Weil-Petersson metric". *Invent. Math.* 173.2, pp. 365–425. MR: 2415311 (cit. on p. 2125).
Welington de Melo and Sebastian van Strien (1993). *One-dimensional dynamics*. Vol. 25. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, pp. xiv+605. MR: 1239171 (cit. on p. 2110).
John Milnor (2006). "On Lattès maps". In: *Dynamics on the Riemann sphere*. Eur. Math. Soc., Zürich, pp. 9–43. MR: 2348953 (cit. on p. 2122).
M. Misiurewicz and W. Szlenk (1980). "Entropy of piecewise monotone mappings". *Studia Math.* 67.1, pp. 45–63. MR: 579440 (cit. on p. 2110).
Volodymyr Nekrashevych (2011). "Iterated monodromy groups". In: *Groups St Andrews 2009 in Bath. Volume 1*. Vol. 387. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, pp. 41–93. MR: 2858850 (cit. on p. 2120).
Tomasz Nowicki and Feliks Przytycki (1998). "Topological invariance of the Collet-Eckmann property for S-unimodal maps". *Fund. Math.* 155.1, pp. 33–43. MR: 1487986 (cit. on p. 2113).
Tomasz Nowicki and Duncan Sands (1998). "Non-uniform hyperbolicity and universal bounds for S-unimodal maps". *Invent. Math.* 132.3, pp. 633–680. MR: 1625708 (cit. on pp. 2113, 2118).
Yakov Pesin (2014). "On the work of Sarig on countable Markov chains and thermodynamic formalism". *J. Mod. Dyn.* 8.1, pp. 1–14. MR: 3296563 (cit. on p. 2117).

Yakov Pesin and Samuel Senti (2008). "Equilibrium measures for maps with inducing schemes". *J. Mod. Dyn.* 2.3, pp. 397–430. MR: 2417478 (cit. on p. 2115).

Feliks Przytycki (1985). "Hausdorff dimension of harmonic measure on the boundary of an attractive basin for a holomorphic map". *Invent. Math.* 80.1, pp. 161–179. MR: 784535 (cit. on pp. 2120, 2122, 2123).

– (1986). "Riemann map and holomorphic dynamics". *Invent. Math.* 85.3, pp. 439–455. MR: 848680 (cit. on pp. 2120, 2123, 2124).

– (1989). "On the law of iterated logarithm for Bloch functions". *Studia Math.* 93.2, pp. 145–154. MR: 1002917 (cit. on p. 2125).

– (1990). "On the Perron-Frobenius-Ruelle operator for rational maps on the Riemann sphere and for Hölder continuous functions". *Bol. Soc. Brasil. Mat. (N.S.)* 20.2, pp. 95–125. MR: 1143178 (cit. on p. 2111).

– (1993). "Lyapunov characteristic exponents are nonnegative". *Proc. Amer. Math. Soc.* 119.1, pp. 309–317. MR: 1186141 (cit. on p. 2114).

– (1994). "Accessibility of typical points for invariant measures of positive Lyapunov exponents for iterations of holomorphic maps". *Fund. Math.* 144.3, pp. 259–278. MR: 1279481 (cit. on p. 2126).

– (1998). "Iterations of holomorphic Collet-Eckmann maps: conformal and invariant measures. Appendix: on non-renormalizable quadratic polynomials". *Trans. Amer. Math. Soc.* 350.2, pp. 717–742. MR: 1407501.

– (1999). "Conical limit set and Poincaré exponent for iterations of rational functions". *Trans. Amer. Math. Soc.* 351.5, pp. 2081–2099. MR: 1615954 (cit. on p. 2118).

– (2000). "Hölder implies Collet-Eckmann". *Astérisque* 261. Géométrie complexe et systèmes dynamiques (Orsay, 1995), pp. xiv, 385–403. MR: 1755448 (cit. on p. 2113).

– (2005). "Expanding repellers in limit sets for iterations of holomorphic functions". *Fund. Math.* 186.1, pp. 85–96. MR: 2163105 (cit. on p. 2122).

– (2006). "On the hyperbolic Hausdorff dimension of the boundary of a basin of attraction for a holomorphic map and of quasirepellers". *Bull. Pol. Acad. Sci. Math.* 54.1, pp. 41–52. MR: 2270793 (cit. on pp. 2123, 2126, 2127).

– (2018). "Geometric pressure in real and complex 1-dimensional dynamics via trees of pre-images and via spanning sets". *Monatsh. Math.* 185.1, pp. 133–158. arXiv: 1602.08573. MR: 3745705 (cit. on pp. 2110, 2118, 2119).

Feliks Przytycki and Juan Rivera-Letelier (2011). "Nice inducing schemes and the thermodynamics of rational maps". *Comm. Math. Phys.* 301.3, pp. 661–707. MR: 2784276 (cit. on pp. 2115, 2116).

– (May 2014). "Geometric pressure for multimodal maps of the interval". To appear in Memoirs of the Amer. Math. Soc. arXiv: 1405.2443 (cit. on pp. 2110, 2113, 2115, 2118).

Feliks Przytycki, Juan Rivera-Letelier, and Stanislav Smirnov (2003). "Equivalence and topological invariance of conditions for non-uniform hyperbolicity in the iteration of rational maps". *Invent. Math.* 151.1, pp. 29–63. MR: 1943741 (cit. on pp. 2113, 2118).

– (2004). "Equality of pressures for rational functions". *Ergodic Theory Dynam. Systems* 24.3, pp. 891–914. MR: 2062924 (cit. on p. 2119).

Feliks Przytycki and Steffen Rohde (1998). "Porosity of Collet-Eckmann Julia sets". *Fund. Math.* 155.2, pp. 189–199. MR: 1606527 (cit. on p. 2113).

– (1999). "Rigidity of holomorphic Collet-Eckmann repellers". *Ark. Mat.* 37.2, pp. 357–371. MR: 1714763 (cit. on p. 2114).

Feliks Przytycki and Jan Skrzypczak (1991). "Convergence and pre-images of limit points for coding trees for iterations of holomorphic maps". *Math. Ann.* 290.3, pp. 425–440. MR: 1116230 (cit. on p. 2121).

Feliks Przytycki and Mariusz Urbański (2010). *Conformal fractals: ergodic theory methods*. Vol. 371. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, pp. x+354. MR: 2656475 (cit. on pp. 2106–2109, 2117, 2118, 2122, 2124, 2125).

Feliks Przytycki, Mariusz Urbański, and Anna Zdunik (1989). "Harmonic, Gibbs and Hausdorff measures on repellers for holomorphic maps. I". *Ann. of Math. (2)* 130.1, pp. 1–40. MR: 1005606 (cit. on pp. 2120, 2121, 2123, 2124, 2127).

– (1991). "Harmonic, Gibbs and Hausdorff measures on repellers for holomorphic maps. II". *Studia Math.* 97.3, pp. 189–225. MR: 1100687 (cit. on pp. 2120, 2124).

Feliks Przytycki and A. Zdunik (1994). "Density of periodic sources in the boundary of a basin of attraction for iteration of holomorphic maps: geometric coding trees technique". *Fund. Math.* 145.1, pp. 65–77. MR: 1295160 (cit. on pp. 2122, 2123).

Juan Rivera-Letelier (Apr. 2012). "Asymptotic expansion of smooth interval maps". arXiv: 1204.3071 (cit. on pp. 2113, 2114).

Juan Rivera-Letelier and Weixiao Shen (2014). "Statistical properties of one-dimensional maps under weak hyperbolicity assumptions". *Ann. Sci. Éc. Norm. Supér. (4)* 47.6, pp. 1027–1083. MR: 3297154 (cit. on p. 2116).

David Ruelle (1978). *Thermodynamic formalism*. Vol. 5. Encyclopedia of Mathematics and its Applications. The mathematical structures of classical equilibrium statistical mechanics, With a foreword by Giovanni Gallavotti and Gian-Carlo Rota. Addison-Wesley Publishing Co., Reading, Mass., pp. xix+183. MR: 511655 (cit. on p. 2106).

Weixiao Shen and Sebastian van Strien (2014). "Recent developments in interval dynamics". In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. III*. Kyung Moon Sa, Seoul, pp. 699–719. MR: 3729048 (cit. on p. 2116).

Ja. G. Sinaĭ (1972). "Gibbs measures in ergodic theory". *Uspehi Mat. Nauk* 27.4(166), pp. 21–64. MR: 0399421 (cit. on p. 2106).

Bernd O. Stratmann and Mariusz Urbanski (2003). "Real analyticity of topological pressure for parabolically semihyperbolic generalized polynomial-like maps". *Indag. Math. (N.S.)* 14.1, pp. 119–134. MR: 2015603 (cit. on p. 2117).

D. P. Sullivan (1982). "Seminar on Conformal and Hyperbolic geometry, Preliminary Publ". *IHES, notes by M. Baker and J. Seade* (cit. on p. 2109).

Michał Szostakiewicz, Mariusz Urbański, and Anna Zdunik (2014). "Stochastics and thermodynamics for equilibrium measures of holomorphic endomorphisms on complex projective spaces". *Monatsh. Math.* 174.1, pp. 141–162. MR: 3190774 (cit. on p. 2111).

– (2015). "Fine inducing and equilibrium measures for rational functions of the Riemann sphere". *Israel J. Math.* 210.1, pp. 399–465. MR: 3430280 (cit. on pp. 2111, 2116, 2121, 2127).

Anna Zdunik (1990). "Parabolic orbifolds and the dimension of the maximal measure for rational maps". *Invent. Math.* 99.3, pp. 627–649. MR: 1032883 (cit. on p. 2122).

– (1991). "Harmonic measure versus Hausdorff measures on repellers for holomorphic maps". *Trans. Amer. Math. Soc.* 326.2, pp. 633–652. MR: 1031980 (cit. on p. 2122).

Received 2017-11-29.

Feliks Przytycki: Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland
feliksp@impan.gov.pl
feliksp@impan.pl

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2131–2154)

QUANTITATIVE ALMOST REDUCIBILITY AND ITS APPLICATIONS

JIANGONG YOU (尤建功)

Abstract

We survey the recent advances of almost reducibility and its applications in the spectral theory of one dimensional quasi-periodic Schrödinger operators.

1 Quasi-periodic operators, cocycles and systems

1.1 One dimensional quasi-periodic Schrödinger operators. One dimensional quasi-periodic discrete time Schrödinger operators are operators defined on $l^2(\mathbb{Z})$ as

$$(H_{V,\omega,\theta}u)_n = u_{n+1} + u_{n-1} + V(\theta + n\omega)u_n, \quad \forall n \in \mathbb{Z}, \tag{1-1}$$

where $\theta \in \mathbb{T}^d := (\mathbb{R}/\mathbb{Z})^d$ is called *phase*, $V : \mathbb{T}^d \to \mathbb{R}$ is called *potential*, rationally independent $\omega \in \mathbb{T}^d$ is called *frequency*(when ω is one dimensional, we will replace it by α to respect the traditional notation in literatures). The simplest but the most important special case is the *almost Mathieu operators*(AMO), i.e., the three-parameter family:

$$(H_{\lambda,\alpha,\theta}u)_n = u_{n+1} + u_{n-1} + 2\lambda \cos 2\pi(\theta + n\alpha)u_n, \quad \forall n \in \mathbb{Z}. \tag{1-2}$$

Due to the rich backgrounds in quantum physics, quasi-periodic Schrödinger operators especially the almost Mathieu operators have been extensively studied Last [2005]. In 1980's, there was an almost periodic flu which already swept the world Simon [1982]. In 2000's, people found that one can use ideas from the dynamical systems (mainly linear cocycles) to study the operators (1-1), and many important progresses have been made since then (Avila [2008, 2015a], Avila and Krikorian [2006], Avila and Jitomirskaya [2009], and Puig [2004]). This survey will focus on how almost reducibility is used to give a systematical study of various delicate spectral properties of (1-1).

This work was partially supported by NNSF of China (11471155) and 973 projects of China (2014CB340701).
MSC2010: primary 37C55; secondary 37J40, 47B36.

It is well known that $H = H_{V,\omega,\theta}$ is a bounded self-adjoint operator, its spectrum $\Sigma_{V,\omega,\theta}$ is a compact perfect subset of $\mathbb{R}$ which is independent of θ if ω is rationally indepentent. In the following, when no danger of confusion, we sometimes simply denote $H_{V,\omega,\theta}$ and $\Sigma_{V,\omega,\theta}$ by H and Σ.

Given an operator $H = H_{V,\omega,\theta}$ and a $\phi \in l^2(\mathbb{Z})$, we define a measure μ^ϕ on Σ such that

$$\langle \phi, f(H)\phi \rangle = \int f(E) d\mu^\phi(E),$$

holds for any $f \in C^0(\Sigma)$, $d\mu = d\mu^{e_0} + d\mu^{e_1}$ is called the *spectral measure* of H. And the *integrated density of states* (IDS) $N : \mathbb{R} \to [0, 1]$ of H is defined as

$$N(E) := \int_{\mathbb{T}} d\mu(-\infty, E]\, d\theta. \tag{1-3}$$

$N(E)$ is always monotone and continuous no matter what $d\mu$ is. Any bounded connected component of $\mathbb{R}\backslash\Sigma$ is called a *spectral gap* of the operator $H_{V,\omega,\theta}$. By Gap-Labelling Theorem (Johnson and Moser [1982]), there is a unique $k \in \mathbb{Z}^d$ such that $N(E) = \langle k, \omega \rangle \mod \mathbb{Z}$ for all E in a gap. In other words, the spectral gaps can be labelled by $k \in \mathbb{Z}^d$. We denote by $G_k(V) = (E_k^-(V), E_k^+(V))$ the gap with labelling k. If $G_k(V)$ is not empty for all k, we say *all gaps of H are open*. When Σ is a Cantor set, we say the operator H has *Cantor spectrum*.

The continuous time quasi-periodic Schrödinger operators $\mathfrak{L} = \mathfrak{L}_{q,\omega,\theta}$ are defined on $L^2(\mathbb{R})$ as

$$(\mathfrak{L}_{q,\omega,\theta}\, y)(t) = -y''(t) + q(\theta + \omega t)y(t) \tag{1-4}$$

where $q : \mathbb{T}^d \to \mathbb{R}$, and $\omega \in \mathbb{T}^d$ is rationally independent. It is known that $\mathfrak{L}_{q,\omega,\theta}$ is selfajoint and unbounded, its spectrum is an unbounded perfect subset of $\mathbb{R}$ independent of θ. All concepts above for the discrete time Schrödinger operators can be defined similarly for the continuous time quasi-periodic Schödinger operators.

The spectrum and spectral measure are two central subjects in spectral theory. For the spectrum, people are mainly interested in the Lebesgue measure of Σ, Cantor spectrum, homogeneity of the spectrum, opening gaps and gap estimates. For the spectral measure, people are interested in the nature of the measure: when it is absolute continuous, when it is singular continuous or pure point; if it is pure point, when it has Anderson localization (pure point with exponential decay eigenfunctions) or dynamical localization. What is the modulus of the continuity of IDS and the spectral measure? If the operator contains parameters, the phase transition is also an important issue.

1.2 Quasi-periodic cocycles and quasi-periodic linear systems. Note that $(u_n)_{n\in\mathbb{Z}}$ is a formal solution of the eigenvalue equation $H_{V,\omega,\theta}u = Eu$ if and only if it satisfies

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = S_E^V(\theta + n\omega)\cdot\begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}, \tag{1-5}$$

where

$$S_E^V(\theta) = \begin{pmatrix} E - V(\theta) & -1 \\ 1 & 0 \end{pmatrix} \in SL(2,\mathbb{R}).$$

The dynamical systems (ω, S_E^V) defined on $\mathbb{T}^d \times \mathbb{R}^2$ by

$$(\theta, v) \mapsto (\theta + \omega, S_E^V(\theta)v) \tag{1-6}$$

are called Schrödinger cocycles.

In general, an analytic quasi-periodic linear cocycle (ω, A) on $\mathbb{T}^d$ with coefficients in Lie group $\mathbf{G}$ (its Lie algebra will be denoted by $\mathbf{g}$) is defined by

$$\begin{aligned} \mathbb{T}^d \times \mathbb{R}^N &\to \mathbb{T}^d \times \mathbb{R}^N \\ (\theta, v) &\mapsto (\theta + \omega, A(\theta)\cdot v). \end{aligned} \tag{1-7}$$

where $A \in C^\omega(\mathbb{T}^d, \mathbf{G})$, $\mathbf{G}$ will be usually taken as $GL(N,\mathbb{R})$, $Sp(2N,\mathbb{R})$. The iterate of the cocycle is defined as

$$\mathcal{A}_n(\theta) := \begin{cases} A(\theta + (n-1)\omega)\cdots A(\theta+\omega)A(x), & n \geq 0 \\ A^{-1}(\theta + n\omega)A^{-1}(\theta + (n+1)\omega)\cdots A^{-1}(\theta - \omega), & n < 0. \end{cases}$$

Let $\lambda_1(\theta), \lambda_2(\theta), \cdots, \lambda_N(\theta)$ be the singular values of $\mathcal{A}_n(\theta)$. By Oseledets theory,

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \ln \lambda_i(\theta)\, d\theta, \quad i = 1, \cdots, N$$

exsit and are same for almost all θ. λ_i's are called *Lyapunov exponents* of (1-7), among them $L(\omega, A) = \lim_{n\to\infty} \frac{1}{n}\int_{\mathbb{T}^d} \ln \|\mathcal{A}_n(\theta)\| d\theta$ is the largest. When ω has been fixed, we simply write $L(A)$ for $L(\omega, A)$. If $A(\theta)$ are in $SL(2,\mathbb{R})$, the two Lyapunov exponents are $\pm L(\omega, A)$. If furthermore $A(\theta) \in C^\omega(\mathbb{T}, \mathrm{SL}(2,\mathbb{R}))$, Avila [2015a] proved that

$$\omega(A) := \lim_{\epsilon\to 0} \frac{L(A_\epsilon) - L(A)}{2\pi\epsilon}, \quad \text{where } A_\epsilon = A(\theta + i\epsilon),$$

exists and moreover it is an integer. The quantity $\omega(A)$, called *acceleration*, plays an important role in Avila's global theory of one frequency analytic quasi-periodic Schrödinger cocycles (Avila [ibid.]).

The cocycle (1-7) is said to be *uniformly hyperbolic* if there exists a continuous splitting $\mathbb{R}^N = E^s(\theta) \oplus E^u(\theta)$ such that for every $n \geq 0$,

$$|\mathfrak{A}_n(\theta)\, v| \leq Ce^{-cn}|v|, \;\; v \in E^s(\theta),$$
$$|\mathfrak{A}_n(\theta)^{-1} v| \leq Ce^{-cn}|v|, \;\; v \in E^u(\theta + n\omega),$$

for some constants $C, c > 0$. Moreover, this splitting is invariant, i.e.,

$$A(\theta)E^s(\theta) = E^s(\theta+\omega), \quad A(\theta)E^u(\theta) = E^u(\theta+\omega), \quad \forall\, \theta \in \mathbb{T}^d.$$

A cocycle is said to be *non-uniformly hyperbolic* if it is not uniformly hyperbolic and all the Lyapunov exponents are not zero.

If $G = SL(2,\mathbb{R})$, another dynamical quantity, the rotation number can be defined. Assume that $A \in C(\mathbb{T}^d, SL(2,\mathbb{R}))$ is homotopic to the identity. It introduces the projective skew-product $F_A \colon \mathbb{T}^d \times \mathbb{S}^1 \to \mathbb{T}^d \times \mathbb{S}^1$ with

$$F_A(\theta, w) := \left(\theta + \omega,\, \frac{A(\theta)v}{|A(\theta)v|}\right),$$

which is also homotopic to the identity. Thus we can lift F_A to a map $\widetilde{F}_A \colon \mathbb{T}^d \times \mathbb{R} \to \mathbb{T}^d \times \mathbb{R}$ of the form $F_A(\theta, y) = (\theta + \omega, y + \psi_\theta(y))$, where for every $\theta \in \mathbb{T}^d$, ψ_θ is $\mathbb{Z}$-periodic. The map $\psi \colon \mathbb{T}^d \times \mathbb{T} \to \mathbb{R}$ is called a *lift* of A. Let μ be any probability measure on $\mathbb{T}^d \times \mathbb{R}$ which is invariant by $\widetilde{F}_A$, and whose projection on the first coordinate is given by Lebesgue measure. The number

$$\rho(\omega, A) := \int_{\mathbb{T}^d \times \mathbb{R}} \psi_\theta(y)\, d\mu(\theta, y) \bmod \mathbb{Z}$$

which depends neither on the lift ψ nor on the measure μ, is called the *fibered rotation number* of (ω, A) (see Herman [1983] and Johnson and Moser [1982] for more details). It is known that $\rho(\omega, A) \in [0, \frac{1}{2}]$.

The continuous counterpart of quasi-periodic cocycles is the quasi-periodic linear systems, i.e., the ordinary differential equations

$$\dot{x} = A(\theta)x, \quad \dot{\theta} = \omega, \tag{1-8}$$

where A is assumed to be in a Lie algebra $\mathfrak{g}$. The eigenvalue equations of continuous quasi-periodic Schrödinger operator

$$(\mathfrak{L}_{q,\omega,\theta}\, y)(t) = -y''(t) + q(\theta + \omega t)y(t) = Ey(t) \tag{1-9}$$

are equivalent to the linear systems

$$\begin{cases} \dot{x} = V_{E,q}(\theta)x \\ \dot{\theta} = \omega \end{cases} \tag{1-10}$$

where

$$V_{E,q}(\theta) = \begin{pmatrix} 0 & 1 \\ q(\theta) - E & 0 \end{pmatrix} \in sl(2, \mathbb{R}).$$

The Poincaré map of the flow of (1-8) in fact defines a quasi-periodic cocycle. In converse, quasi-periodic cocycles close to constant can be embedded into the flow of quasi-periodic linear systems (You and Zhou [2013]). So there are paralell concepts, methods and theories for cocycles and systems.

Uniform hyperbolicity, the Lyapunov exponents, the rotation number and the acceleration are important concepts and quantities in the study of the dynamics of quasi-periodic cocycles and quasi-periodic linear systems. The central problems include positivity, continuity and regularity of the Lyapunov exponents, absolute continuity and Hölder continuity of the rotation number. Avila's acceleration is an important new index, its relation with dynamics and spectral theory has not been sufficiently explored.

1.3 Relations between operators and dynamical systems. For simplicity, the Lyapunov exponent, the rotation number and the IDS of (1-6) or (1-10) will be simply denoted by $L(E), \rho(E)$ and $N(E)$ when V and ω are fixed. The spectral theory of (1-1) (respectively (1-4)) are closely related to the dynamics of the one parameter family Schödinger cocycles (1-6) (respectively (1-10)) where the energy $E \in \mathbb{R}$ serves as parameter. Full understanding of the one parameter family of dynamical systems (1-6) or (1-10) would lead to a full understanding of the spectral theory of the Schrödinger operators (1-1) or (1-4).

There are some classical relationships between the spectrum of (1-1) (respectively (1-4)) and the dynamics of (1-6) (respectively (1-10)). It is known that $E \notin \Sigma$ if and only if the corresponding Schrödinger cocycle $(\omega, S_E^V(\cdot))$ is uniformly hyperbolic. IDS is the average of the spectral measure, which relates transparently to the rotation number by the formula $N(E) = 1 - 2\rho(E)$ and relates to the Lyapunov exponent through the Thouless formula

$$L(E) = \int \log|E - E'| \, dN(E').$$

Moreover, by Kotani's theory (Kotani [1984]), the absolutely continuous spectrum is the essential closure of the energies E such that (ω, S_E^V) has zero Lyapunov exponent. To obtain more precise information of the spectrum and the spectral measure, we need another tool: almost reducibility, which has been proved to be very powerful. In this survey, we will emphasize the applications of almost reducibility in the study of the spectral theory of the quasi-periodic Schrödinger operators.

2 Almost Reducibility

An analytic cocycle (ω, A) defined in (1-7) is said to be *reducible* if it can be conjugated to a constant cocycle, i.e., there exist $B \in C^\omega(2\mathbb{T}^d, \mathbf{G})$ and $C \in \mathbf{G}$ such that

$$B(\cdot+\omega)^{-1}A(\cdot)B(\cdot) = C.$$

Similarly, an analytic quasi-periodic linear system defined in (1-8) is said to be reducible if there exist $B \in C^\omega(2\mathbb{T}^d, \mathbf{G})$, $C \in \mathbf{g}$ such that

$$\partial_\omega B + BA - CB = C.$$

There are obstructions to the reducibility. The first obstruction is the presence of non-uniformly hyperbolicity. The second obstruction comes from the arithmetic condition on ω. Usually, reducibility requires that ω is Diophantine, i.e.,

$$\min_{l\in\mathbb{Z}} |\langle k, \omega\rangle - l| > \frac{\gamma^{-1}}{|k|^\tau}, \quad 0 \neq k \in \mathbb{Z}^d, \tag{2-1}$$

with fixed $\gamma, \tau > 1$. Here (γ, τ) are called the Diophantine constants of ω. Denote by $DC(\gamma, \tau)$ the set of all (γ, τ)-type Diophantine ω and $DC = \cup_{\gamma,\tau>1} DC(\gamma, \tau)$ (DC is of full measure). If $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, let $\frac{p_n}{q_n}$ be the $n-$th continued fraction convergents of irrational α, then we define

$$\beta(\alpha) = \limsup_{n\to\infty} \frac{\ln q_{n+1}}{q_n}.$$

$\beta(\alpha)$ measures how Liouvillean α is. Obviously if $\alpha \in DC$, then $\beta(\alpha) = 0$.

A quasi-periodic cocycle is in general not reducible when α is Liouvillean. A weaker concept than reducibility is almost reducibility. An analytic cocycle (ω, A) is said to be *almost reducible* if there exist a sequence of conjugations $B_n \in C^\omega_{h_n}(2\mathbb{T}^d, \mathbf{G})$, a sequence of constant matrices $A_n \in \mathbf{G}$ and a sequence of $F_n \in C^\omega_{h_n}(\mathbb{T}^d, \mathbf{g})$ converging to zero in C^ω topology, such that B_n conjugates (ω, A) to $(\omega, A_n e^{F_n(\cdot)})$. (ω, A) is called *weak almost reducible* if $h_n \to 0$, and *strong almost reducible* if $h_n \to h' > 0$. Some problems in the spectral theory needs strong almost reducibility and even require very precise estimates on B_n and F_n. Almost reducibility with more precise estimates is refered to as *quantitative almost reducibility*. For applications in the spectral theory of Schrödinger operators, the most interesting cases are $\mathbf{G} = SL(2,\mathbb{R})$ and $\mathbf{g} = sl(2,\mathbb{R})$.

Almost reducibility is useful and important since it prescribes a domain of applicability of local theories of cocycles close to constant (Avila [2010]). Due to its importance in the theory of dynamical systems and the spectral theory of quasi-periodic Schrödinger operators, reducibility has received much attention.

2.1 Perturbative reducibility. The rotation number of quasi-periodic linear system or quasi-periodic cocycle plays an important role in reducibility theory and its application to the the spectrum theory. We say that the rotation number ρ is rational with respect to (w.r.t. for short) ω if $\rho = \frac{1}{2}\langle k_0, \omega\rangle$ for some $k_0 \in \mathbb{Z}^2$, and to be Diophantine w.r.t. ω, with constants $\gamma, \tau > 1$, if

$$\min_{l\in\mathbb{Z}} |\langle k, \omega\rangle - 2\rho - l| \geq \frac{\gamma^{-1}}{|k|^{\tau}}, \quad k \in \mathbb{Z}^2.$$

We denote by $DC_\omega(\gamma, \tau)$ the set of all such ρ. It is well known that the union $DC_\omega = \cup_{\gamma,\tau>1} DC_\omega(\gamma, \tau)$ is a full measure subset of $\mathbb{R}$.

The reducibility of quasi-periodic linear systems (1-10) and its applications in the spectral theory were initiated by Dinaburg and Sinaĭ [1975], based on classical KAM theory, they proved that if q is analytic and sufficiently small, then (1-10) is reducible for $\rho(E) \in DC_\omega(\gamma, \tau)$. Dinaburg and Sinai's reducibility result implies the existence of absolutely continuous spectrum of the Schrödinger operator (1-4). The first breakthrough was due to (Eliasson [1992]), who proved the following:

Theorem 2.1. *Eliasson [ibid.] Suppose that $\omega \in DC(\tau, \gamma)$ and q is analytic and sufficiently small, then* (1-10) *is weak almost reducible for all E. Moreover* (1-10) *is reducible if $\rho(E) \in DC_\omega$ or rational w.r.t ω.*

The proof in Eliasson [ibid.] uses a crucial resonance-cancelation technique which was introduced by Moser and Pöschel [1984] earlier. Eliasson [1992] work has profound impact: Theorem 2.1 can describe the dynamical behavior for all parameters E, while the classical KAM theory can only describe a positive measure set of E. Theorem 2.1 implies that, when the potential is analytic and small, the spectral measure of (1-9) is purely absolutely continuous for all phases θ, which shows that almost reducibility could play an important role in the study of the spectrum of quasi-periodic Schrödinger operators. Moreover, the later non-perturbative and quantitative versions of Theorem 2.1 have been found useful in the study of Cantor spectrum, gap estimates, Anderson localization, Hölder continuity of IDS and even more, which we will review separately in the following sections.

Theorem 2.1 holds for more general quasi-periodic cocycles $A \in C^\omega(\mathbb{T}^d, \mathbf{G})$ with A close to some constant (Chavaudret [2013] and Krikorian [1999b,a]), even for finite smooth case (Cai, Chavaudret, You, and Zhou [2017]). We remark that all the above mentioned results are *perturbative*, i.e., the smallness of q depends on the frequencies ω through the Diophantine constants (γ, τ). The perturbative reducibility result is optimal when $d \geq 2$ in the discrete case and $d > 2$ in the continuous case as shown by a counterexample of Bourgain [2002]. However, when $d = 2$ in the continuous case and $d = 1$ in the discrete case, one can expect more. In the following we shall restrict our attention to these cases.

2.2 Non-perturbative reducibility. The non-perturbative reducibility means that the smallness of the perturbation does not depend on the Diophantine constants (γ, τ) of α. The non-perturbative reducibility was first proved by Puig [2006] for Schrödinger cocycles $(\alpha, S_E^V(\cdot))$ with one frequency $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. However, the proof, which is based on Aubry duality (Aubry and André [1980] and Gordon, Jitomirskaya, Last, and Simon [1997] and Anderson localization results of Bourgain and Jitomirskaya [2002a], doesn't work for the continuous linear systems. Hou and You [2012] gave a non-perturbative version of Theorem 2.1 in the continuous case.

Theorem 2.2. *Hou and You [ibid.] Let $h > 0$ and $\omega = (\alpha, 1)$ with $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. Consider*

$$\begin{cases} \dot{x} = (A + F(\theta))x \\ \dot{\theta} = \omega \end{cases} \tag{2-2}$$

with $A \in sl(2,\mathbb{R})$ and $F \in C_h^{\omega}(\mathbb{T}^2, sl(2,\mathbb{R}))$. Then there exists $\delta = \delta(A,h) > 0$ depending on A, h but not on α, such that system (2-2) is weak almost reducible if $\sup_{|Im\theta|<h} |F(\theta)| < \delta$. Moreover (2-2) is reducible if ω is Diophantine and $\rho(E) \in DC_{\omega}$ or rational w.r.t ω.

Remark 2.1. *By an embedding theorem of You and Zhou [2013], one sees that the same result in Theorem 2.2 holds for $SL(2,\mathbb{R})$ cocycles with one frequency.*

We remark that Theorem 2.2 works for any $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ not merely Diophantine frequency, its proof is based on KAM and Floquet theory. Before Hou and You [2012], Avila, Fayad, and Krikorian [2011] proved that for any analytic $SL(2,\mathbb{R})$ cocycles (α, A) which is close to constant, for any $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, the cocycles are analytic rotations reducible (analytic conjugacy to a cocycle with values in $SO(2,\mathbb{R})$) for full measure rotation number, their proof is based on "algebraic conjugacy trick" which was first developed by Fayad and Krikorian [2009].

2.3 Strong almost reducibility and Quantitative almost reducibility. We remark that in the above mentioned almost reducibility results, the convergence to constants occurs on analytic strips of width going to zero. Breakthrough belongs to Avila and Jitomirskaya [2010]. Based on almost localization and Aubry duality, Avila and Jitomirskaya [ibid.] gave a non-perturbative strong almost reducibility result for Schrödinger cocycles with a single frequency $\alpha \in DC$ and small potentials. Avila [2008] generalized the result to $\beta(\alpha) = 0$ with much more delicate estimates. Chavaudret [2013] proved a strong almost reducibility result in the local regime for multiple Diophantine frequencies. However, as we mentioned, in spectral applications, we need quite delicate quantitative estimates, while Chavaudret's estimates are not enough to give interesting consequence in applications. Recently, Leguil, You, Zhao, and Zhou [2017] gave another strong almost reducibility result with more precise estimates. As a consequence, several interesting spectral applications were obtained. We will explain the applications in Section 3.

2.4 Global reducibility. In all the results above, we assume that the cocycle or system is close to a constant. For cocycles not close to a constant, Kotani's theory (Kotani [1984]) essentially asserts that there is an almost surely dichotomy between non-uniform hyperbolicity and L^2 rotations-reducibility of the cocycles (α, S_E^V). L^2-conjugation can be further proved to be smooth by a renormalization scheme developed by Avila and Krikorian [2006, 2015] and Krikorian [2004]. Thus for any $\alpha \in \mathbb{R}\backslash\mathbb{Q}$, and for $a.e. E \in \mathbb{R}$, (α, S_E^V) is either analytically rotations reducible or non-uniformly hyperbolic (Avila and Krikorian [2006] and Avila, Fayad, and Krikorian [2011]).

The real breakthrough is Avila's global theory for one frequency analytic $\mathrm{SL}(2,\mathbb{R})$ cocycles Avila [2015a]. Avila classified $(\alpha, A) \in \mathbb{R}\backslash\mathbb{Q}\times C^\omega(\mathbb{T}, SL(2,\mathbb{R}))$, which is not uniformly hyperbolic, into three classes according to the Lyapunov exopents and acceleration: supercritical, subcritical and critical. A cocycle (α, A) is supercritical, if $L(\alpha, A) > 0$, it is called subcritical, if $L(\alpha, A(z)) = 0$ for $|\Im z| \leq \delta$, it is called critical otherwise. We say that H is acritical if (α, S_E^v) is not critical for all $E \in \Sigma$. The main result in Avila's global theory is the following:

Theorem 2.3. *Avila [ibid.] Let $\alpha \in \mathbb{R}\backslash\mathbb{Q}$. Then for a (measure-theoretically) typical analytic potential $V \in C^\omega(\mathbb{T}, \mathbb{R})$, the operator $H_{V,\alpha,\theta}$ is acritical.*

Avila's global theory is crucial in the study of the spectral theory of Schrödinger operators, especially when the potential is not small but the corresponding Schrödinger cocycles are still in the subcritical region. The cornerstone in Avila's global theory is his *Almost Reducibility Conjecture(ARC)*: subcriticity implies almost reducibility, which has been solved completely by Avila [2010, n.d.(b)].

Theorem 2.4 (Avila [2010, n.d.(b)]). *Let $A \in C^\omega(\mathbb{T}, \mathrm{SL}(2,\mathbb{R}))$. Then (α, A) is strong almost reducible if it is subcritical.*

However, we remark that sometimes Avila's global almost reducibility does not contain sufficient estimates on the conjugations(since it deals with global cocycle directly). In applications, we have to cook it with finer local quantitative almost reducibility results.

3 Applications to quasi-periodic Schrödinger operators

Quasi-periodic Schrödinger operators are mathematical models for many subjects in quantum physics including quantum Hall effect and the nature of quasi-crystal. It is also a subject to test the power of mathematical theories and methods, thus has attracted much attentions, which we refer to survey articles of Damanik [2017], Jitomirskaya [2007], Marx and Jitomirskaya [2017], Last [2005], and Simon [1982]. In this survey we will only present some of them which are closely related to the theory of almost reducibility, readers are invited to consult the former references for other interesting results.

3.1 Spectrum of quasi-periodic Schrödinger operators. The spectrum Σ is one of most important objects in the spectral theory of quasi-periodic Schrödinger operators.

3.1.1 Cantor spectrum. Cantor spectrum was conjectured to be a generic phenomenon for one dimensional almost periodic Schrödinger operator (Problem 6 of Simon [1982]). There are few exceptions in this case (the so called finite gap potentials). In one frequency case, there is no counter-example with big potential so far, but the recent work of Goldstein-Schlag-Voda shows that finite gap happens for multi-frequency case (Goldstein, Schlag, and Voda [2017]).

However, to prove the existence of Cantor spectrum is not an easy task. Eliasson [1992] proved that for any given $\omega \in DC(\gamma,\tau)$, $H_{V,\omega,\theta}$ has Cantor spectrum for *generic* small analytic potentials. His proof is based on Moser-Pöschel argument (Moser and Pöschel [1984]) and the fact: if $\rho(E)$ is rational w.r.t ω, then (ω, S_E^V) is reducible. Eliasson's proof applies to the cocycle case, however his proof is not constructive, which can not provide any concrete example.

In the discrete case, the situation is better. Goldstein and Schlag proved that for any fixed non-constant analytic potential, in the supercritical region, the spectrum is a Cantor set for almost all $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ (Goldstein and Schlag [2011]). In C^0-topology, Avila-Bochi-Damanik proved Cantor spectrum for any fixed totally irrational vector $\omega \in \mathbb{T}^d$ and generic $V \in C^0(\mathbb{T}^d,\mathbb{R})$ (Avila, Bochi, and Damanik [2009]). In the case of C^k-topology ($1 \le k \le \infty$ or even in analytic category), based on Avila [2011] and Goldstein and Schlag [2011], one can prove that for generic $\alpha \in \mathbb{R}\backslash\mathbb{Q}$ and generic $V \in C^k(\mathbb{T},\mathbb{R})$, the spectrum of $H_{V,\omega,\theta}$ is a Cantor set, one can consult footnote 1 of J. Wang, Zhou, and Jäger [2016] for the outline of this proof.

The most remarkable progress is for almost Mathieu operator. AMO $H_{\lambda,\alpha,\theta}$ has been conjectured for a long time to have Cantor spectrum for irrational α (consult a 1964 paper of Azbel [1964]). This conjecture has been dubbed the Ten Martini Problem by Simon [2000], after an offer of Kac [n.d.] in 1981. Since when it was posed, Ten Martini Problem became the central problem in the spectral theory of quasi-periodic Schödinger operators and attracted a lot of attentions. Finally, it was completely solved by Avila and Jitomirskaya [2009] (see the references therein for partial advances). One main ingredient of the proof is the reducibility: the cocycle cann't be reducible to rotations for all E in an interval.

We should also mention the works by Sinaĭ [1987] and Y. Wang and Z. Zhang [2017] where the Cantor spectrum was proved for $H_{\lambda V,\alpha,\theta}$ with sufficiently large λ and C^2 cosine-like V. So far, those are the only known concrete operators having Cantor spectrum. Particularly, there is no concrete examples of Cantor spectrum in the continuous time case. In a forthcoming paper, we will give more concrete examples of Cantor spectrum for both discrete and continuous quasi-periodic Schrödinger operators based on reducibility (Hou, Shan, and You [n.d.]).

3.1.2 All gaps are open. Certainly, one should not expect that all gaps are open for all analytic quasi-periodic Schrödinger operators because of, as we mentioned before, the existence of the finite gap potentials. But AMO is special. Motivated by Hofstadter's numerical result (Hofstadter [1976]), Kac [n.d.] raised the well-known question: Are all possible spectral gaps of AMO open? The problem was named as the Dry Ten Martini Problem by Simon [1982]. The Dry Ten Martini Problem was also named as Ten Martini Conjecture by physicists, which has particular importance in quantum physics such as the Integer Quantum Hall effect. Some works in physics have been done under the assumption that the Dry Ten Martini Problem is true, see i.g. Osadchy and J. E. Avron [2001].

Obviously, Dry Ten Martini Problem automatically implies Ten Martini Problem. Certainly people want to solve this original problem of Kac. In the last thirty years, substantial progresses were made by Choi, Elliott, and Yui [1990], Puig [2004], Avila and Jitomirskaya [2009, 2010]. However, the problem has not been completely solved for any fixed λ.

Recently, Avila, You, and Zhou [2016] gave a complete answer to the Dry Ten Martini Problem for the non-critical case $\lambda \neq 1$ by quantitative almost reducibility.

Theorem 3.1. *Avila, You, and Zhou [ibid.] $H_{\lambda,\alpha,\theta}$ has all spectral gaps open for all irrational α and all $\lambda \neq 1$.*

The strategy of Puig's proof Puig [2004] is to prove that (α, S_E^{λ}) is reducible but it cann't be reduced to (α, Id) if $N_{\lambda,\alpha}(E)$ is rational w.r.t α. Developing this idea, Avila and Jitomirskaya solved the problem for $\alpha \in DC$, and $\lambda \neq 1$ (Avila and Jitomirskaya [2010]). However, if $\beta(\alpha) > 0$ as in our case, one cann't expect that the cocycle is still reducible. However, we can show that the cocycle can not be almost reducible to (α, Id) with fast decay of $\|B_n\|_h \|F_n\|_h$. Once we have this, Theorem 3.1 can be proved by a modified Moser-Pöschel argument. We finally remark that the Dry Ten Martini problem has not been completely solved for $\lambda = 1$.

3.1.3 Estimate of the spectral gaps. Recall G_k, the spectral gaps with labelling k. The well-known Dry Ten Martini Problem asks whether G_k is empty or not for AMO. Further problem is: how big the gaps are? More precisely, can we give any lower bound or upper bound for G_k? For AMO, Leguil, You, Zhao, and Zhou [2017] proved the following result by quantitative almost reducibility:

Theorem 3.2. *Leguil, You, Zhao, and Zhou [ibid.] For $\alpha \in \mathrm{DC}$, and for any $0 < \xi < 1$, there exist constants $C = C(\lambda, \alpha, \xi) > 0$, $\tilde{C} = \tilde{C}(\lambda, \alpha)$, such that for all $k \in \mathbb{Z} \backslash \{0\}$,*

$$\tilde{C}\lambda^{\tilde{\xi}|k|} \leq |G_k(\lambda)| \leq C\lambda^{\xi|k|}, \qquad \text{if} \quad 0 < \lambda < 1,$$

$$\tilde{C}\lambda^{-\tilde{\xi}|k|} \leq |G_k(\lambda)| \leq C\lambda^{-\xi|k|}, \qquad \text{if} \quad 1 < \lambda < \infty,$$

where $\tilde{\xi} > 1$ is a numerical constant, $|G_k(\lambda)|$ denotes the length of $G_k(\lambda)$.

For general analytic potential, Leguil, You, Zhao, and Zhou [ibid.] also proved that $|G_k(V)| \leq \varepsilon_0^{\frac{2}{3}} e^{-r|k|}$ for all $k \in \mathbb{Z}^d \backslash \{0\}$ and any $r \in (0, h)$ if $\sup_{|\Im x| < h} |V(x)| < \varepsilon_0$ is small enough and $\omega \in \mathrm{DC}$ and $V \in C_h^{\omega}(\mathbb{T}^d, \mathbb{R})$. Before Leguil, You, Zhao, and Zhou [ibid.], Damanik and Goldstein [2014] have shown that $|G_k(V)| \leq \epsilon_0 e^{-\frac{h}{2}|k|}$. We remark that the proof in Damanik and Goldstein [ibid.] is based on the localization argument, which cannot be directly applied to the discrete case, while the proof in Leguil, You, Zhao, and Zhou [2017] is based on reducibility, so it works equally well both for the continuous time operators and discrete time operators. For more history on the study of the upper bounds, one may consult Leguil, You, Zhao, and Zhou [ibid.] and the references therein.

In methodology, for estimating of the spectral gaps we need to analyze the behavior of Schrödinger cocycle at the edge points of the spectral gaps, where the cocycles are reducible to constant parabolic cocycles. The crucial points for the gap estmate are the proof of the exponential decay of the off-diagonal element of the parabolic matrix and the exponential growth of the conjugacy with respect to the labelling k. Furthermore, in order to prove the decay rate to be uniform with respect to the labelling k, we need the strong quantitative almost reducibility result, i.e. the cocycle is almost reducible in a fixed band, with precise estimates on the conjugations and the off-diagonal element of the (reduced) parabolic matrix (Leguil, You, Zhao, and Zhou [ibid.]).

3.1.4 Homogeneous spectrum. we say that Σ is μ-homogenuous if for any $E \in \Sigma$ and any $0 < \epsilon \leq \operatorname{diam} \Sigma$, we have $|\mathcal{S} \cap (E - \epsilon, E + \epsilon)| > \mu\epsilon$ for $\mu > 0$. We say that $H_{V,\alpha,\theta}$ has homogenuous spectrum if Σ is homogenuous. The homogeneity of the spectrum plays an essential role in the inverse spectral theory of almost periodic potentials(refer to the fundamental work of Sodin and Yuditskii [1995, 1997]).

The exponential decay of the spectral gaps can be used to prove the homogeneity of the spectrum.

Theorem 3.3 (Leguil, You, Zhao, and Zhou [2017]). *Let $\alpha \in$ SDC [1]. For a (measure-theoretically) typical analytic potential $V \in C^{\omega}(\mathbb{T}, \mathbb{R})$, the spectrum $\Sigma_{V,\alpha}$ is μ-homogeneous for some $0 < \mu < 1$. Especially, the spectrum is homogenuous for small analytic potentials.*

Homogeneity of the spectrum $\Sigma_{V,\alpha}$ in the subcritical regime is derived from the upper bounds of the spectral gaps and Hölder continuity of the IDS (Leguil, You, Zhao, and Zhou [ibid.]). While the homogeneity in supercritical region was proved by Damanik, Goldstein, Schlag, and Voda [2015]. Together with Theorem 2.3, one sees that the homogeneity of the spectrum is a typical phenomenon for analytic Schrödinger operators when α is strong Diophantine. We remark that, at least in the subcritical region, strong Diophantine is not necessary (Leguil, You, Zhao, and Zhou [2017]), however some kind of arithmetic property is necessary. After Leguil, You, Zhao, and Zhou [ibid.], Avila, Last, Shamis, and Zhou [n.d.] proved that there exists a dense set of Liouvillean frequencies α such that $\Sigma_{\lambda,\alpha}$ of AMO is not homogeneous.

3.2 The spectral measure, IDS and Lyapunov exponent.

The nature of the spectral measure, IDS and Lyapunov exponent are central subjects in the spectral theory of quasi-periodic Schrödinger operators, while the study of Lyapunov exponents is also a central subject in smooth dynamical systems.

3.2.1 Anderson localization.

There are two important results concerning Anderson localization in supercritical regime. The first result belongs to Jitomirskaya [1999], who proved that for almost Mathieu operator, $H_{\lambda,\alpha,\theta}$ has Anderson localization for a.e. θ if $|\lambda| > 1$ and $\alpha \in DC$. Another result belongs to Bourgain and Goldstein [2000], who proved that up to a typical perturbation of the frequency, Anderson localization holds through the supercritical regime. Comparing the two results above, the result of Jitomirskaya [1999] is for fixed frequency and typical phases (depending on the frequency), while Bourgain and Goldstein [2000] is for fixed phase and typical frequencies (depending on the phase). Both results are proved by the positivity of the Lyapunov exponent, which is classical method for studying the pure point spectrum of the Schrödinger operators.

At first glance, the reducibility has no business with the Anderson localization spectrum since the point spectrum corresponds to the non-uniformly hyperbolicity which is definitely not almost reducible. Surprisingly, one can use reducibility to study the point spectrum, even Anderson localization and dynamical localization. This idea was first

[1] SDC is the set of strong Diophantine numbers, i.e., there exist $\gamma, \tau > 0$ such that

(3-1) $$\inf_{j \in \mathbb{Z}} |n\alpha - j| \geq \frac{\gamma}{|n|(\log|n|)^{\tau}}, \quad \forall n \in \mathbb{Z}\backslash\{0\}.$$

appeared in You and Zhou [2013], and completely built in Avila, You, and Zhou [2017]. The bridge is the Aubry duality (Aubry and André [1980] and Gordon, Jitomirskaya, Last, and Simon [1997]): Suppose that the quasi-periodic Schrödinger operator (1-1) with one frequency has an analytic quasi-periodic Bloch wave $u_n = e^{2\pi i n\varphi}\overline{\psi}(n\alpha+\phi)$ for some $\overline{\psi} \in C^{\omega}(\mathbb{T},\mathbb{C})$ and $\varphi \in [0,1)$, then the Fourier coefficients of $\overline{\psi}(\theta)$ satisfy the following long range operator:

$$(\widehat{H}_{V,\alpha,\varphi}x)_n = \sum_{k\in\mathbb{Z}} V_k x_{n-k} + 2cos2\pi(\varphi+n\alpha)x_n = Ex_n, \tag{3-2}$$

where V_k is the Fourier coefficients of $V(\theta)$. The converse is also true. We remark that the almost Mathieu family $\{H_{2\lambda\cos,\alpha,\theta}\}_{\lambda>0}$ is self-dual. The reducibility of $(\alpha, S_E^V(\theta))$ will provide analytic quasi-periodic Bloch waves of the operator (1-1) and thus will provide eigenfunctions for its dual operator, so the general philosophy is that the full measure reducibility of $(\alpha, S_E^V(\theta))$ will imply Anderson localization of the dual operator $\widehat{H}_{V,\alpha,\varphi}$ for almost every phases. Let us mention a recent work by Avila, You, and Zhou [2017] for the almost Mathieu operators as an example.

Theorem 3.4. *Avila, You, and Zhou [ibid.] If $|\lambda| > e^{\beta(\alpha)}$, then $H_{\lambda,\alpha,\theta}$ has Anderson localization for a.e. θ.*

Note that $H_{\lambda,\alpha,\theta}$ has purely absolutely continuous spectrum for all θ if $|\lambda| < 1$ (Avila [2008], Avila and Damanik [2008], Avila and Jitomirskaya [2010], and Jitomirskaya [1999]), and $H_{\lambda,\alpha,\theta}$ has purely singular continuous spectrum for all θ if $1 \le |\lambda| < e^{\beta(\alpha)}$ (Avila, You, and Zhou [2017], J. Avron and Simon [1982], and Gordon [1976]). Now one sees the sharp phase transition scenario for three types of the spectral measure for a.e. θ, and solves a conjecture of Jitomirskaya [1995], which is the corrected version of a conjecture by Aubry and André [1980]. We remark that based on localization method, before Avila, You, and Zhou [2017], Avila and Jitomirskaya [2009] proved that $H_{\lambda,\alpha,\theta}$ has Anderson localization for a.e. θ, if $|\lambda| > e^{16\beta(\alpha)/9}$. More recently, Jitomirskaya and Liu [n.d.] proved a refined result of Avila, You, and Zhou [2017] with precise description on the localized phases and the hierarchical structure of eigenfunctions. We also refer to Jitomirskaya and S. Zhang [2015] for another interesting phase transition result, valid for general analytic potentials.

For the proof of Theorem 3.4, a new criterion (which reveals the fact that nice asymptotical distribution of the eigenfunctions implies Anderson localization) for establishing the purity of the point spectrum was developed in Avila, You, and Zhou [2017], which applies to general ergodic family of operators (one may consult Jitomirskaya and Kachkovskiy [2016] for another proof but with same spirit). Compared with traditional localization argument, the trade off is the loss of precise arithmetic control on the localization phases.

However, by this approach, we indeed establish a kind of equivalence between full measure reducibility of Schrödinger cocycles and Anderson localization of the dual operators. The methods developed in Avila, You, and Zhou [2017] has further applications. For example, it can be used to study the spectral properties at the transition line $\lambda = e^{\beta}$ (Avila, Jitomirskaya, and Zhou [2018]). In a forthcoming paper, we even show that it can be used to study the dynamical localization of the long-range operator and a family of Schrödinger operators on $l^2(\mathbb{Z}^d)$. Finally, we remark that it has been an open question for a long time whether in the supercritical regime, $H_{V,\alpha,\theta}$ with fixed Diophantine α has Anderson localiztion for a.e. phase (consult Eliasson [1997] for partial advances). Our method might provide a way to study this problem when the potential is a trigonometric polynomial. In this case, Equation (3-2) naturally defines a $2d$-dimensional cocycles, while the full measure reducibility is easy to establish, the difficulty remains in the proof is the purity of the point spectrum.

3.2.2 Absolutely continuous spectrum. Absolute continuity of the spectral measure is a traditional territory of reducibility. If the cocycles (α, S_E^V) are reducible for positive measure of $E \in \Sigma$, then the operator has absolutely continuous spectrum (Dinaburg and Sinaĭ [1975]). Based on Theorem 2.1, Eliasson [1992] proved directly that the spectral measure $H_{V,\alpha,\theta}$ is absolutely continuous spectrum if the potential V is small enough and $\alpha \in DC$. Recently, Avila [2008] gave a new understanding of Eliasson's result based on Gilbert-Person's subordinacy theory Gilbert and Pearson [1987]: $\mu_{\alpha,\theta,V}|_{\mathfrak{B}}$ is absolutely continuous for all $\theta \in \mathbb{R}$ where $\mathfrak{B}$ is the set of $E \in \mathbb{R}$ such that the cocycle (α, S_E^V) is bounded. Thus to obtain purely absolutely continuous spectrum, one only needs to show that $\mu_{\alpha,\theta,V}(\Sigma \backslash \mathfrak{B}) = 0$. Note that by reducibility, we can prove that $\mathfrak{A}_n$ are bounded for almost all E in the spectrum Σ. However the bounds are not uniform, the proof of $\mu_{\alpha,\theta,V}(\Sigma \backslash \mathfrak{B}) = 0$ relies on the measure estimate of $E \in \Sigma$ for any given bound. Based on this idea, Avila [2008] proved purely absolutely continuous spectrum for general one frequency analytic Schödinger operators if the potential is small and $\beta(\alpha) = 0$. Recent breakthrough also belongs to Avila [n.d.(b)], he shows that almost reducibility actually implies pure absolutely continuous spectrum. Together with Theorem 2.3 and formly mentioned Bourgain-Goldstein's result (Bourgain and Goldstein [2000]), it implies that typical one frequency analytic Schödinger operators don't have singular continuous spectrum.

Concerning the absolutely continuous spectrum, we also mention the well known Kotani-Last conjecture (Kotani and Krishna [1988]). It says that if an one-dimensional ergodic Schrödinger operator has absolutely continuous spectrum, then its potential is almost periodic. By periodic approximation, Avila [2015b] constructed non-almost periodic Schrödinger operators with absolutely continuous spectrum both for the discrete and continuous cases. Another independent work was due to Volberg and Yuditskii [2014], they constructed, by inverse spectral theory, counter-examples in the discrete case (see also the example Damanik and Yuditskii [2016] for the continuous case). The reducibility theory can also provide another approach to construct counterexamples in the continuous case (You and Zhou [2015]). The idea is that reducibility theory and subordinacy theory ensures the existence of ac spectrum, while time scaling make the potential non-almost periodic.

3.2.3 Continuity of Lyapunov exponent and IDS. By Thouless formula and the non-negativity of $L(E)$, one knows that $N(E)$ is always Log-Hölder continuous and that the Hölder continuity of $L(E)$ is equivalent to the Hölder continuity of $N(E)$. IDS is the average of the spectral measure, in general it is more regular than the spectral measure, in fact it is always continuous. However, behavior of the Lyapunov exponents of quasi-periodic cocycles is very complicated. They could be discontinuous in the space of smooth $SL(2,\mathbb{R})$ cocycles (Bochi [2002] and Furman [1997] for C^0 case, Y. Wang and You [2013] for smooth case). Different from the smooth case, the Lyapunov exponent is alway continuous in the space of analytic $SL(2,\mathbb{C})$ cocycles (Bourgain [2005b], Bourgain and Jitomirskaya [2002b], and Jitomirskaya, Koslover, and Schulteis [2009]), even in the space of higher dimensional $GL(d,\mathbb{C})$ cocycles (Avila, Jitomirskaya, and Sadel [2014]). The continuity of Lyapunov exponents implies that the set of the cocycles with positive Lyapunov exponent is open in analytic topology. Together with the denseness result by Avila [2011], one knows that the set of quasi-periodic cocycles with positive Lyapunov exponent is open and dense in analytic topology, but this result is not true in the space of smooth quasi-periodic cocycles (Y. Wang and You [2015]).

One could expect the Hölder continuity in analytic case when the freqencies satisfy some arithmetic conditions. In the supercritical region, Goldstein and Schlag [2001] proved the Hölder continuity of $L(E)$ if $V(x)$ is analytic and α is strong Diophantine. You and S. Zhang [2014] generalized Goldstein-Schlag's result to all Diophantine α and some weaker Liouvillean α, which shows that the Diophantine condition on ω is not necessary for the Hölder continuity of $L(E)$. However, some kind of arithmetic assumptions on α is neccessary (Bourgain [2005a]). Recently, Avila, Last, Shamis, and Zhou [n.d.] proved that IDS of $H_{\lambda,\alpha,\theta}$ is not even weak Hölder if α is extremely Liouvillean.

For Diophantine frequency, the modulus of the Hölder continuity is not very clear so far. It is already known that it can not be better than $\frac{1}{2}$-Hölder. In the supercritical regime, and if furthermore the potential V is in a small L^∞ neighborhood of a trigonometric polynomial of degree d, then the IDS is $(\frac{1}{2d}-\epsilon)$-Hölder for all $\epsilon>0$ (Goldstein and Schlag [2008]), and it is exactly $\frac{1}{2}$-Hölder for AMO (Avila [2008] and Avila and Jitomirskaya [2010]). However, $(\frac{1}{2d}-\epsilon)$-Hölder continuity is surely not optimal. By reducibility argument, we conjecture that the modulus of Hölder continuity of $L(E)$ is at least $\frac{1}{2N}$, where N is the acceleration of the Schrödinger cocycle $(\alpha, S_E^V(\cdot))$.

For small analytic potentials, the reducibility argument was used by Hadj Amor [2009] to prove the $\frac{1}{2}$-Hölder continuity of the IDS and the Lyapunov exponent if ω is Diophantine. However when dealing with subcritical regime, her approach does not work since the estimates need explicit dependence on the parameters. In fact, when reducing the global potential to local regimes by Avila's global theory, the explicit dependence of the parameters is lost. Based on Thouless's formula, Avila and Jitomirskaya [2010] developed a new understanding of the problem. In order to prove that IDS is $\frac{1}{2}$ Hölder, it is sufficient to prove

$$L(E+i\epsilon)-L(E)\leq \epsilon^{1/2},$$

which relates to the growth of cocycles $\|\mathcal{A}_n\|_{C^0}$. Thus in the almost reducible scheme, one only needs to estimate the C^0 norm of $\|B_n\|$ and $\|F_n\|$. By this method and Avila's global theory, one can show in the subcritical region for $\beta(\alpha)=0$, then IDS is $\frac{1}{2}$-Hölder continuous (Avila [2008], Avila and Jitomirskaya [2010], Avila [n.d.(b)], and Leguil, You, Zhao, and Zhou [2017]). The method also works for finite smooth potentials, recently Cai, Chavaudret, You, and Zhou [2017] proved the $\frac{1}{2}$-Hölder continuity of IDS for operators with finite smooth small potentials and Diophantine frequency.

3.2.4 Positivity of Lyapunov exponent. The positivity of $L(E)$ is also a big issue. Actually, it is difficult to compute. Herman [1983] proved that, by the subharmonicity method, $L(E)\geq \ln|\lambda|$ for almost Mathieu operator $H_{\lambda,\alpha,\theta}$ with $|\lambda|>1$. By continuity of Lyapunov exponent (Bourgain and Jitomirskaya [2002b]), it was further proved that

$$L(E)=\max\{0,\ln|\lambda|\},$$

for $E\in\Sigma$ (consult Avila [2015a] for another elegant proof). Herman's subharmonicity trick also works for trigonometric polynomials $\lambda V(x)$ with large λ (Herman [1983]). The generalization to arbitrary one-frequency nonconstant real analytic potentials was given by Sorets and Spencer [1991], who proved that if $|\lambda|\geq\lambda_0$, then

$$L(E)\geq \ln|\lambda|-C, \tag{3-3}$$

Here, C and λ depend on V but not on α (consult Bourgain [2005b], Bourgain and Goldstein [2000], Duarte and Klein [2014], Goldstein and Schlag [2001], and Z. Zhang [2012] for simplified proofs and generalizations) .

Compared with the very precise estimate of $L(E)$ for the almost Mathieu operators, the formula (3-3) is too rough. Based on a generalized Thouless formula by Haro and Puig [2013] and the large deviation theorem for identically singular cocycles, Duarte and Klein [n.d.] showed that

$$L(E)=\log\lambda+\int|V(\theta)-\frac{E}{\lambda}|\,d\theta+O(e^{-c(\ln|\lambda|)^b}),$$

where $c > 0$ and $0 < b < 1$. Han and Marx [2018] further improved the bound to

$$L(E) = \log \lambda + \int |V(\theta) - \frac{E}{\lambda}| \, d\theta + O(|\lambda|^{-\frac{2}{2N+1}}),$$

where $N = N(V)$ is a large number. The proof of Han and Marx [ibid.] relies on estimating the acceleration of the cocycle which is defined by Avila [2015a]. In a forthcoming paper, we will show that, based on almost reducibility and Aubry duality,

$$L(E) = \log \lambda + \int |V(\theta) - \frac{E}{\lambda}| \, d\theta + O(\lambda^{-\frac{1}{2d}}),$$

for trigonometric polynomial potentials of order d and sufficiently large λ.

Acknowledgments. The author would like to thank Qi Zhou for his careful reading of the manuscript and many valuable suggestions.

References

Serge Aubry and Gilles André (1980). "Analyticity breaking and Anderson localization in incommensurate lattices". In: *Group theoretical methods in physics (Proc. Eighth Internat. Colloq., Kiryat Anavim, 1979)*. Vol. 3. Ann. Israel Phys. Soc. Hilger, Bristol, pp. 133–164. MR: 626837 (cit. on pp. 2138, 2144).

Artur Avila (n.d.[a]). "Almost reducibility and absolute continuity II". In preparations.

– (n.d.[b]). "KAM, Lyapunov exponents and spectral dichotomy for one-frequency Schrödinger operators" (cit. on pp. 2139, 2145, 2147).

– (Oct. 2008). "The absolutely continuous spectrum of the almost Mathieu operator". arXiv: 0810.2965 (cit. on pp. 2131, 2138, 2144–2147).

– (June 2010). "Almost reducibility and absolute continuity I". arXiv: 1006.0704 (cit. on pp. 2136, 2139).

– (2011). "Density of positive Lyapunov exponents for $SL(2, \mathbb{R})$-cocycles". *J. Amer. Math. Soc.* 24.4, pp. 999–1014. MR: 2813336 (cit. on pp. 2140, 2146).

– (2015a). "Global theory of one-frequency Schrödinger operators". *Acta Math.* 215.1, pp. 1–54. MR: 3413976 (cit. on pp. 2131, 2133, 2139, 2147, 2148).

– (2015b). "On the Kotani-Last and Schrödinger conjectures". *J. Amer. Math. Soc.* 28.2, pp. 579–616. MR: 3300702 (cit. on p. 2145).

Artur Avila, Jairo Bochi, and David Damanik (2009). "Cantor spectrum for Schrödinger operators with potentials arising from generalized skew-shifts". *Duke Math. J.* 146.2, pp. 253–280. MR: 2477761 (cit. on p. 2140).

Artur Avila and David Damanik (2008). "Absolute continuity of the integrated density of states for the almost Mathieu operator with non-critical coupling". *Invent. Math.* 172.2, pp. 439–453. MR: 2390290 (cit. on p. 2144).

Artur Avila, Bassam Fayad, and Raphaël Krikorian (2011). "A KAM scheme for $SL(2, \mathbb{R})$ cocycles with Liouvillean frequencies". *Geom. Funct. Anal.* 21.5, pp. 1001–1019. MR: 2846380 (cit. on pp. 2138, 2139).

Artur Avila and Svetlana Jitomirskaya (2009). "The Ten Martini Problem". *Ann. of Math. (2)* 170.1, pp. 303–342. MR: 2521117 (cit. on pp. 2131, 2140, 2141, 2144).

– (2010). "Almost localization and almost reducibility". *J. Eur. Math. Soc. (JEMS)* 12.1, pp. 93–131. MR: 2578605 (cit. on pp. 2138, 2141, 2144, 2146, 2147).

Artur Avila, Svetlana Jitomirskaya, and Christian Sadel (2014). "Complex one-frequency cocycles". *J. Eur. Math. Soc. (JEMS)* 16.9, pp. 1915–1935. MR: 3273312 (cit. on p. 2146).

Artur Avila, Svetlana Jitomirskaya, and Qi Zhou (2018). "Second phase transition line". *Math. Ann.* 370.1-2, pp. 271–285. MR: 3747487 (cit. on p. 2145).

Artur Avila and Raphaël Krikorian (2006). "Reducibility or nonuniform hyperbolicity for quasiperiodic Schrödinger cocycles". *Ann. of Math. (2)* 164.3, pp. 911–940. MR: 2259248 (cit. on pp. 2131, 2139).

– (2015). "Monotonic cocycles". *Invent. Math.* 202.1, pp. 271–331. MR: 3402800 (cit. on p. 2139).

Artur Avila, Y. Last, M. Shamis, and Q. Zhou (n.d.). *On the abominable properties of the Almost Mathieu operator with well approximated frequencies*. In preparation (cit. on pp. 2143, 2146).

Artur Avila, J You, and Q Zhou (2016). "Dry Ten Martini problem in non-critical case". Preprint (cit. on p. 2141).

Artur Avila, Jiangong You, and Qi Zhou (2017). "Sharp phase transitions for the almost Mathieu operator". *Duke Math. J.* 166.14, pp. 2697–2718. MR: 3707287 (cit. on pp. 2144, 2145).

Joseph Avron and Barry Simon (1982). "Singular continuous spectrum for a class of almost periodic Jacobi matrices". *Bull. Amer. Math. Soc. (N.S.)* 6.1, pp. 81–85. MR: 634437 (cit. on p. 2144).

Mark Ya Azbel (1964). "Energy spectrum of a conduction electron in a magnetic field". *Sov. Phys. JETP* 19.3, pp. 634–645 (cit. on p. 2140).

Jairo Bochi (2002). "Genericity of zero Lyapunov exponents". *Ergodic Theory Dynam. Systems* 22.6, pp. 1667–1696. MR: 1944399 (cit. on p. 2146).

J. Bourgain (2002). "On the spectrum of lattice Schrödinger operators with deterministic potential. II". *J. Anal. Math.* 88. Dedicated to the memory of Tom Wolff, pp. 221–254. MR: 1984594 (cit. on p. 2137).

J. Bourgain (2005a). *Green's function estimates for lattice Schrödinger operators and applications*. Vol. 158. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, pp. x+173. MR: 2100420 (cit. on p. 2146).

– (2005b). "Positivity and continuity of the Lyapounov exponent for shifts on $\mathbb{T}^d$ with arbitrary frequency vector and real analytic potential". *J. Anal. Math.* 96, pp. 313–355. MR: 2177191 (cit. on pp. 2146, 2147).

J. Bourgain and M. Goldstein (2000). "On nonperturbative localization with quasi-periodic potential". *Ann. of Math. (2)* 152.3, pp. 835–879. MR: 1815703 (cit. on pp. 2143, 2145, 2147).

J. Bourgain and Svetlana Jitomirskaya (2002a). "Absolutely continuous spectrum for 1D quasiperiodic operators". *Invent. Math.* 148.3, pp. 453–463. MR: 1908056 (cit. on p. 2138).

– (2002b). "Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential". *J. Statist. Phys.* 108.5-6. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays, pp. 1203–1218. MR: 1933451 (cit. on pp. 2146, 2147).

Ao Cai, Claire Chavaudret, Jiangong You, and Qi Zhou (June 2017). "Sharp Holder continuity of the Lyapunov exponent of finitely differentiable quasi-periodic cocycles". arXiv: 1706.08649 (cit. on pp. 2137, 2147).

Claire Chavaudret (2013). "Strong almost reducibility for analytic and Gevrey quasi-periodic cocycles". *Bull. Soc. Math. France* 141.1, pp. 47–106. MR: 3031673 (cit. on pp. 2137, 2138).

Man Duen Choi, George A. Elliott, and Noriko Yui (1990). "Gauss polynomials and the rotation algebra". *Invent. Math.* 99.2, pp. 225–246. MR: 1031901 (cit. on p. 2141).

David Damanik (2017). "Schrödinger operators with dynamically defined potentials". *Ergodic Theory Dynam. Systems* 37.6, pp. 1681–1764. MR: 3681983 (cit. on p. 2140).

David Damanik, M Goldstein, W Schlag, and M Voda (2015). "Homogeneity of the spectrum for quasi-periodic Schrödinger operators". To appear in J. Eur. Math. Soc. (cit. on p. 2143).

David Damanik and Michael Goldstein (2014). "On the inverse spectral problem for the quasi-periodic Schrödinger equation". *Publ. Math. Inst. Hautes Études Sci.* 119, pp. 217–401. MR: 3210179 (cit. on p. 2142).

David Damanik and Peter Yuditskii (2016). "Counterexamples to the Kotani-Last conjecture for continuum Schrödinger operators via character-automorphic Hardy spaces". *Adv. Math.* 293, pp. 738–781. MR: 3474334 (cit. on p. 2146).

E. I. Dinaburg and Ya. G. Sinaĭ (1975). "The one-dimensional Schrödinger equation with quasiperiodic potential". *Funkcional. Anal. i Priložen.* 9.4, pp. 8–21. MR: 0470318 (cit. on pp. 2137, 2145).

Pedro Duarte and Silvius Klein (n.d.). “Continuity, positivity and simplicity of the Lyapunov exponents for quasi-periodic cocycles”. To appear in *J. Eur. Math. Soc.* (cit. on p. 2147).

– (2014). “Positive Lyapunov exponents for higher dimensional quasiperiodic cocycles”. *Comm. Math. Phys.* 332.1, pp. 189–219. MR: 3253702 (cit. on p. 2147).

L. H. Eliasson (1992). “Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation”. *Comm. Math. Phys.* 146.3, pp. 447–482. MR: 1167299 (cit. on pp. 2137, 2140, 2145).

– (1997). “Discrete one-dimensional quasi-periodic Schrödinger operators with pure point spectrum”. *Acta Math.* 179.2, pp. 153–196. MR: 1607554 (cit. on p. 2145).

Bassam Fayad and Raphaël Krikorian (2009). “Rigidity results for quasiperiodic $\mathrm{SL}(2,\mathbb{R})$-cocycles”. *J. Mod. Dyn.* 3.4, pp. 497–510. MR: 2587083 (cit. on p. 2138).

Alex Furman (1997). “On the multiplicative ergodic theorem for uniquely ergodic systems”. *Ann. Inst. H. Poincaré Probab. Statist.* 33.6, pp. 797–815. MR: 1484541 (cit. on p. 2146).

D. J. Gilbert and D. B. Pearson (1987). “On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operators”. *J. Math. Anal. Appl.* 128.1, pp. 30–56. MR: 915965 (cit. on p. 2145).

Michael Goldstein and Wilhelm Schlag (2001). “Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions”. *Ann. of Math. (2)* 154.1, pp. 155–203. MR: 1847592 (cit. on pp. 2146, 2147).

– (2008). “Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues”. *Geom. Funct. Anal.* 18.3, pp. 755–869. MR: 2438997 (cit. on p. 2146).

– (2011). “On resonances and the formation of gaps in the spectrum of quasi-periodic Schrödinger equations”. *Ann. of Math. (2)* 173.1, pp. 337–475. MR: 2753606 (cit. on p. 2140).

Michael Goldstein, Wilhelm Schlag, and Mircea Voda (Aug. 2017). “On the Spectrum of Multi-Frequency Quasiperiodic Schrödinger Operators with Large Coupling”. arXiv: 1708.09711 (cit. on p. 2140).

A. Ya. Gordon (1976). “The point spectrum of the one-dimensional Schrödinger operator”. *Uspehi Mat. Nauk* 31.4(190), pp. 257–258. MR: 0458247 (cit. on p. 2144).

A. Ya. Gordon, Svetlana Jitomirskaya, Y. Last, and B. Simon (1997). “Duality and singular continuous spectrum in the almost Mathieu equation”. *Acta Math.* 178.2, pp. 169–183. MR: 1459260 (cit. on pp. 2138, 2144).

Sana Hadj Amor (2009). “Hölder continuity of the rotation number for quasi-periodic cocycles in $\mathrm{SL}(2,\mathbb{R})$”. *Comm. Math. Phys.* 287.2, pp. 565–588. MR: 2481750 (cit. on p. 2147).

Rui Han and Chris A. Marx (2018). “Large Coupling Asymptotics for the Lyapunov Exponent of Quasi-Periodic Schrödinger Operators with Analytic Potentials”. *Ann. Henri Poincaré* 19.1, pp. 249–265. arXiv: 1612.04321. MR: 3743760 (cit. on p. 2148).

Alex Haro and Joaquim Puig (2013). “A Thouless formula and Aubry duality for long-range Schrödinger skew-products”. *Nonlinearity* 26.5, pp. 1163–1187. MR: 3043377 (cit. on p. 2147).

Michael-R. Herman (1983). “Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d’un théorème d’Arnold et de Moser sur le tore de dimension 2”. *Comment. Math. Helv.* 58.3, pp. 453–502. MR: 727713 (cit. on pp. 2134, 2147).

Douglas R Hofstadter (1976). “Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields”. *Physical review B* 14.6, pp. 2239–2249 (cit. on p. 2141).

X. Hou, Y. Shan, and J. You (n.d.). *Explict construction of quasi-periodic Schrödinger operators with Cantor spectrum*. In preparation (cit. on p. 2141).

Xuanji Hou and Jiangong You (2012). “Almost reducibility and non-perturbative reducibility of quasi-periodic linear systems”. *Invent. Math.* 190.1, pp. 209–260. MR: 2969277 (cit. on p. 2138).

Svetlana Jitomirskaya (1995). “Almost everything about the almost Mathieu operator. II”. In: *XIth International Congress of Mathematical Physics (Paris, 1994)*. Int. Press, Cambridge, MA, pp. 373–382. MR: 1370694 (cit. on p. 2144).

– (1999). “Metal-insulator transition for the almost Mathieu operator”. *Ann. of Math. (2)* 150.3, pp. 1159–1175. MR: 1740982 (cit. on pp. 2143, 2144).

– (2007). “Ergodic Schrödinger operators (on one foot)”. In: *Spectral theory and mathematical physics: a Festschrift in honor of Barry Simon’s 60th birthday*. Vol. 76. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 613–647. MR: 2307750 (cit. on p. 2140).

Svetlana Jitomirskaya and Ilya Kachkovskiy (2016). “L^2-reducibility and localization for quasiperiodic operators”. *Math. Res. Lett.* 23.2, pp. 431–444. MR: 3512893 (cit. on p. 2144).

Svetlana Jitomirskaya, D. A. Koslover, and M. S. Schulteis (2009). “Continuity of the Lyapunov exponent for analytic quasiperiodic cocycles”. *Ergodic Theory Dynam. Systems* 29.6, pp. 1881–1905. MR: 2563096 (cit. on p. 2146).

Svetlana Jitomirskaya and Wencai Liu (n.d.). “Universal hierarchical structure of quasiperiodic eigenfunctions”. To appear in Ann. Math. arXiv: 1609.08664 (cit. on p. 2144).

Svetlana Jitomirskaya and Shiwen Zhang (Oct. 2015). "Quantitative continuity of singular continuous spectral measures and arithmetic criteria for quasiperiodic Schrödinger operators". arXiv: 1510.07086 (cit. on p. 2144).

R. Johnson and J. Moser (1982). "The rotation number for almost periodic potentials". *Comm. Math. Phys.* 84.3, pp. 403–438. MR: 667409 (cit. on pp. 2132, 2134).

M. Kac (n.d.). public commun, at 1981 AMS Annual Meeting (cit. on pp. 2140, 2141).

S. Kotani and M. Krishna (1988). "Almost periodicity of some random potentials". *J. Funct. Anal.* 78.2, pp. 390–405. MR: 943504 (cit. on p. 2145).

Shinichi Kotani (1984). "Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators". In: *Stochastic analysis (Katata/Kyoto, 1982)*. Vol. 32. North-Holland Math. Library. North-Holland, Amsterdam, pp. 225–247. MR: 780760 (cit. on pp. 2135, 2139).

Raphaël Krikorian (1999a). "Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts". *Astérisque* 259, pp. vi+216. MR: 1732061 (cit. on p. 2137).

– (1999b). "Réductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts". *Ann. Sci. École Norm. Sup. (4)* 32.2, pp. 187–240. MR: 1681809 (cit. on p. 2137).

– (Feb. 2004). "Reducibility, differentiable rigidity and Lyapunov exponents for quasi-periodic cocycles on $T \times SL(2, R)$". arXiv: math/0402333 (cit. on p. 2139).

Yoram Last (2005). "Spectral theory of Sturm-Liouville operators on infinite intervals: a review of recent developments". In: *Sturm-Liouville theory*. Birkhäuser, Basel, pp. 99–120. MR: 2145079 (cit. on pp. 2131, 2140).

Martin Leguil, Jiangong You, Zhiyan Zhao, and Qi Zhou (2017). "Asymptotics of spectral gaps of quasi-periodic Schrödinger operators". arXiv: 1712.04700 (cit. on pp. 2138, 2142, 2143, 2147).

C. A. Marx and Svetlana Jitomirskaya (2017). "Dynamics and spectral theory of quasi-periodic Schrödinger-type operators". *Ergodic Theory Dynam. Systems* 37.8, pp. 2353–2393. MR: 3719264 (cit. on p. 2140).

Jürgen Moser and Jürgen Pöschel (1984). "An extension of a result by Dinaburg and Sinaĭ on quasiperiodic potentials". *Comment. Math. Helv.* 59.1, pp. 39–85. MR: 743943 (cit. on pp. 2137, 2140).

D. Osadchy and J. E. Avron (2001). "Hofstadter butterfly as quantum phase diagram". *J. Math. Phys.* 42.12, pp. 5665–5671. MR: 1866679 (cit. on p. 2141).

Joaquim Puig (2004). "Cantor spectrum for the almost Mathieu operator". *Comm. Math. Phys.* 244.2, pp. 297–309. MR: 2031032 (cit. on pp. 2131, 2141).

– (2006). "A nonperturbative Eliasson's reducibility theorem". *Nonlinearity* 19.2, pp. 355–376. MR: 2199393 (cit. on p. 2138).

Barry Simon (1982). "Almost periodic Schrödinger operators: a review". *Adv. in Appl. Math.* 3.4, pp. 463–490. MR: 682631 (cit. on pp. 2131, 2140, 2141).

Barry Simon (2000). "Schrödinger operators in the twenty-first century". In: *Mathematical physics 2000*. Imp. Coll. Press, London, pp. 283–288. MR: 1773049 (cit. on p. 2140).

Ya. G. Sinaĭ (1987). "Anderson localization for one-dimensional difference Schrödinger operator with quasiperiodic potential". *J. Statist. Phys.* 46.5-6, pp. 861–909. MR: 893122 (cit. on p. 2141).

Mikhail Sodin and Peter Yuditskii (1995). "Almost periodic Sturm-Liouville operators with Cantor homogeneous spectrum". *Comment. Math. Helv.* 70.4, pp. 639–658. MR: 1360607 (cit. on p. 2142).

– (1997). "Almost periodic Jacobi matrices with homogeneous spectrum, infinite-dimensional Jacobi inversion, and Hardy spaces of character-automorphic functions". *J. Geom. Anal.* 7.3, pp. 387–435. MR: 1674798 (cit. on p. 2142).

Eugene Sorets and Thomas Spencer (1991). "Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials". *Comm. Math. Phys.* 142.3, pp. 543–566. MR: 1138050 (cit. on p. 2147).

A. Volberg and P. Yuditskii (2014). "Kotani-Last problem and Hardy spaces on surfaces of Widom type". *Invent. Math.* 197.3, pp. 683–740. MR: 3251833 (cit. on p. 2145).

Jing Wang, Qi Zhou, and Tobias Jäger (July 2016). "Genericity of mode-locking for quasiperiodically forced circle maps". arXiv: 1607.01700 (cit. on p. 2140).

Yiqian Wang and Jiangong You (2013). "Examples of discontinuity of Lyapunov exponent in smooth quasiperiodic cocycles". *Duke Math. J.* 162.13, pp. 2363–2412. MR: 3127804 (cit. on p. 2146).

– (Jan. 2015). "Quasi-Periodic Schrödinger Cocycles with Positive Lyapunov Exponent are not Open in the Smooth Topology". arXiv: 1501.05380 (cit. on p. 2146).

Yiqian Wang and Zhenghe Zhang (2017). "Cantor spectrum for a class of C^2 quasiperiodic Schrödinger operators". *Int. Math. Res. Not. IMRN* 8, pp. 2300–2336. MR: 3658199 (cit. on p. 2141).

Jiangong You and Shiwen Zhang (2014). "Hölder continuity of the Lyapunov exponent for analytic quasiperiodic Schrödinger cocycle with weak Liouville frequency". *Ergodic Theory Dynam. Systems* 34.4, pp. 1395–1408. MR: 3227161 (cit. on p. 2146).

Jiangong You and Qi Zhou (2013). "Embedding of analytic quasi-periodic cocycles into analytic quasi-periodic linear systems and its applications". *Comm. Math. Phys.* 323.3, pp. 975–1005. MR: 3106500 (cit. on pp. 2135, 2138, 2144).

– (2015). "Simple counter-examples to Kotani-Last conjecture via reducibility". *Int. Math. Res. Not. IMRN* 19, pp. 9450–9455. MR: 3431598 (cit. on p. 2146).

Zhenghe Zhang (2012). "Positive Lyapunov exponents for quasiperiodic Szegő cocycles". *Nonlinearity* 25.6, pp. 1771–1797. MR: 2929602 (cit. on p. 2147).

Received 2017-11-27.

JIANGONG YOU (尤建功)

Chern Institute of Mathematics and LPMC
Nankai University
Tianjin 300071
China
jyou@nankai.edu.cn

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2155–2184)

THE ORR MECHANISM: STABILITY/INSTABILITY OF THE COUETTE FLOW FOR THE 2D EULER DYNAMIC

Jacob Bedrossian, Yu Deng and Nader Masmoudi

Abstract

We review our works on the nonlinear asymptotic stability and instability of the Couette flow for the 2D incompressible Euler dynamic. In the fits part of the work we prove that perturbations to the Couette flow which are small in Gevrey spaces G^s of class $1/s$ with $s > 1/2$ converge strongly in L^2 to a shear flow which is close to the Couette flow. Moreover in a well chosen coordinate system, the solution converges in the same Gevrey space to some limit profile. In a later work, we proved the existence of small perturbations in G^s with $s < 1/2$ such that the solution becomes large in Sobolev regularity and hence yields instability. In this note we discuss the most important physical and mathematical aspects of these two results and the key ideas of the proofs.

Contents

MSC2010: primary 35Q31; secondary 35B35.

1 Introduction and physical background

The theory of hydrodynamic stability at high Reynolds number started already in the 19th century, with the likes of Stokes, Reynolds, Kelvin, Orr and others. Some of the first early theoretical works were done by Rayleigh [1879/80, 1887/88, 1895/96], including for example, the inflection point theorem of the spectral instability on inviscid planar shear flows, and the exact solutions for Coutte flow in the absence of boundaries constructed by Kelvin [1887] which showed linear stability independent of Reynolds number. These solutions were later followed up by Orr [1907], Dikiĭ [1961], and Case [1960] to show linearized stability of the inviscid Couette flow also in a channel. The early experiments of Reynolds [1883] clearly showed instability at all high Reynolds number for flow in a pipe, which although that linearized problem still to this day has not been conclusively solved, seemed slightly in contradiction with most of the theoretical results of the time. Kelvin proposed the solution to this 'paradox' (sometimes called the 'Sommerfeld paradox' Li and Lin [2011]): although the fluid might be stable at all Reynolds numbers, as the Reynolds number increases, the fluid becomes increasing sensitive to small perturbations. This phenomena now often called *subcritical transition* and it is ubiquitous in 3D fluid mechanics (as well as plasma physics). It has also been observed in cases with spectral instability at high Reynolds number in the sense that the flows often go unstable at lower Reynolds number than that predicted by linearized theory or in a way completely unrelated to the unstable eigenvalues; see e.g. Drazin and Reid [1982], Schmid and Henningson [2001], and Yaglom [2012] and the references therein for discussions on this effect in the physics literature. In the case of linearized stability at all Reynolds number, we could then phrase the following question: given a norm $\|\cdot\|_X$, and an initial perturbation $\tilde{u}$, we could look for a $\gamma = \gamma(X)$ such that:

(1a) $\|\tilde{u}\|_X \lesssim \mathbf{Re}^{-\gamma} \Rightarrow$ stability and linear-dominated behavior $\forall t$,

(1b) $\|\tilde{u}\|_X \gg \mathbf{Re}^{-\gamma} \Rightarrow$ (potential) nonlinear-dominated behavior and instability.

The exponent γ is sometimes called the *transition threshold*. More recent insights show that there are two interesting aspects of this effect:

- a/ The linearized problem may contain transient growth, and these could be triggering nonlinear instabilities, especially if the growth becomes larger as Reynolds number goes to infinity.
- b/ The dependence on on the stability threshold may depend strongly on the topology in which one measures perturbations. It is this aspect that we will be reviewing here.

Modern research shows several fundamental differences between 2D and 3D hydrodynamic stability, both at the linear and nonlinear level. This is due to the fact that (A) the 3D linearized equations have a wider range of more extreme transient growth mechanisms than 2D and (B) 3D has a much more complicated structure of 'resonances', that is, the weakly nonlinear structure is much more complicated (see below for more discussion). Unlike what might be suggested by Squire's theorem Squire [1933], as a result, 2D stability studies do not seem to give significant physical insight into 3D fluids in the sense that theoretical or numerical results on 2D hydrodynamic stability give no specific, direct information on 3D flows. However, many important physical applications are well-approximated to leading order by 2D fluids, such as many atmospheric and oceanic phenomena, so it is still important to give careful consideration to 2D fluids. Moreover, when it comes to hydrodynamic stability questions, the dynamics of 2D fluids are significantly simpler than 3D fluids in many ways, and hence it is reasonable to begin mathematical studies with the former rather than the latter: a theorem about 2D flow in a channel does not give much specific physical insight into 3D flow in a pipe, but the mathematics developed therein hopefully will. Indeed, this was clear in the works Bedrossian and Masmoudi [2013] and Bedrossian, Masmoudi, and Vicol [2016] vs Bedrossian, Germain, and Masmoudi [2015a,b, 2017a]: the dynamics and nonlinear structures might be very different in 2D and 3D, but nevertheless, the subsequent proofs all used certain mathematical tools originally designed for the stability of 2D Euler in Bedrossian and Masmoudi [2013], or at least used ideas heavily influenced by the insight obtained therein.

With the above discussions in mind, we will focus in this paper on the 2D case, and even mostly on the simpler case of infinite Reynolds number, e.g. the 2D incompressible Euler equations (we will see that the inviscid problem is a reasonable place to start, even though (1) is phrased in terms of Reynolds number). See Bedrossian, Germain, and Masmoudi [2017b] for a review of the related 3D stability problems. Moreover, we are interested in nonlinear questions, and progress has mostly only been made on one shear flow: the Couette flow $u = (y, 0)$ on $(x, y) \in \mathbb{T} \times \mathbb{R}$. In this case, the 2D Euler system in the vorticity formulation with the background shear flow becomes:

$$\begin{cases} \omega_t + y\partial_x\omega + U \cdot \nabla\omega = 0, \\ U = \nabla^{\perp}(\Delta)^{-1}\omega, \qquad \omega(t=0) = \omega_{in}. \end{cases} \tag{2}$$

Here, $(x, y) \in \mathbb{T} \times \mathbb{R}$, $\nabla^{\perp} = (-\partial_y, \partial_x)$ and (U, ω) are periodic in the x variable with period normalized to 2π. The physical velocity is $(y, 0) + U$ where $U = (U^x, U^y)$ denotes the velocity perturbation and the total vorticity is $-1 + \omega$. In what follows, we define the streamfunction $\psi = \Delta^{-1}\omega$.

We first state the result of Bedrossian and Masmoudi [2013] and then attempt to elucidate some of the interesting physical and mathematical concepts which are involved in the proof. We then state the instability result of Deng and Masmoudi [2018] which shows the criticality of the $G^{1/2}$ space, namely the Gevrey space of class 2. It is worth noting that this space appears here due to a nonlinear mechanism, not a linear one. The relationship with Landau damping in the Vlasov equations of plasma physics and the recent work of Mouhot and Villani [2011] will also be discussed.

2 Linear dynamics

Linearizing the 2D Euler equations as written in (2) just means dropping the quadratic term:

$$\begin{cases} \partial_t \omega + y\partial_x \omega = 0 \\ \Delta\phi = \omega. \end{cases} \tag{3}$$

This is readily solved:

$$\begin{aligned} \omega(t,x,y) &= \omega_{in}(x-ty,y) \\ \widehat{\omega}(t,k,\eta) &= \widehat{\omega_{in}}(k,\eta+kt). \end{aligned} \tag{4}$$

From (4) we can see a linear-in-time transfer of enstrophy to high frequencies. Since Δ^{-1} gains two derivatives, we should intuitively guess that $P_{k\neq 0}\phi$ decays like $\langle t\rangle^{-2}$ in L^2. This might seem circuitous for solving the linear problem, but let us follow Kelvin and Orr and introduce the following change of variables:

$$z = x - ty \tag{5a}$$

$$f(t,z,y) = \omega(t,z+ty,y) = \omega(t,x,y) \tag{5b}$$

$$\phi(t,z,y) = \psi(t,z+ty,y) = \psi(t,x,y). \tag{5c}$$

This is nothing more than rewinding by the linear propagator associated to the Couette flow. From (3) we have

$$\partial_t f = 0 \tag{6a}$$

$$\partial_{zz}\phi + (\partial_y - t\partial_x)^2\phi = f, \tag{6b}$$

which gives:

$$\hat{\phi}(t,k,\eta) = -\frac{\hat{f}(t,k,\eta)}{k^2+|\eta-kt|^2} = -\frac{\hat{\omega}_{in}(k,\eta)}{k^2+|\eta-kt|^2}. \tag{7}$$

From (7) we derive the fundamental decay-by-mixing estimate: for any $\sigma \in [0,\infty)$ and $\beta \in [0,2]$,

$$\|P_{\neq 0}\phi\|_{H^\sigma} \lesssim \frac{1}{\langle t\rangle^\beta}\|f\|_{H^{\sigma+\beta}} = \frac{1}{\langle t\rangle^\beta}\|\omega_{in}\|_{H^{\sigma+\beta}}, \tag{8}$$

where we are using H^σ to denote the L^2 Sobolev norm of order σ. Due to

$$U^x(t,x,y) = -\partial_y \psi(t,x,y) = -\partial_y\left(\phi(t,x-ty,y)\right) = \left((\partial_y - t\partial_x)\phi\right)(t,x-ty,y)$$
$$U^y(t,x,y) = \partial_x \psi(t,x,y) = \partial_x\left(\phi(t,x-ty,y)\right) = (\partial_x\phi)(t,x-ty,y),$$

we get the inviscid damping predicted by Orr [1907]:

$$\|P_{\neq 0}U^x\|_{L^2} \lesssim \langle t\rangle \|\nabla\phi\|_{L^2} \lesssim \langle t\rangle^{-1}\|\omega_{in}\|_{H^3} \tag{9a}$$

$$\|U^y\|_{L^2} \lesssim \|\partial_x\phi\|_{L^2} \lesssim \langle t\rangle^{-2}\|\omega_{in}\|_{H^3}. \tag{9b}$$

This shows that on the linear level, we have the convergence $(y+U^x, U^y) \to (y + < U^x_{in} > (y), 0)$ in L^2 as time goes to infinity. Hence, the velocity field converges strongly back to a shear flow but not back to the Couette flow. As discussed in the introduction, Orr had a second observation from (7), which is that modes with $\eta k > 0$ undergo first a transient growth in ϕ before decaying. Indeed, the denominator (7) is minimal the time $t = \frac{\eta}{k}$, which Orr called the *critical time*. Therefore, if $|\eta| >> |k|$, the velocity field is *amplified* by a large factor between $t = 0$ and the critical time $t = \frac{\eta}{k}$. In physical terms, this transient growth is due to the fact that the mode of the vorticity in question is initially well-mixed, and then proceeds to *unmix* under the Couette flow evolution. See figure (1) for how this mixing/un-mixing effect appears on each Fourier mode of the vorticity. The relevance of the Orr mechanism to hydrodynamic stability has been debated over the years; see e.g. Orr [1907], Boyd [1983], and Lindzen [1988] and Yaglom [2012] for a detailed account of how the literature on the topic developed. However, our work verifies the crucial importance of the Orr mechanism for nonlinear stability problems at high Reynolds numbers in 2D fluid mechanics, or at least for the Couette flow.

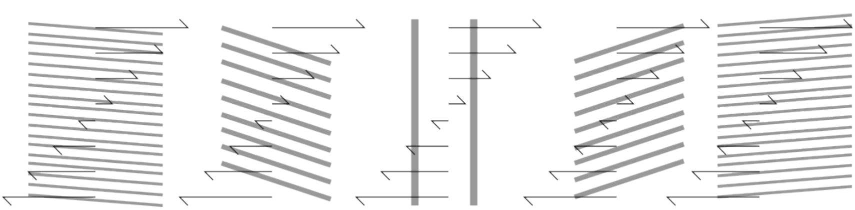


Figure 1: A mode-by-mode visualization of the Orr mechanism: the arrows represent the background flow, while the stripes are the level sets of the function $e^{ik(x-ty)+i\eta y}$ with $\eta/k \gg 1$. Time increases from left to right, and the center image is the critical time $t = \eta/k$. The full linearized solution is simply a superposition of these tilting waves.

3 Overview of the mathematical results

In this section we will summarize the results of Bedrossian and Masmoudi [2013] and Deng and Masmoudi [2018] in the 2D Euler equations. Each has analogies in the Vlasov-Poisson equations of kinetic theory. The original work of Mouhot and Villani proved the nonlinear Landau damping in $\mathbb{T}^d \times \mathbb{R}^d$, as predicted by the linearized Vlasov (see also Bedrossian, Masmoudi, and Mouhot [2016b]). These results are the analogue of the positive stability results of Bedrossian and Masmoudi [2013] (Theorem 1 below), though broadly speaking, Theorem 1 seems to require a much more subtle proof for several reasons. For the instability results, something related to Theorem 2 below was proved previously for Vlasov-Poisson in Bedrossian [2016] in Sobolev spaces; both the proof and the nature of the nonlinear behavior demonstrated by the solutions are different, but closely related.

3.1 Stability result. The main result of Bedrossian and Masmoudi [2013] is the following, which shows that if one is willing to take Gevrey-1/2 regularity on the initial data, the *nonlinear* 2D Euler equations display an inviscid damping essentially the same as that predicted by Orr.

Theorem 1. *For all $1/2 < s \leq 1$, $\lambda_0 > \lambda' > 0$ there exists an $\epsilon_0 = \epsilon_0(\lambda_0, \lambda', s) \leq 1/2$ such that for all $\epsilon \leq \epsilon_0$ if ω_{in} satisfies*

$$\int_{\mathbb{T}\times\mathbb{R}} (1+|y|)\cdot|\omega_{in}(x,y)|\,\mathrm{d}x\mathrm{d}y \leq \varepsilon, \quad \int_{\mathbb{T}\times\mathbb{R}} \omega_{in}(x,y)\,\mathrm{d}x\mathrm{d}y = 0 \quad \textit{and} \tag{10}$$

$$\|\omega_{in}\|_{\lambda_0}^2 = \sum_k \int |\hat{\omega}_{in}(k,\eta)|^2\, e^{2\lambda_0|k,\eta|^s}\, d\eta \leq \epsilon^2,$$

then there exists f_∞ with $\int f_\infty dxdy = 0$ and $\|f_\infty\|_{\lambda'} \lesssim \epsilon$ such that

$$\|\omega(t, x+ty+\Phi(t,y), y) - f_\infty(x,y)\|_{\lambda'} \lesssim \frac{\epsilon^2}{\langle t\rangle}, \tag{11}$$

where $\Phi(t,y)$ is given explicitly by

$$\Phi(t,y) = \frac{1}{2\pi}\int_0^t \int_{\mathbb{T}} U^x(s,x,y)dxds = u_\infty(y)t + \theta(t,y), \tag{12}$$

with $u_\infty = \partial_y \partial_{yy}^{-1} \frac{1}{2\pi} \int_{\mathbb{T}} f_\infty(x,y)dx$ *and* $|\theta(t,y)| \lesssim \epsilon^2$*. Moreover, the velocity field* U *decays as*

$$\left\| \int U^x(t,x,\cdot)dx - u_\infty \right\|_{\lambda'} \lesssim \frac{\epsilon^2}{\langle t \rangle}, \tag{13a}$$

$$\left\| U^x(t) - \int U^x(t,x,\cdot)dx \right\|_{L^2} \lesssim \frac{\epsilon}{\langle t \rangle}, \tag{13b}$$

$$\|U^y(t)\|_{L^2} \lesssim \frac{\epsilon}{\langle t \rangle^2}. \tag{13c}$$

The above result was extended to a uniform-in-Reynolds number statement about the 2D Navier-Stokes equations in Bedrossian, Masmoudi, and Vicol [2016] (note, that such a statement is always strictly harder than an infinite Reynolds number result). This work also shows that the mixing due to the Couette flow enhances the effect of the viscosity, an effect which plays an important role in 3D as well Bedrossian, Germain, and Masmoudi [2015a,b, 2017a]. These papers show that for 2D Couette flow, in Gevrey-2 regularity, there is *no subcritical transition*. Note that in 3D, there *is* subcritical transition, even in Gevrey-2 Bedrossian, Germain, and Masmoudi [2015b].

3.2 Instability result. It is known that the dynamics of Theorem 1 may not happen in low regularities. In Lin and Zeng [2011a], time periodic solutions to Equation (2) are constructed in Sobolev spaces H^s where $s < 3/2$; for Vlasov-Poisson equations, the same result was proved in Lin and Zeng [2011b] and, as mentioned before, Bedrossian [2016] has proved instability in any Sobolev space H^s.

The main result of Deng and Masmoudi [2018] fills the gap between these stability and instability results, by proving instability of the $2D$ Couette flow in any Gevrey-s regularity for $s < 1/2$. In fact something slightly stronger is proved, where Gevrey-s is replaced by a log-corrected version of Gevrey-$1/2$.

Theorem 2. *Let* $N_0 = 9000$*,* $N_1 = 60000$*, and denote* $\log^+(x) = \log(2+|x|)$*. For a function* $f : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$*, define the Gevrey-type norm* $\mathcal{G}^*$ *by*

$$\|F\|_{\mathcal{G}^*}^2 = \sum_{k\in\mathbb{Z}} \int_{\mathbb{R}} e^{2\kappa(k,\xi)} |\widehat{F}(k,\xi)|^2 \, d\xi, \quad \kappa(k,\xi) = \frac{(|k|+|\xi|)^{1/2}}{(\log^+(|k|+|\xi|))^{N_1}}. \tag{14}$$

Then, for any sufficiently small $\varepsilon > 0$*, there exists a solution* $\omega = \omega(t,x,y)$ *to* (2)*, such that:*

1. *The initial data* $\omega(0)$ *satisfies the assumptions (10), and that*

$$\|\omega(0)\|_{\mathcal{G}^*} \le \varepsilon; \tag{15}$$

2. *At some later time T, the solution ω satisfies that*

$$\|\langle\partial_x\rangle^{N_0}\omega(T,x,y)\|_{L^2(\mathbb{T}\times\mathbb{R})} \geq \frac{1}{\varepsilon}. \tag{16}$$

4 The nonlinear dynamics: The toy model

In both of the theorems above, we are interested in understanding the *weakly nonlinear* effects. It is a classical idea that transient growth in a linear problem can interact badly with the nonlinearity to trigger instability, for instance see the discussion in L. N. Trefethen, A. E. Trefethen, Reddy, and T. A. Driscoll [1993]. The basic mechanism is as follows. Heuristically, in the weakly nonlinear regime we can imagine the solution as an interacting superposition of waves undergoing linear shear. Through the nonlinear term, each mode has a strong effect at its critical time during which it strongly forces the others, potentially putting information into modes which have not yet reached their critical time and are hence still growing. At a later time, these modes have a large effect and continue to excite other growing modes and so forth, perpetuating a so-called self-sustaining 'nonlinear bootstrap' (see L. N. Trefethen, A. E. Trefethen, Reddy, and T. A. Driscoll [1993], Baggett, T. A. Driscoll, and L. N. Trefethen [1995], Vanneste, Morrison, and Warn [1998], and Vanneste [2001/02] and the references therein for discussions in the fluid mechanics context). Since the measurable effect of a nonlinear interaction can occur long after the event, this mechanism permits *nonlinear echoes*, in which the electric field of the plasma, or kinetic energy of the fluid disturbance, is highly concentrated at specific times. These spectacular displays of reversibility were captured experimentally for Vlasov, there known as *plasma echoes*, in the work of Malmberg, Wharton, Gould, and O'Neil [1968]. The analogous 'Euler echoes' were recently studied and observed both numerically Vanneste, Morrison, and Warn [1998] and Vanneste [2001/02] and experimentally Yu and C. Driscoll [2002] and Yu, C. Driscoll, and O'Neil [2005].

The careful analysis of plasma echoes in the Vlasov equations is crucial in the proof of Mouhot and Villani [2011] (and also in Bedrossian, Masmoudi, and Mouhot [2016b,a]), as these are the dominant weakly nonlinear effect that could lead to instability. Later, the work of Bedrossian [2016] confirmed this viewpoint by constructing arbitrarily small perturbations in H^s (for any finite s) which give rise to arbitrarily many distinct nonlinear oscillations in the electric field (similar to the experiments of Malmberg, Wharton, Gould, and O'Neil [1968] but with arbitrarily long chains of echoes). This resonance is also used to prove Theorem 2.

Let us try to begin the analysis in the natural way, by first making the change $f(t,z,y) = \omega(t, z+ty, y)$. The nonlinear Euler equations then become (note the gratuitious cancellation)

$$\partial_t f + \nabla^{\perp}\phi \cdot \nabla f = 0 \tag{17a}$$

$$\partial_{zz}\phi + (\partial_y - t\partial_x)^2\phi = f. \tag{17b}$$

There are actually two immediate problems. First of all, the contribution of the velocity field $\frac{1}{2\pi}\int_0^{2\pi} \partial_y \phi(t,z,y)dz$ will not decay, indeed, the linear problem leaves these modes invariant. Hence, we will unavoidably have $\|\partial_y^j f\| \approx \epsilon\langle t\rangle^j$ in general. This clearly shows that for long-time estimates, we are working in the wrong coordinate system.

Even if we ignore this problem, we have another problem. In order to obtain inviscid damping, the goal is simply to obtain uniform-in-time H^s estimates on f. Imagine we forget about the non-decaying modes and write (up to commutators that are not important here):

$$\begin{aligned}\frac{1}{2}\partial_t \|\langle\nabla\rangle^s f\|_2^2 &= \langle \langle\nabla\rangle^s f, \langle\nabla\rangle^s \left(P_{k\neq 0}\nabla^{\perp}\phi\cdot\nabla f\right)\rangle \approx \\ &\approx \langle \langle\nabla\rangle^s f, \langle\nabla\rangle^s P_{k\neq 0}\nabla^{\perp}\phi\cdot\nabla f\rangle + \langle \langle\nabla\rangle^s f, P_{k\neq 0}\,|\nabla|\,\nabla^{\perp}\phi\cdot\nabla\langle\nabla\rangle^{s-1} f\rangle + \cdots\end{aligned} \tag{18}$$

The second term we can pay regularity to get decay of the velocity field provided s is large enough, however, it is far from clear how to obtain any kind of decay on the first term. Indeed, getting decay from inviscid damping seems to always require more regularity than we have; and it seems to require so much that even if we are willing to work in analytic regularity via some kind of Cauchy-Kovalevskaya argument, it would still not be enough to close any estimates.

Let us now look closer. Since we must pay regularity to deduce decay on the velocity u, it is natural to consider the frequency interactions in the product $u\cdot\nabla f$ with the frequencies of u much larger than f. This leads us to study a simpler model

$$\partial_t f = -u\cdot\nabla f_{lo}, \tag{19}$$

where f_{lo} is a given function that we think of as much smoother than f. Let us just focus on what should be the worst:

$$\partial_t f = \partial_v P_{\neq 0}\phi\,\partial_z f_{lo}.$$

This problem is linear on the Fourier side:

$$\partial_t \hat{f}(t,k,\eta) = \frac{1}{2\pi}\sum_{l\neq 0}\int_{\xi} \frac{\xi(k-l)}{l^2+|\xi - lt|^2}\hat{f}(l,\xi)\hat{f}_{lo}(t,k-l,\eta-\xi)\,d\xi.$$

Since f_{lo} weakens interactions between well-separated frequencies, let us consider a discrete model with η as a fixed parameter:

$$\partial_t \hat{f}(t,k,\eta) = \frac{1}{2\pi} \sum_{l \neq 0} \frac{\eta(k-l)}{l^2 + |\eta - lt|^2} \hat{f}(l,\eta) f_{lo}(t,k-l,0). \tag{20}$$

As time advances this system of ODEs will go through resonances or "critical times" given by $t = \frac{\eta}{k}$, at which time the k mode strongly forces the others. If $|\eta|\, k^{-2} \ll 1$ then the critical time does not have a serious detriment. Henceforth only consider $|\eta|\, k^{-2} > 1$. The scenario we are most concerned with is a high-to-low cascade in which the k mode has a strong effect at time η/k that excites the $k-1$ mode which has a strong effect at time $\eta/(k-1)$ that excites the $k-2$ mode and so on. Now focus near one critical time η/k on a time interval of length roughly η/k^2, namely $I_k = [\eta/k - \eta/k^2, \eta/k + \eta/k^2]$ and consider the interaction between the mode k and a nearby mode l with $l \neq k$. If one takes absolute values and retains only the leading order terms, then this reduces to the much simpler system of two ODEs (thinking of $f_{lo} = O(\kappa)$) which we refer to as the *toy model*:

$$\partial_t f_R = \kappa \frac{k^2}{|\eta|} f_{NR}, \tag{21a}$$

$$\partial_t f_{NR} = \kappa \frac{|\eta|}{k^2 + |\eta - kt|^2} f_R, \tag{21b}$$

where we think of f_R as being the evolution of the k mode and f_{NR} being the evolution of a nearby mode l with $l \neq k$. The factor $k^2/|\eta|$ in the ODE for f_R is an upper bound on the strongest interaction a non-resonant mode, for example the $k-1$ mode, can have with the resonant mode. It is important to note that if at the beginning of the interval I_k, we have $f_R = f_{NR}$, then over the interval I_k, both f_R and f_{NR} are at most amplified by roughly the same factor $C\,(\frac{\eta}{k^2})^{1+2C\kappa}$ (though they crucially are not amplified by the same amount on the left and right parts of the interval). Taking the product of these amplifications for $k = E(\sqrt{\eta}), E(\sqrt{\eta}) - 1, ..., 1$ yields a total amplification which is $O(e^{C\sqrt{\eta}})$. This indicates that unless there is some special structure or cancellation not taken into account, the growth of high frequencies will cause a loss of Gevrey-2 regularity of the solution as $t \to \infty$. Therefore, in order to maintain control, the initial data must have at least this much regularity to lose, and this is the origin of the requirement $s > 1/2$ (or at least $s \geq 1/2$).

5 Proof of the stability result

Hence, we have two main challenges to overcome. The first is to choose a coordinate system that is properly adapted to the shear flow which is mixing the solution. Note that this shear flow is changing in time and cannot be determined directly from the initial data. We carry this out in Section 5.1 below. The next step is to get global-in-time, uniform regularity estimates on the resulting f. To do this we will design a special norm with which to measure the solution that accounts for the nonlinear Orr mechanism described above.

5.1 Coordinate transform. The original equations in vorticity form are (2), and we are trying essentially to prove that

$$\omega(t,x,y) \to f_\infty\left(x - ty - u_\infty(y)t, y\right),$$

as $t \to \infty$, where $u_\infty(y)$ is the correction to the shear flow determined by f_∞. From the initial data alone, there is no simple way to determine u_∞; it is chosen by the nonlinear evolution. In order to deal with this lack of information about how the final state evolves we choose a coordinate system which adapts to the solution and converges to the expected form as $t \to \infty$. The change of coordinates used is $(t,x,y) \to (t,z,v)$, where

$$z(t,x,y) = x - tv \tag{22a}$$

$$v(t,y) = y + \frac{1}{t}\int_0^t < U^x > (\tau, y)d\tau, \tag{22b}$$

where we recall $< w >$ denotes the average of w in the x variable (or equivalently in the z variable), namely $< w >= \frac{1}{2\pi}\int_{\mathbb{T}} w dx$. The reason for the change $y \to v$ is not immediately clear, however v is named as such since it is an approximation for the background shear flow. If the velocity field in the integrand were constant in time, then we are simply transforming the y variables so that the shear appears linear. It will turn out that this choice of v ensures that the Biot-Savart law is in a form amenable to Fourier analysis in the variables (z,v); in particular, even when the shear is time-varying we may still study the *Orr critical times*. In this light, the motivation for the shift in z is clear: as suggested by the discussion in Section 2, we are eliminating the contribution of $< U^x >$ and following the flow in the horizontal variable to guarantee compactness.

Define $f(t,z,v) = \omega(t,x,y)$ and $\phi(t,z,v) = \psi(t,x,y)$, hence

$$\partial_t\omega = \partial_t f + \partial_t z \partial_z f + \partial_t v \partial_v f, \qquad \partial_x \omega = \partial_z f, \qquad \partial_y \omega = \partial_y v\left(\partial_v f - t\partial_z f\right),$$

where

$$
\begin{aligned}
\partial_t z &= -y - < U^x > (t,y) \\
\partial_t v &= \frac{1}{t}\left[< U^x > (t,y) - \frac{1}{t}\int_0^t < U^x > (s,y)ds \right] \\
\partial_y v &= 1 - \frac{1}{t}\int_0^t < \omega > (s,y)ds \\
\partial_{yy} v &= -\frac{1}{t}\int_0^t \partial_y < \omega > (s,y)ds.
\end{aligned}
$$

Expressing $[\partial_t v](t,v) = \partial_t v(t,y)$, $v'(t,v) = \partial_y v(t,y)$ and $v''(t,v) = \partial_{yy} v(t,y)$, we get the following evolution equation for f,

$$
\partial_t f + [\partial_t v]\partial_v f + \partial_t z \partial_z f = -y\partial_z f + v'\left[\partial_v \phi + \partial_z\phi\partial_v z - \partial_z\phi\partial_v z\right]\partial_z f - v'\partial_z\phi\partial_v f.
$$

Using the definition of $\partial_t z$ and the Biot-Savart law to transform $< U^x >$ to $-v'\partial_v < \phi >$ in the new variables, this becomes

$$
\partial_t f - \left(v'\partial_v(\phi - < \phi >)\right)\partial_z f + \left([\partial_t v] + v'\partial_z\phi\right)\partial_v f = 0.
$$

The Biot-Savart law also gets transformed into:

$$
f = \partial_{zz}\phi + (v')^2\left(\partial_v - t\partial_z\right)^2\phi + v''\left(\partial_v - t\partial_z\right)\phi = \Delta_t\phi. \tag{23}
$$

The original 2D Euler system (2) is expressed as

$$
\begin{cases}
\partial_t f + u\cdot\nabla_{z,v} f = 0, \\
u = (0,[\partial_t v]) + v'\nabla^{\perp}_{z,v} P_{\neq 0}\phi, \\
\phi = \Delta_t^{-1}[f].
\end{cases} \tag{24}
$$

It what follows we will write $\nabla_{z,v} = \nabla$ and specify when other variables are used. Next we transform the momentum equation to allow us to express $[\partial_t v]$ in a form amenable to estimates. Denoting $\tilde{u}(t,z,v) = U^x(t,x,y)$ and $p(t,z,v) = P(t,x,y)$ we have by the same derivation on f,

$$
\partial_t\tilde{u} + [\partial_t v]\partial_v\tilde{u} + \partial_z P_{\neq 0}\phi + v'\nabla^{\perp}P_{\neq 0}\phi\cdot\nabla\tilde{u} = -\partial_z p.
$$

Taking averages in z we isolate the zero mode of the velocity field,

$$
\partial_t\tilde{u}_0 + [\partial_t v]\partial_v\tilde{u}_0 + v' < \nabla^{\perp}P_{\neq 0}\phi\cdot\nabla\tilde{u} >= 0. \tag{25}
$$

Finally, one can express v' and $[\partial_t v]$ as solutions to a system of PDE in the (t, v) variables coupled to (24) (see Bedrossian and Masmoudi [2013] for details):

$$\partial_t(t(v'-1)) + [\partial_t v]t\partial_v v' = -f_0 \tag{26a}$$

$$\partial_t[\partial_t v] + \frac{2}{t}[\partial_t v] + [\partial_t v]\partial_v[\partial_t v] = -\frac{v'}{t} < \nabla^{\perp} P_{\neq 0}\phi \cdot \nabla \tilde{u} > \tag{26b}$$

$$v''(t,v) = v'(t,v)\partial_v v'(t,v). \tag{26c}$$

Note that to leading order in ϵ, one can express $v'-1$ as a time average of $-f_0$. Note also that we have a simple expression for $\partial_v \tilde{u}_0$ from the Biot-Savart law:

$$\partial_v \tilde{u}_0(t,v) = \frac{1}{v'(t,v)}\partial_y U_0^x(t,y) = -\frac{1}{v'(t,v)}\omega_0(t,y) = -\frac{1}{v'(t,v)} f_0(t,v). \tag{27}$$

Given a priori estimates on the system (24), (26), we can recover estimates on the original system (2) by the inverse function theorem as long as $v'-1$ remains sufficiently small (see Bedrossian and Masmoudi [ibid.] for details).

5.2 Construction of the toy model norm. For simplicity of notation in this section we usually take $\eta, k > 0$ but the work applies equally well to $\eta, k < 0$. Note that modes where $\eta k < 0$ do not have resonances for positive times. Keeping with the intuition from the derivation of (21), in this section we will think of η as a fixed parameter and time varying. Accordingly, in this section we will use $I_{k,\eta} = [\frac{\eta}{k} - \frac{\eta}{2k(k+1)}, \frac{\eta}{k} + \frac{\eta}{2k(k-1)}]$ to denote any resonant interval with $\eta/k^2 \geq 1$ (with the modification $[\eta - \frac{\eta}{4}, 2\eta]$ if $k = 1$). A key feature of the methods in Bedrossian and Masmoudi [ibid.] is how the toy model is used to construct a norm which precisely matches the estimated worst-case behavior that the reaction terms create, done by choosing $w_k(t,\eta)$ to be an approximate solution to (21). First we have the following (easy to check) Proposition.

Proposition 1. *Let $\tau = t - \frac{\eta}{k}$ and consider the solution $(f_R(\tau), f_{NR}(\tau))$ to (21) with $f_R\left(-\frac{\eta}{k^2}\right) = f_{NR}\left(-\frac{\eta}{k^2}\right) = 1$. There exists a constant C such that for all $\kappa < 1/2$ and $\frac{\eta}{k^2} \geq 1$,*

$$f_R(\tau) \leq C\left(\frac{k^2}{\eta}(1+|\tau|)\right)^{-C\kappa} \qquad -\frac{\eta}{k^2} \leq \tau \leq 0,$$

$$f_{NR}(\tau) \leq C\left(\frac{k^2}{\eta}(1+|\tau|)\right)^{-C\kappa-1} \qquad -\frac{\eta}{k^2} \leq \tau \leq 0,$$

$$f_R(\tau) \leq C\left(\frac{\eta}{k^2}\right)^{C\kappa}(1+|\tau|)^{1+C\kappa} \qquad 0 \leq \tau \leq \frac{\eta}{k^2},$$

$$f_{NR}(\tau) \leq C\left(\frac{\eta}{k^2}\right)^{C\kappa+1}(1+|\tau|)^{C\kappa} \qquad 0 \leq \tau \leq \frac{\eta}{k^2}.$$

For the remainder of the paper we fix κ such that $3/2 < (1+2C\kappa) < 10$.

Remark 1. It is important to notice that over the whole interval $\left[-\frac{\eta}{k^2}, \frac{\eta}{k^2}\right]$, both f_R and f_{NR} are at most amplified by roughly the same factor $C(\frac{\eta}{k^2})^{1+2C\kappa}$. Over the interval $[-\frac{\eta}{k^2}, 0]$, f_{NR} is amplified at most by $C(\frac{\eta}{k^2})^{1+C\kappa}$ and f_R is amplified at most by $C(\frac{\eta}{k^2})^{C\kappa}$. Whereas, over the interval $[0, \frac{\eta}{k^2}]$, f_{NR} is amplified at most by $C(\frac{\eta}{k^2})^{C\kappa}$ and f_R is amplified at most by $C(\frac{\eta}{k^2})^{1+C\kappa}$. Near the critical time, the imbalance between f_{NR} and f_R is the largest - in particular, the resonant mode f_R is a factor of $\frac{\eta}{k^2}$ less than f_{NR} at this time. However by the end of the interval, the total growth of the resonant and non-resonant modes are comparable. The fact that f_R and f_{NR} are amplified the same over that interval will simplify the construction of w.

On each interval $I_{k,\eta}$, growth of the resonant mode (k, η) will be modeled by w_R and the rest of the modes (which are non-resonant) will be modeled by w_{NR}. By Proposition 1, we will able to choose w such that the total growth of w_R and w_{NR} exactly agree.

The construction is done backward in time, starting with $k = 1$. For $t \in I_{k,\eta}$ and $\tau = t - \frac{\eta}{k}$, we will choose (w_{NR}, w_R) such that over the interval $I_{k,\eta}$ they approximately satisfy (21):

$$\begin{aligned} \partial_\tau w_R &\approx \kappa \frac{k^2}{\eta} w_{NR}, \\ \partial_\tau w_{NR} &\approx \kappa \frac{\eta}{k^2(1+\tau^2)} w_R, \end{aligned} \tag{28}$$

We first construct the non-resonant component w_{NR} and then explain how we should modify it over each interval $I_{k,\eta}$ to construct w_R.

Let w_{NR} be a non-decreasing function of time with $w_{NR}(t, \eta) = 1$ for $t \geq 2\eta$. For $k \geq 1$, we assume that $w_{NR}(t_{k-1,\eta})$ was computed. To compute w_{NR} on the interval $I_{k,\eta}$, we use the growth predicted by Proposition 1: for $k = 1, 2, 3, ..., E(\sqrt{\eta})$, we define

(29a)
$$w_{NR}(t, \eta) = \left(1 + a_{k,\eta}|t - \frac{\eta}{k}|\right)^{-1-C\kappa} w_{NR}\left(\frac{\eta}{k}\right), \qquad \forall t \in I^L_{k,\eta} = \left[t_{k,\eta}, \frac{\eta}{k}\right],$$
(29b)
$$w_{NR}(t, \eta) = \left(\frac{k^2}{\eta}\left[1 + b_{k,\eta}|t - \frac{\eta}{k}|\right]\right)^{C\kappa} w_{NR}(t_{k-1,\eta}), \qquad \forall t \in I^R_{k,\eta} = \left[\frac{\eta}{k}, t_{k-1,\eta}\right].$$

The constant $b_{k,\eta}$ is chosen to ensure that $\frac{k^2}{\eta}\left[1 + b_{k,\eta}|t_{k-1,\eta} - \frac{\eta}{k}|\right] = 1$, hence for $k \geq 2$, we have

$$b_{k,\eta} = \frac{2(k-1)}{k}\left(1 - \frac{k^2}{\eta}\right) \tag{30}$$

and $b_{1,\eta} = 1 - 1/\eta$. Similarly, $a_{k,\eta}$ is chosen to ensure $\frac{k^2}{\eta}\left[1 + a_{k,\eta}|t_{k,\eta} - \frac{\eta}{k}|\right] = 1$, which implies

$$a_{k,\eta} = \frac{2(k+1)}{k}\left(1 - \frac{k^2}{\eta}\right). \tag{31}$$

Hence, $w_{NR}(\frac{\eta}{k}) = w_{NR}(t_{k-1,\eta})\left(\frac{k^2}{\eta}\right)^{C\kappa}$ and $w_{NR}(t_{k,\eta}) = w_{NR}(t_{k-1,\eta})\left(\frac{k^2}{\eta}\right)^{1+2C\kappa}$. The choice of $a_{k,\eta}$ and $b_{k,\eta}$ was made to ensure that the ratio between $w_{NR}(t_{k,\eta})$ and $w_{NR}(t_{k-1,\eta})$ is exactly $\left(\frac{k^2}{\eta}\right)^{1+2C\kappa}$. Finally, we take w_{NR} to be constant on the interval $[0, t_{E(\sqrt{\eta}),\eta}]$, namely $w_{NR}(t,\eta) = w(t_{E(\sqrt{\eta}),\eta}, \eta)$ for $t \in [0, t_{E(\sqrt{\eta}),\eta}]$. Note that we always have $0 \le b_{k,\eta} < 1$ and $0 \le a_{k,\eta} < 4$, but that $a_{k,\eta}$ and $b_{k,\eta}$ approach 0 when k approaches $E(\sqrt{\eta})$. This will present minor technical difficulties in the sequel since this implies that $\partial_t w$ vanishes near this time and hence a loss of the lower bounds in (28).

On each interval $I_{k,\eta}$, we define $w_R(t,\eta)$ by

$$w_R(t,\eta) = \frac{k^2}{\eta}\left(1 + a_{k,\eta}\left|t - \frac{\eta}{k}\right|\right) w_{NR}(t,\eta), \qquad \forall t \in I^L_{k,\eta} = \left[t_{k,\eta}, \frac{\eta}{k}\right], \tag{32a}$$

$$w_R(t,\eta) = \frac{k^2}{\eta}\left(1 + b_{k,\eta}\left|t - \frac{\eta}{k}\right|\right) w_{NR}(t,\eta), \qquad \forall t \in I^R_{k,\eta} = \left[\frac{\eta}{k}, t_{k-1,\eta}\right]. \tag{32b}$$

Due to the choice of $b_{k,\eta}$ and $a_{k,\eta}$, we get that $w_R(t_{k,\eta},\eta) = w_{NR}(t_{k,\eta},\eta)$ and $w_R(\frac{\eta}{k},\eta) = \frac{k^2}{\eta} w_{NR}(\frac{\eta}{k},\eta)$.

To define the full $w_k(t,\eta)$, we then have

$$w_k(t,\eta) = \begin{cases} w_k(t_{E(\sqrt{\eta}),\eta},\eta) & t < t_{E(\sqrt{\eta}),\eta} \\ w_{NR}(t,\eta) & t \in [t_{E(\sqrt{\eta}),\eta}, 2\eta] \setminus I_{k,\eta} \\ w_R(t,\eta) & t \in I_{k,\eta} \\ 1 & t \ge 2\eta. \end{cases} \tag{33}$$

Since w_R and w_{NR} agree at the end-points of $I_{k,\eta}$, $w_k(t,\eta)$ is Lipschitz continuous in time. This completes the construction of w which appears in the J defined above (36).

The following lemma shows that the toy model predicts a growth of high frequencies which amounts to a loss of Gevrey-2 regularity, which is the primary origin of the restriction $s > 1/2$ in Theorem 1, though naturally the main application of the toy model is to design the norm.

Lemma 1 (Growth of w). For $\eta > 1$, we have for $\mu = 4(1 + 2C\kappa)$,

$$\frac{1}{w_k(0,\eta)} = \frac{1}{w_k(t_{E(\sqrt{\eta}),\eta},\eta)} \sim \frac{1}{\eta^{\mu/8}} e^{\frac{\mu}{2}\sqrt{\eta}}. \tag{34}$$

Here $\sim$ is in the sense of asymptotic expansion.

Proof. Counting the growth over each interval implied by (33) gives the exact formula:

$$\frac{1}{w_k(0,\eta)} = \left(\frac{\eta}{N^2}\right)^c \left(\frac{\eta}{(N-1)^2}\right)^c \dots \left(\frac{\eta}{1^2}\right)^c = \left[\frac{\eta^N}{(N!)^2}\right]^c,$$

where $c = 1 + 2C\kappa$. Recall Stirling's formula $N! \sim \sqrt{2\pi N}(N/e)^N$, which implies

$$(w_k(0,\eta))^{-1/c} \sim \frac{\eta^N}{(2\pi N)(N/e)^{2N}} \sim \frac{1}{2\pi\sqrt{\eta}} e^{2\sqrt{\eta}} \left[\frac{\sqrt{\eta}}{N} e^{2N-2\sqrt{\eta}} \left(\frac{\eta}{N^2}\right)^N\right]$$

and the result follows since the term between [..] is ≈ 1 since $\left|N - \sqrt{\eta}\right| \le 1$. □

5.3 Main energy estimate. Our goal is to control solutions to (24) uniformly in a suitable norm as $t \to \infty$. The key idea we use for this is the carefully designed time-dependent norm written as

$$\|A(t,\nabla) f\|_2^2 = \sum_k \int_\eta \left|A_k(t,\eta)\hat{f}_k(t,\eta)\right|^2 d\eta.$$

The multiplier A has several components,

$$A_k(t,\eta) = e^{\lambda(t)|k,\eta|^s} \langle k,\eta\rangle^\sigma J_k(t,\eta).$$

The index $\lambda(t)$ is the bulk Gevrey$-\frac{1}{s}$ regularity and will be chosen to satisfy

$$\lambda(t) = \frac{3}{4}\lambda_0 + \frac{1}{4}\lambda', \quad t \le 1 \tag{35a}$$

$$\dot{\lambda}(t) = -\frac{\delta_\lambda}{\langle t\rangle^{2\tilde{q}}}(1 + \lambda(t)), \quad t > 1 \tag{35b}$$

where $\delta_\lambda \approx \lambda_0 - \lambda'$ is a small parameter that ensures $\lambda(t) > \lambda_0/2 + \lambda'/2$ and $1/2 < \tilde{q} \le s/8 + 7/16$ is a parameter chosen by the proof. The reason for (35a) is to account for the behavior of the solution on the time-interval $[0,1]$; see Bedrossian and Masmoudi [2013] for this minor detail. The main multiplier for dealing with the Orr mechanism and the associated nonlinear growth is

$$J_k(t,\eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)} + e^{\mu|k|^{1/2}}, \tag{36}$$

where $w_k(t,\eta)$ was constructed above and describes the expected 'worst-case' growth due to nonlinear interactions at the critical times. What will be important is that J imposes more regularity on modes which satisfy $t \sim \frac{\eta}{k}$ (the 'resonant modes') than those that do not (the 'non-resonant modes'). by controlled loss of regularity and is reminiscent of the notion of losing regularity estimates used in e.g. Bahouri and Chemin [1994] and Chemin and Masmoudi [2001]. One of the main differences is that here we have to be more precise in the sense that the loss of regularity occurs for different frequencies during different time intervals.

With this special norm, we can define our main energy:

$$E(t) = \frac{1}{2}\|A(t)f(t)\|_2^2 + E_v(t), \tag{37}$$

where, for some constants K_v, K_D depending only on s, λ, λ' fixed by the proof,

$$\begin{aligned} E_v(t) = \langle t\rangle^{2+2s}\|\frac{A}{\langle\partial_v\rangle^s}v'\partial_v[\partial_t v](t)\|_2^2 + \\ + \langle t\rangle^{4-K_D\epsilon}\|[\partial_t v](t)\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2 + \frac{1}{K_v}\|A^R(v'-1)(t)\|_2^2. \end{aligned} \tag{38}$$

In a sense, there are two coupled energy estimates: the one on Af and the one on E_v. The latter quantity is encoding information about the coordinate system, or equivalently, the evolution of the background shear flow. It turns out $v'\partial_v[\partial_t v]$ is a physical quantity that measures the convergence of the x-averaged vorticity to its time average and satisfies a useful PDE. It will be convenient to get two separate estimates on $[\partial_t v]$ as opposed to just one ($[\partial_t v]$ is essentially measuring how rapidly the x-averaged velocity is converging to its time average).

By the well-posedness theory for 2D Euler in Gevrey spaces Bardos and Benachour [1977], Ferrari and Titi [1998], Foias and Temam [1989], Levermore and Oliver [1997], and Kukavica and Vicol [2009] we may safely ignore the time interval (say) $[0, 1]$ by further restricting the size of the initial data. See Bedrossian and Masmoudi [2013] for a slightly more detailed discussion. The goal is next to prove by a continuity argument that the energy $E(t)$ (together with some related quantities) is uniformly bounded for all time if ϵ is sufficiently small. We define the following controls referred to in the sequel as the bootstrap hypotheses for $t \geq 1$,

(B1) $E(t) \leq 4\epsilon^2$;

(B2) $\|v' - 1\|_\infty \leq \frac{3}{4}$

(B3) 'CK' integral estimates (for 'Cauchy-Kovalevskaya'):

$$\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + \\ + K_v^{-1}\left(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau)\right) + K_v^{-1}\sum_{i=1}^{2}\left(CCK_w^i(\tau) + CCK_\lambda^i(\tau)\right)\Big] d\tau \leq 8\epsilon^2$$

The CK terms above that appear without the K_v^{-1} prefactor arise from the time derivatives of $A(t)$ and are naturally controlled by the energy estimates we are making. The others are related quantities that are controlled separately in Proposition 6 below. These both will be defined below when discussing the energy estimates.

Let I_E be the connected set of times $t \geq 1$ such that the bootstrap hypotheses (B1-B3) are all satisfied. We will work on regularized solutions for which we know $E(t)$ takes values continuously in time, and hence I_E is a closed interval $[1, T^\star]$ with $T^\star > 1$. The bootstrap is complete if we show that I_E is also open, which is the purpose of the following proposition, the proof of which constitutes the majority of this work.

Proposition 2 (Bootstrap)**.** *There exists an $\epsilon_0 \in (0, 1/2)$ depending only on λ, λ', s and σ such that if $\epsilon < \epsilon_0$, and on $[1, T^\star]$ the bootstrap hypotheses (B1-B3) hold, then for $\forall\, t \in [1, T^\star]$,*

1. $E(t) < 2\epsilon^2$,
2. $\|1 - v'\|_\infty < \frac{5}{8}$,
3. *and the CK controls satisfy:*

$$\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + \\ + K_v^{-1}\left(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau)\right) + K_v^{-1}\sum_{i=1}^{2}\left(CCK_w^i(\tau) + CCK_\lambda^i(\tau)\right)\Big] d\tau \leq 6\epsilon^2,$$

from which it follows that $T^\star = +\infty$.

The remainder of the section is devoted to the proof of Proposition 2, the primary step being to show that on $[1, T^\star]$, we have

$$E(t) + \frac{1}{2}\int_1^t \Big[CK_\lambda(\tau) + CK_w(\tau) + CK_w^{v,2}(\tau) + CK_\lambda^{v,2}(\tau) + \\ + K_v^{-1}\left(CK_w^{v,1}(\tau) + CK_\lambda^{v,1}(\tau)\right) + K_v^{-1}\sum_{i=1}^{2}\left(CCK_w^i(\tau) + CCK_\lambda^i(\tau)\right)\Big] d\tau \leq E(1) + K\epsilon^3 \tag{39}$$

for some constant K which is independent of ϵ and $T^\star$. If ϵ is sufficiently small then (39) implies Proposition 2. Indeed, the control $\|1 - v'\| < 5/8$ is an immediate consequence of (B1) by Sobolev embedding for ϵ sufficiently small.

To prove (39), it is natural to compute the time evolution of $E(t)$,

$$\frac{d}{dt}E(t) = \frac{1}{2}\frac{d}{dt}\int |Af|^2\,dx + \frac{d}{dt}E_v(t)$$

The first contribution is of the form

$$\frac{1}{2}\frac{d}{dt}\int |Af|^2\,dx = -CK_\lambda - CK_w - \int AfA(u\cdot\nabla f)dx, \tag{40}$$

where the CK stands for 'Cauchy-Kovalevskaya' since these three terms arise from the progressive weakening of the norm in time, and are expressed as

$$CK_\lambda = -\dot{\lambda}(t)\|\,|\nabla|^{s/2}Af\|_2^2 \tag{41a}$$

$$CK_w = \sum_k \int \frac{\partial_t w_k(t,\eta)}{w_k(t,\eta)} e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)} A_k(t,\eta)\left|\hat{f}_k(t,\eta)\right|^2 d\eta. \tag{41b}$$

In what follows we define

$$\tilde{J}_k(t,\eta) = \frac{e^{\mu|\eta|^{1/2}}}{w_k(t,\eta)}, \tag{42a}$$

$$\tilde{A}_k(t,\eta) = e^{\lambda(t)|k,\eta|^s}\langle k,\eta\rangle^\sigma \tilde{J}_k(t,\eta). \tag{42b}$$

Note that $\tilde{A} \le A$ and if $|k| \le |\eta|$ then $A \lesssim \tilde{A}$.

Strictly speaking, equality (40) is not quite rigorous since it involves a derivative of Af, which is not a priori well-defined. To make this calculation rigorous, we have first to approximate the initial data of (2) by (for instance) analytic initial data and use that the global solutions of (2) stay analytic for all time (see Bardos and Benachour [1977], Foias and Temam [1989], and Ferrari and Titi [1998]). Hence, we can perform all calculations on these solutions with regularized initial data and then perform a passage to the limit to infer that (39) still holds.

To treat the main term in (40), begin by integrating by parts, as in the techniques Foias and Temam [1989], Levermore and Oliver [1997], Kukavica and Vicol [2009], and Gerard-Varet and Masmoudi [2015]

$$\int AfA(u\cdot\nabla f)dx = -\frac{1}{2}\int \nabla\cdot u\,|Af|^2\,dx + \int Af\,[A(u\cdot\nabla f) - u\cdot\nabla Af]\,dx. \tag{43}$$

Notice that the relative velocity is not divergence free:

$$\nabla \cdot u = \partial_v [\partial_t v] + \partial_v v' \partial_z \phi.$$

The first term is controlled by the bootstrap hypothesis (B1). For the second term we pay regularity and show that under the bootstrap hypotheses we have

$$\| P_{\neq 0} \phi(t) \|_{\mathcal{G}^{\lambda(t),\sigma-3}} \lesssim \frac{\epsilon}{\langle t \rangle^2}. \tag{44}$$

Therefore, by Sobolev embedding, $\sigma > 5$ and the bootstrap hypotheses,

$$\left| \int \nabla \cdot u \, |Af|^2 \, dx \right| \leq \|\nabla u\|_\infty \|Af\|_2^2 \lesssim \frac{\epsilon}{\langle t \rangle^{2-K_D\epsilon/2}} \|Af\|_2^2 \lesssim \frac{\epsilon^3}{\langle t \rangle^{2-K_D\epsilon/2}}. \tag{45}$$

To handle the commutator, $\int Af \, [A(u \cdot \nabla f) - u \cdot \nabla Af] \, dx$, we use a paraproduct decomposition (see e.g. Bony [1981] and Bahouri, Chemin, and Danchin [2011]). Precisely, we define three main contributions: *transport*, *reaction* and *remainder*:

$$\int Af \, [A(u \cdot \nabla f) - u \cdot \nabla Af] \, dx = \frac{1}{2\pi} \sum_{N \geq 8} T_N + \frac{1}{2\pi} \sum_{N \geq 8} R_N + \frac{1}{2\pi} \mathfrak{R}, \tag{46}$$

where (the factors of 2π are for future notational convenience)

$$\begin{aligned} T_N &= 2\pi \int Af \left[A(u_{<N/8} \cdot \nabla f_N) - u_{<N/8} \cdot \nabla A f_N \right] dx \\ R_N &= 2\pi \int Af \left[A(u_N \cdot \nabla f_{<N/8}) - u_N \cdot \nabla A f_{<N/8} \right] dx \\ \mathfrak{R} &= 2\pi \sum_{N \in \mathbb{D}} \sum_{\frac{1}{8}N \leq N' \leq 8N} \int Af \, [A(u_N \cdot \nabla f_{N'}) - u_N \cdot \nabla A f_{N'}] \, dx. \end{aligned}$$

Here $N \in \mathbb{D} = \{\frac{1}{2}, 1, 2, 4, ..., 2^j, ...\}$ and g_N denotes the N-th Littlewood-Paley projection and $g_{<N}$ means the Littlewood-Paley projection onto frequencies less than N. Formally, the paraproduct decomposition (46) represents a kind of 'linearization' for the evolution of higher frequencies around the lower frequencies. The terminology 'reaction' is borrowed from Mouhot and Villani [2011].

To control the transport term, we use

Proposition 3 (Transport)**.** *Under the bootstrap hypotheses,*

$$\sum_{N \geq 8} |T_N| \lesssim \epsilon C K_\lambda + \epsilon C K_w + \frac{\epsilon^3}{\langle t \rangle^{2-K_D\epsilon/2}}.$$

The proof of Proposition 3 uses ideas from the works of Foias and Temam [1989], Levermore and Oliver [1997], and Kukavica and Vicol [2009]. Since the velocity u is restricted to 'low frequency', we will have the available regularity required to apply (44). However, the methods of Foias and Temam [1989], Levermore and Oliver [1997], and Kukavica and Vicol [2009] do not adapt immediately since $J_k(t,\eta)$ is imposing slightly different regularities to certain frequencies, which is problematic. Physically speaking, we need to ensure that resonant frequencies do not incur a very large growth due to nonlinear interactions with non-resonant frequencies (which are permitted to be slightly larger than the resonant frequencies). Controlling this imbalance is why CK_w appears in Proposition 3.

Controlling the reaction contribution in (46) is one of the main tasks. Here we cannot apply (44), as an estimate on this term requires u in the highest norm on which we have control, and hence we have no regularity to spare. Physically, here in the reaction term is where the dangerous nonlinear effects are expressed and a great deal of precision is required to control them.

Proposition 4 (Reaction)**.** *Under the bootstrap hypotheses,*

$$\begin{aligned}\sum_{N\geq 8}|R_N| &\lesssim \epsilon CK_\lambda + \epsilon CK_w + \frac{\epsilon^3}{\langle t\rangle^{2-K_D\epsilon/2}} + \epsilon CK_\lambda^{v,1} + \epsilon CK_w^{v,1} \\ &\quad + \epsilon\|\langle\frac{\partial_v}{t\partial_z}\rangle^{-1}\left(\partial_z^2+(\partial_v-t\partial_z)^2\right)\left(\frac{|\nabla|^{s/2}}{\langle t\rangle^s}A+\sqrt{\frac{\partial_t w}{w}}\tilde{A}\right)P_{\neq 0}\phi\|_2^2. \end{aligned} \tag{47}$$

The $CK^{v,1}$ terms are defined below in (51). The first step to controlling the term in (47) involving ϕ is Proposition 5. This proposition treats Δ_t as a perturbation of $\partial_{zz}+(\partial_v - t\partial_z)^2$ and passes the multipliers in the last term of (47) onto f *and* the coefficients of Δ_t. Physically, these latter contributions are indicating the nonlinear interactions between the higher modes of f and the coefficients v', v'' (which involve time-averages of f_0 (26)). Analogous lemmas have continued to play important roles in the theory.

Proposition 5 (Precision elliptic control)**.** *Under the bootstrap hypotheses,*

$$\begin{aligned}&\|\langle\frac{\partial_v}{t\partial_z}\rangle^{-1}\left(\partial_z^2+(\partial_v-t\partial_z)^2\right)\left(\frac{|\nabla|^{s/2}}{\langle t\rangle^s}A+\sqrt{\frac{\partial_t w}{w}}\tilde{A}\right)P_{\neq 0}\phi\|_2^2 \\ &\qquad \lesssim CK_\lambda + CK_w + \epsilon^2\sum_{i=1}^{2} CCK_\lambda^i + CCK_w^i,\end{aligned} \tag{48}$$

where the 'coefficient Cauchy-Kovalevskaya' terms are given by

$$CCK_{\lambda}^{1} = -\dot{\lambda}(t)\| \, |\partial_v|^{s/2} A^R \left(1-(v')^2\right) \|_2^2, \tag{49a}$$

$$CCK_{w}^{1} = \| \sqrt{\frac{\partial_t w}{w}} A^R \left(1-(v')^2\right) \|_2^2, \tag{49b}$$

$$CCK_{\lambda}^{2} = -\dot{\lambda}(t)\| \, |\partial_v|^{s/2} \frac{A^R}{\langle \partial_v \rangle} v'' \|_2^2, \tag{49c}$$

$$CCK_{w}^{2} = \| \sqrt{\frac{\partial_t w}{w}} \frac{A^R}{\langle \partial_v \rangle} v'' \|_2^2. \tag{49d}$$

The next step in the bootstrap is to provide good estimates on the coordinate system and the associated CK and CCK terms. The following proposition provides controls on $v'-1$, the CCK terms arising in (49), the pair $[\partial_t v]$, $v'\partial_v[\partial_t v]$ and finally all of the $CK^{v,i}$ terms. The norm defined by $A^R(t)$ is stronger than that defined by $A(t)$, which we use to measure f. It turns out that we will be able to propagate this stronger regularity on $v'-1$ due to a time-averaging effect, derived via energy estimates on (26). By contrast, $[\partial_t v]$ is expected basically to have the regularity of $\tilde{u}_0$ and hence even (50b) has s fewer derivatives than expected. On the other hand, it has a significant amount of time decay, which near critical times can be converted into regularity.

Proposition 6 (Coordinate system controls)**.** *Under the bootstrap hypotheses, for ϵ sufficiently small and K_v sufficiently large there is a $K > 0$ such that*

(50a)

$$\|A^R(v'-1)(t)\|_2^2 + \frac{1}{2}\int_1^t \sum_{i=1}^{2} CCK_w^i(\tau)d\tau + \frac{1}{2}\int_1^t \sum_{i=1}^{2} CCK_\lambda^i(\tau)d\tau \le \frac{1}{2}K_v\epsilon^2$$

$$\langle t\rangle^{2+2s}\|\frac{A}{\langle \partial_v\rangle^s} v'\partial_v[\partial_t v]\|_2^2 + \frac{1}{2}\int_1^t CK_\lambda^{v,2}(\tau) + CK_w^{v,2}(\tau)d\tau \le \epsilon^2 + K\epsilon^3 \tag{50b}$$

$$\langle t\rangle^{4-K_D\epsilon}\|[\partial_t v]\|_{\mathcal{G}^{\lambda(t),\sigma-6}}^2 \le \epsilon^2 + K\epsilon^3 \tag{50c}$$

$$\int_1^t CK_\lambda^{v,1}(\tau) + CK_w^{v,1}(\tau)d\tau \le \frac{1}{2}K_v\epsilon^2, \tag{50d}$$

where the $CK^{v,i}$ terms are given by

$$CK_w^{v,2}(\tau) = \langle \tau \rangle^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v \rangle^s} v' \partial_v [\partial_t v](\tau) \right\|_2^2 \tag{51a}$$

$$CK_\lambda^{v,2}(\tau) = \langle \tau \rangle^{2+2s} (-\dot{\lambda}(\tau)) \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} v' \partial_v [\partial_t v](\tau) \right\|_2^2 \tag{51b}$$

$$CK_w^{v,1}(\tau) = \langle \tau \rangle^{2+2s} \left\| \sqrt{\frac{\partial_t w}{w}} \frac{A}{\langle \partial_v \rangle^s} [\partial_t v](\tau) \right\|_2^2 \tag{51c}$$

$$CK_\lambda^{v,1}(\tau) = \langle \tau \rangle^{2+2s} (-\dot{\lambda}(\tau)) \left\| |\partial_v|^{s/2} \frac{A}{\langle \partial_v \rangle^s} [\partial_t v](\tau) \right\|_2^2. \tag{51d}$$

Note that neither (50b) nor (50c) controls the other: at higher frequencies the former is stronger than the latter and at lower frequencies the opposite is true. One of the advantages of this scheme is that $v' \partial_v [\partial_t v]$ satisfies an equation that is simpler than $[\partial_t v]$ and so is easier to get good estimates on. Both (50b) and (50c) are linked to the convergence of the background shear flow; in particular, they rule out that the background flow oscillates or wanders due to nonlinear effects.

Finally we need to control the remainder term in (46). This is straightforward and is detailed in Bedrossian and Masmoudi [2013].

Proposition 7 (Remainders)**.** *Under the bootstrap hypotheses,*

$$\mathfrak{R} \lesssim \frac{\epsilon^3}{\langle t \rangle^{2-K_D \epsilon/2}}.$$

Collecting Propositions 3, 4, 5, 6, 7 with (46) and (45), we have finally (39) for ϵ sufficiently small with constants independent of both ϵ and $T^\star$; hence for ϵ sufficiently small we may propagate the bootstrap control and prove Proposition 2.

6 Proof of the instability result

6.1 Ideas of the proof. The proof starts by performing the same coordinate change $(x, y) \mapsto (z, v)$, as defined in Equation (22a) and Equation (22b). Following the calculations of Section 5.1, this then gets rid of the badly behaving zeroth mode, and reduces Equation (2) to a system for f, $v' - 1$, and $[\partial_t v]$. These are recorded in Equation (52) $\sim$ Equation (53) below, where for simplicity we denote $h := v' - 1$, $\theta := [\partial_t v]$ and $g := (f, h, \theta)$:

The goal is to find a solution (f, h, θ) to the system

$$\begin{cases} \partial_t f = -\theta \cdot \partial_v f - (h+1)\nabla^{\perp}\phi \cdot \nabla f, \\ \partial_t h = -\theta \partial_v h - \dfrac{\mathbb{P}_0 f + h}{t}, \\ \partial_t \theta = -\dfrac{2\theta}{t} - \theta \partial_v \theta + \dfrac{1}{t}\mathbb{P}_0(f \cdot \partial_z \phi), \end{cases} \tag{52}$$

where the relevant quantities are defined as

$$\begin{cases} \phi = \mathbb{P}_{\neq 0}\Delta_t^{-1} f, \\ \Delta_t = \partial_z^2 + (h+1)^2(\partial_v - t\partial_z)^2 + (h+1)(\partial_v h)(\partial_v - t\partial_z). \end{cases} \tag{53}$$

Moreover, for $t \in I$ we have

$$t(h+1)\partial_v \theta(t) + \mathbb{P}_0 f(t) + h(t) = 0; \quad \int_{\mathbb{R}} \frac{h(t,v)}{h(t,v)+1}\, \mathrm{d}v = 0. \tag{54}$$

6.1.1 The choice of data, and setup. We will construct (f, h, θ) that satisfies the required instability assumptions. This solution will be constructed as the superposition of a background solution $(\underline{f}, \underline{h}, \underline{\theta})$, and a perturbation (f^*, h^*, θ^*) (which is a second order perturbation of a much smaller size but lower regularity). It turns out that (h, θ) plays a relatively less important role in the proof, so for simplicity, here we will consider f only.

The background solution $\underline{f}$ is guaranteed to exist by Theorem 1; we will assume it has analytic regularity (i.e. $s = 1$ in Theorem 1), and has size $\varepsilon_0 \ll \varepsilon$. More precisley, we define the background solution $\underline{g} := (\underline{f}, \underline{h}, \underline{\theta})$, which solves (52)~(53) with initial data

$$\underline{f}(1, z, v) = \varepsilon_0 \cos z \cdot \varphi_b(v), \quad \underline{h}(1, z, v) = \underline{\theta}(1, z, v) = 0, \tag{55}$$

where

$$\varphi_b(v) = e^{-(C_0^{-1} v)^{18}}.$$

By Theorem 1, we know that $\underline{g}$ exists on $[1, +\infty)$, and satisfies the following properties, where recall that all constants here depend on C_0:

1. $\underline{f}$ and $\underline{h}$ are real-valued and even, $\underline{\theta}$ is real-valued and odd, and

$$\|\underline{f}(t)\|_{\mathfrak{A}_{C_0}} + \|\underline{h}(t)\|_{\mathfrak{A}_{C_0}} + \|\underline{\theta}(t)\|_{\mathfrak{A}_{C_0}} \lesssim \varepsilon_0, \quad \|\underline{\theta}(t)\|_{\mathfrak{A}_{C_0 - 1}} \lesssim \frac{\varepsilon_0}{t^2}; \tag{56}$$

2. $\underline{f}$ and $\underline{h}$ converge as $t \to \infty$,

$$\|\underline{f}(t) - f_\infty\|_{\mathfrak{A}_{C_0 - 1}} \lesssim \frac{\varepsilon_0^2}{t}, \quad \|\underline{h}(t) + \mathbb{P}_0 f_\infty\|_{\mathfrak{A}_{C_0 - 1}} \lesssim \frac{\varepsilon_0}{t}; \tag{57}$$

3. The limit f_∞ is close to the specific profile we choose, namely

$$\left\| f_\infty - \varepsilon_0 \cos z \cdot \varphi_b(v) \right\|_{\mathfrak{G}_{C_0-1}} \lesssim \varepsilon_0^2. \tag{58}$$

We also define the function $\underline{\phi}$ and the operator $\underline{\Delta_t}$, corresponding to $(\underline{f}, \underline{h}, \underline{\theta})$, as in (53). In practice we will think of $\underline{f}$ as only having low frequency components, as it is much more regular than the perturbation f^* we will construct.

The perturbation f^* will be fixed by assigning the data at some time $t = T_0$:

$$f^*(T_0) = \varepsilon_1 \cos(k_0 z + \eta_0 v)\varphi_p(k_0\sqrt{\sigma} v),$$

where φ_p is a suitable Schwartz function. The parameters $(\varepsilon_0, k_0, T_0, \eta_0, \sigma)$ are related by (where $N_2 = 30$, $N_3 = 30000$):

$$\begin{aligned} &\sigma = (\log k_0)^{-N_2}, \quad \alpha = 1 + (\log k_0)^{-2N_2}, \quad \varepsilon_0 = (\log k_0)^{-N_3}; \\ &\eta_0 = \frac{2k_0^2\alpha}{\pi\varepsilon_0}, \qquad T_0 = \frac{2\eta_0}{2k_0+1}. \end{aligned} \tag{59}$$

Note that $\alpha - 1 = \sigma^2$ and $\varepsilon_0 = \sigma^{1000}$. For now it suffices to note that $\varepsilon_1 \ll \varepsilon_0$, and $\widehat{f^*(T_0)}$ is concentrated near only two frequencies, (k_0, η_0) and $(-k_0, -\eta_0)$, where (k_0, η_0) is considered the high frequency mode compared to $\underline{g}$. We also fix two times $T_1 \in [T_0, 2T_0]$ and $T_2 \le T_0$, define by

$$\begin{aligned} &t_m = \frac{2\eta_0}{2m+1}, \qquad T_j = t_{k_j}, \ 1 \le j \le 2; \\ &k_1 = (1-\sigma)k_0, \quad k_2 = \varepsilon_0^{-1/40}\sqrt{\eta_0} \end{aligned} \tag{60}$$

Note that $k_2 > k_0 > k_1$ and $T_1 > T_0 > T_2$. For simplicity we will assume all the k_i's are integers (otherwise take their integer parts).

6.1.2 The linearized system. Since $\varepsilon_1 \ll \varepsilon_0$, it is natural to first study the linearization of (52)~(53) at the background solution $\underline{f}$. This linear system has the form $\partial_t f' = \mathfrak{L} f'$, where $\mathfrak{L}$ is a linear operator and $f'(T_0) = f^*(T_0)$. Following the observation made in Bedrossian and Masmoudi [2013], we know $\mathfrak{L}$ consists of two parts: the first one is a "transport" part,

$$\mathfrak{L}^T f' = \underline{\Phi} \cdot \nabla f', \tag{61}$$

where $\underline{\Phi}$ is a combination of the background solution, which has much higher regularity than f' (and thus contributes low frequencies only), and moreover decays like t^{-2}.

The second one is a "reaction" term, which is responsible for the Orr growth mechanism,

$$\mathfrak{L}^R f' = \underline{F} \cdot \nabla \underline{\Delta_t}^{-1} f', \tag{62}$$

where $\underline{F}$ again comes from the background solution, but has no decay in time (one can think $\underline{\Phi} \sim t^{-2}\underline{F}$); the operator $\underline{\Delta_t}^{-1}$ is defined, up to some error terms, by

$$\widehat{\Delta_t^{-1} F}(t,k,\xi) = \frac{1}{(\xi - kt)^2 + k^2} \widehat{F}(t,k,\xi);$$

Notice that, if one compares (61) and (62), say at a critical time $t = \xi/k$, and assume that $\underline{\Phi} \sim t^{-2}\underline{F}$, then $\mathfrak{L}^R$ dominates $\mathfrak{L}^T$ if $t \gg k$ or equivalently $t \gtrsim \sqrt{|\xi|}$, and $\mathfrak{L}^T$ dominates $\mathfrak{L}^R$ if $t \lesssim \sqrt{|\xi|}$.

Our strategy is to show that the size of $\widehat{f}'$, say near $(\pm k_0, \pm \eta_0)$, exhibits growth at critical times *between* T_0 *and* T_1 by the Orr mechanism, and in fact saturates the upper bound proved in Bedrossian and Masmoudi [2013]. Note that this also explains the seemingly strange choice of assigning data at $t = T_0$ instead of $t = 1$, since we only know how to saturate the optimal growth on $[T_0, T_1]$.

Moreover, we need to go *backwards* from T_0 and recover the control for f' at time $t = 1$. There are two regimes here: when t is small (namely $t \leq T_2$; note that T_2 is almost $\sqrt{\eta_0}$, see(60)), the transport term dominates, and the growth of f' can be easily controlled by an energy-type inequality for transport equations. When t is large, namely $t \in [T_2, T_0]$, the reaction term dominates and the situation will be much similar to what happens on $[T_0, T_1]$, except that only an upper bound is needed.

Summing up, we need to obtain a lower bound for f' on $[T_0, T_1]$, and an upper bound for f' on $[T_2, T_0]$, of form

$$|\widehat{f}'(T_1, \pm k_0, \pm \eta_0)| \gtrsim e^{c\sqrt{\eta_0}} \varepsilon_1; \quad |\widehat{f}'(T_2, \pm k_0, \pm \eta_0)| \lesssim e^{c'\sqrt{\eta_0}} \varepsilon_1, \tag{63}$$

for some suitable $c > c' > 0$. This would then imply that $f'(T_1)$ is large in H^N, and that $f'(T_2)$ (and hence $f'(1)$) is small in $\mathcal{G}^*$, upon choosing ε_1 appropriately. In both cases it is crucial to obtain precise bounds on the size of $\widehat{f}'$ near frequency $(\pm k_0, \pm \eta_0)$, which is the next step of the proof.

6.1.3 Linear analysis, and a more precise toy model. We may now restrict the linearized system to time $t \in [T_2, T_1]$, where the transport term plays essentially no role, so we will focus on the reaction term only. Recall the expression in (62); for simplicity we assume that $\underline{F}$ is independent of time and has only $k = \pm 1$ modes, say $\widehat{\underline{F}}(t,k,\xi) = \varepsilon_0 \mathbf{1}_{k=\pm 1} \varphi(\xi)/2$ with a Schwartz function φ.

By (62), we then write down the equation

$$\partial_t \widehat{f}'(t,k,\xi) = \int_{\mathbb{R}} \frac{\varepsilon_0 \eta/2}{(\eta - t(k+1))^2 + (k+1)^2} \widehat{\varphi}(\xi-\eta) \widehat{f}'(t,k+1,\eta)\, \mathrm{d}\eta - \int_{\mathbb{R}} \frac{\varepsilon_0 \eta/2}{(\eta - t(k-1))^2 + (k-1)^2} \widehat{\varphi}(\xi-\eta) \widehat{f}'(t,k-1,\eta)\, \mathrm{d}\eta \tag{64}$$

for $\widehat{f}'$. In Bedrossian and Masmoudi [ibid.], the authors replaced the function φ on the right hand side of (64) by the δ function, obtaining and ODE *toy model* which is essentially an "envelope" of (64) and can be solved explicitly. This is perfect for obtaining an *upper bound* for solutions to (64), but in order to get a *lower bound* a more accurate approximation will be needed - which is precisely what we are able to obtain here, under the assumption $\eta \approx \eta_0$ and $t \in [T_2, T_1]$.

For simplicity, let us assume $t \sim T_0$; recall from (59) that $T_0 \sim \sqrt{\eta_0/\varepsilon_0}$. Since $\eta \approx \eta_0$ due to the definition of $f'(T_0)$, we know that (64) plays a significant role only near the critical times η_0/m, where $m \sim \sqrt{\varepsilon_0 \eta_0}$. We thus cut the time interval into subintervals, each containing exactly one critical time. Define, see also (60),

$$t_m = \frac{2\eta_0}{2m+1}; \qquad \frac{\eta_0}{m} \in [t_m, t_{m-1}] := I_m;$$

then on each I_m, according to (64), only the modes $k = m \pm 1$ will be active (i.e. has significant increments), since

$$\frac{1}{(\eta_0 - kt)^2 + k^2} \lesssim \frac{1}{t^2} \ll \frac{1}{m^2}, \qquad \forall t \in I_m, k \neq m.$$

We can therefore solve (64) approximately and explicitly[1], obtaining an approximate recurrence relation (see (59) for definition of parameters):

$$\mathcal{F} f'(t_{m-1},k,v) = \mathcal{F} f'(t_m,k,v) + \mathfrak{R} + \begin{cases} 0, & k \neq m \pm 1; \\ \mp \dfrac{\alpha k_0^2}{m^2} \varphi(v) \cdot \mathcal{F} f'(t_m, m, v), & k = m \pm 1, \end{cases} \tag{65}$$

after taking inverse Fourier transform in ξ, where the error term $\mathfrak{R}$ is small in L^2.

The recurrence relation (65) then plays the role of the toy model in Bedrossian and Masmoudi [ibid.]. In fact, if we choose φ such that $\|\varphi\|_{L^\infty} = 1$, then this already suffices

[1] Note that this argument works precisely when $t \in [T_2, T_1]$: when t si too small transport terms will come in, and when t is too large the $(m \pm 1)$ modes $\widehat{f}'(t, m \pm 1, \xi)$ will grow too much and destroy the approximate decoupling.

to prove the upper bound on $[T_2, T_0]$, since (65) essentially implies that

$$\sup_k \|\mathscr{F} f'(t_m, k, \cdot)\|_{L^2} \lesssim \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \cdot \sup_k \|\mathscr{F} f'(t_{m-1}, k, \cdot)\|_{L^2},$$

and thus by iteration,

$$\sup_k \|\mathscr{F} f'(T_2, k, \cdot)\|_{L^2} \lesssim \varepsilon_1 \prod_{m=k_2}^{k_0} \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \sim e^{c'\sqrt{\eta_0}}\varepsilon_1. \tag{66}$$

We turn to the lower bound for f' on $[T_0, T_1]$. If φ were identically 1, then in view of the *smallness* of $\mathfrak{R}$, we can use the same argument to obtain that

$$\sup_k \|\mathscr{F} f'(t_{m-1}, k, \cdot)\|_{L^2} \gtrsim \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \cdot \sup_k \|\mathscr{F} f'(t_m, k, \cdot)\|_{L^2},$$

and hence

$$\sup_k \|\mathscr{F} f'(T_1, k, \cdot)\|_{L^2} \gtrsim \varepsilon_1 \prod_{m=k_0}^{k_1} \max\left(1, \frac{\alpha k_0^2}{m^2}\right) \sim e^{c\sqrt{\eta_1}}\varepsilon_1. \tag{67}$$

Comparing (66) and (67) we obtain the desired inequality (63) by direct computations, due to our choice of parameters.

Nevertheless φ cannot be identically 1, and moreover the error term $\mathfrak{R}$ is not local. To recover (67), in view of the factor $\varphi(v)$ on the right hand side of (65), we thus need to localize v in the region where $\varphi(v)$ is equal or close to 1. This localization is achieved by going back to physical space and performing an energy-type estimate for an L^2 norm with exponential weight in physical space.

6.1.4 Nonlinear analysis, and the Taylor expansion. Up to now we have only considered f', which is the solution to the linearized system $\partial_t f' = \mathfrak{L} f'$. The full nonlinear system (52)$\sim$(53), in terms of f^*, can be written as

$$\partial_f f^* = \mathfrak{L} f^* + \mathfrak{N}(f^*, f^*), \tag{68}$$

if, say, we consider only quadratic nonlinearities. Note that f' can also be regarded as the first order term in a formal Taylor expansion of f^*; we may write out the higher order terms by $f^{(1)} = f'$ and

$$\partial_t f^{(n)} = \mathfrak{L} f^{(n)} + \sum_{q_1+q_2=n-1} \mathfrak{N}(f^{(q_1)}, f^{(q_2)}), \quad f^{(n)}(T_0) = 0;$$

Our next step is to prove that, in some sense, we have[2]

$$\text{(the size of } f^{(n)}) \lesssim (\text{the size of } f^{(1)})^n, \tag{69}$$

Since the size of $f^{(1)}$ is $O(\varepsilon_1)$, the bound (69) guarantees that the contribution of $f^{(n)}$ with $n \geq 2$ will be negligible, and thus Theorem 2 follows from the estimates for $f^{(1)}$ obtained above.

The proof of (69) follows from an inductive argument, where at each step we combine the multilinear estimates for the nonlinear term $\mathfrak{N}$ with the linear estimates for the inhomogeneous equation

$$\partial_t f = \mathfrak{L} f + \mathfrak{N},$$

which is proved in the same way as the linear homogeneous case. Here the main difficulty is that $f^{(n)}$, being essentially the n-th power of $f^{(1)}$, is supported in Fourier space at (say) the frequency $(nk_0, n\eta_0)$. We thus need to run the arguments above for this particular choice of frequency, instead of (k_0, η_0). Fortunately this just corresponds to changing of parameters in the Orr growth mechanism, and most of the arguments above still go through.

Finally, to avoid the divergence issue caused by doing the Taylor expansion directly, we will close the whole proof by fixing some very large n_0 and claiming that

$$f^{(1)} + f^{(2)} + \cdots + f^{(n_0)}$$

is an approximate solution to (68), with error term so small that an actual solution to (68) can be constructed by a perturbative argument on the interval $[1, T_1]$.

6.2 Further discussions.

We mention two possible further questions related to Theorem 2.

6.2.1 Asymptotic instability.

Given Theorem 2, an immediate question is whether asymtotic instability can also be proved for (2). We believe this can be done by repeatedly applying the arguments in this paper.

Roughly speaking, we fix the background solution $\underline{f}$ and construct the perturbation $f^* = f_1^*$ as in Theorem 2. Note that f_1^* and grows from some time T_0^1 to some later time T_1^1; We now take $\underline{f} + f_1^*$ as the new background and construct a further perturbation f_2^* which grows from time T_0^2 to T_1^2, and so on. We then pile up a sequence of perturbations and define

$$f := \underline{f} + f_1^* + f_2^* + \cdots,$$

[2]Note however that $f^{(n)}$ is supported at higher and higher frequencies, namely $(nk_0, n\eta_0)$; thus this fact cannot be captured by a bootstrap argument in a single Gevrey norm, and this formal Taylor expansion seems necessary.

which we expect to satisfy that

$$\|f(1)\|_{\mathcal{G}^*} \leq \varepsilon, \quad \lim_{t\to\infty} \|\langle\partial_x\rangle^{N_0} f(t)\|_{L^2} = +\infty.$$

The main difficulty here is to control the evolution of f_1^* after time T_1^1; we then have to extend our arguments, which now covers only critical times η_0/m with $m \gtrsim k_0$, to *all* critical times up to $m = 1$. We believe that a suitable combination of the techniques used in this paper and the weighted energy method used in Bedrossian and Masmoudi [2013] should be the key to solving this problem.

6.2.2 Genericity. Another natural question is whether the Orr growth mechanism is generic, i.e., whether the full upper bound of growth can be saturated for "most" solutions in a suitable sense. To study this problem, we have to consider solutions with general distribution in frequencies, instead of the f^* we choose here, which essentially has only two modes. In such cases we no longer have the simple decoupling as in Section 6.1.3, nor the recurrence relation (65); the main challenge is thus to find a substitute to (65) and to approximate (64), and it would be crucial to be able to separate the different components of the solution that evolve differently. It seems that some further physical-space based techniques will be needed.

Another challenge is the possible cancellations for the toy model (if we can find one) in the generic case. This also depends on how well are different frequencies and different physical space locations separated - if they are mixed together then we would have less control of the solution.

References

Jeffrey S. Baggett, Tobin A. Driscoll, and Lloyd N. Trefethen (1995). "A mostly linear model of transition to turbulence". *Phys. Fluids* 7.4, pp. 833–838. MR: 1324952 (cit. on p. 2162).

H. Bahouri and J.-Y. Chemin (1994). "Équations de transport relatives á des champs de vecteurs non-lipschitziens et mécanique des fluides". *Arch. Rational Mech. Anal.* 127.2, pp. 159–181. MR: 1288809 (cit. on p. 2171).

Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin (2011). *Fourier analysis and nonlinear partial differential equations*. Vol. 343. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Heidelberg, pp. xvi+523. MR: 2768550 (cit. on p. 2174).

C. Bardos and S. Benachour (1977). "Domaine d'analycité des solutions de l'équation d'Euler dans un ouvert de R^n". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 4.4, pp. 647–687. MR: 0454413 (cit. on pp. 2171, 2173).

Jacob Bedrossian (May 2016). "Nonlinear echoes and Landau damping with insufficient regularity". arXiv: 1605.06841 (cit. on pp. 2160–2162).

Jacob Bedrossian, Pierre Germain, and Nader Masmoudi (June 2015a). "Dynamics near the subcritical transition of the 3D Couette flow I: Below threshold case". To appear in Mem. Amer. Math. Soc. arXiv: 1506.03720 (cit. on pp. 2157, 2161).

– (June 2015b). "Dynamics near the subcritical transition of the 3D Couette flow II: Above threshold case". arXiv: 1506.03721 (cit. on pp. 2157, 2161).

– (2017a). "On the stability threshold for the 3D Couette flow in Sobolev regularity". *Ann. of Math. (2)* 185.2, pp. 541–608. MR: 3612004 (cit. on pp. 2157, 2161).

– (Dec. 2017b). "Stability of the Couette flow at high Reynolds number in 2D and 3D". arXiv: 1712.02855 (cit. on p. 2157).

Jacob Bedrossian and Nader Masmoudi (2013). "Inviscid damping and the asymptotic stability of planar shear flows in the 2D Euler equations". *Publ. math. de l'IHÉS* 1306, pp. 1–106 (cit. on pp. 2157, 2160, 2167, 2170, 2171, 2177, 2179–2181, 2184).

Jacob Bedrossian, Nader Masmoudi, and Clément Mouhot (2016a). "Landau damping: paraproducts and Gevrey regularity". *Ann. PDE* 2.1, Art. 4, 71. MR: 3489904 (cit. on p. 2162).

– (2016b). "Landau damping: paraproducts and Gevrey regularity". *Ann. PDE* 2.1, Art. 4, 71. MR: 3489904 (cit. on pp. 2160, 2162).

Jacob Bedrossian, Nader Masmoudi, and Vlad Vicol (2016). "Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the two dimensional Couette flow". *Arch. Ration. Mech. Anal.* 219.3, pp. 1087–1159. MR: 3448924 (cit. on pp. 2157, 2161).

Jean-Michel Bony (1981). "Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires". *Ann. Sci. École Norm. Sup. (4)* 14.2, pp. 209–246. MR: 631751 (cit. on p. 2174).

John P Boyd (1983). "The continuous spectrum of linear Couette flow with the beta effect". *Journal of the atmospheric sciences* 40.9, pp. 2304–2308 (cit. on p. 2159).

K. M. Case (1960). "Stability of inviscid plane Couette flow". *Phys. Fluids* 3, pp. 143–148. MR: 0128230 (cit. on p. 2156).

Jean-Yves Chemin and Nader Masmoudi (2001). "About lifespan of regular solutions of equations related to viscoelastic fluids". *SIAM J. Math. Anal.* 33.1, pp. 84–112. MR: 1857990 (cit. on p. 2171).

Yu Deng and Nader Masmoudi (2018). *Long time instability of the couette flow in low gevrey spaces*. Preprint (cit. on pp. 2157, 2160, 2161).

L. A. Dikiĭ (1961). "The stability of plane-parallel flows of an ideal fluid". *Soviet Physics Dokl.* 5, pp. 1179–1182. MR: 0147072 (cit. on p. 2156).

P. G. Drazin and W. H. Reid (1982). *Hydrodynamic stability*. Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press, Cambridge-New York, pp. xiv+527. MR: 684214 (cit. on p. 2156).

Andrew B. Ferrari and Edriss S. Titi (1998). "Gevrey regularity for nonlinear analytic parabolic equations". *Comm. Partial Differential Equations* 23.1-2, pp. 1–16. MR: 1608488 (cit. on pp. 2171, 2173).

C. Foias and R. Temam (1989). "Gevrey class regularity for the solutions of the Navier-Stokes equations". *J. Funct. Anal.* 87.2, pp. 359–369. MR: 1026858 (cit. on pp. 2171, 2173, 2175).

David Gerard-Varet and Nader Masmoudi (2015). "Well-posedness for the Prandtl system without analyticity or monotonicity". *Ann. Sci. Éc. Norm. Supér.* 48.6, pp. 1273–1325 (cit. on p. 2173).

Lord Kelvin (1887). "Stability of fluid motion: rectilinear motion of viscous fluid between two parallel plates". *Phil. Mag* 24.5, pp. 188–196 (cit. on p. 2156).

Igor Kukavica and Vlad Vicol (2009). "On the radius of analyticity of solutions to the three-dimensional Euler equations". *Proc. Amer. Math. Soc.* 137.2, pp. 669–677. MR: 2448589 (cit. on pp. 2171, 2173, 2175).

C. David Levermore and Marcel Oliver (1997). "Analyticity of solutions for a generalized Euler equation". *J. Differential Equations* 133.2, pp. 321–339. MR: 1427856 (cit. on pp. 2171, 2173, 2175).

Y. Charles Li and Zhiwu Lin (2011). "A resolution of the Sommerfeld paradox". *SIAM J. Math. Anal.* 43.4, pp. 1923–1954. MR: 2831254 (cit. on p. 2156).

Zhiwu Lin and Chongchun Zeng (2011a). "Inviscid dynamical structures near Couette flow". *Arch. Ration. Mech. Anal.* 200.3, pp. 1075–1097. MR: 2796139 (cit. on p. 2161).

– (2011b). "Small BGK waves and nonlinear Landau damping". *Comm. Math. Phys.* 306.2, pp. 291–331. MR: 2824473 (cit. on p. 2161).

Richard S Lindzen (1988). "Instability of plane parallel shear flow (toward a mechanistic picture of how it works)". *Pure and applied geophysics* 126.1, pp. 103–121 (cit. on p. 2159).

J. H. Malmberg, C. B. Wharton, R. W. Gould, and T. M. O'Neil (1968). "Plasma wave echo experiment". *Physical Review Letters* 20.3, pp. 95–97 (cit. on p. 2162).

Clément Mouhot and Cédric Villani (2011). "On Landau damping". *Acta Math.* 207.1, pp. 29–201. MR: 2863910 (cit. on pp. 2158, 2162, 2174).

William M'F Orr (1907). "The Stability or Instability of the Steady Motions of a Perfect Liquid and of a Viscous Liquid. Part I: A Perfect Liquid". In: *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*. Vol. 27. JSTOR, pp. 9–68 (cit. on pp. 2156, 2159).

Lord Rayleigh (1879/80). "On the Stability, or Instability, of certain Fluid Motions". *Proc. Lond. Math. Soc.* 11, pp. 57–70. MR: 1575266 (cit. on p. 2156).

– (1895/96). "On the Stability or Instability of certain Fluid Motions (III.)" *Proc. Lond. Math. Soc.* 27, pp. 5–12. MR: 1576484 (cit. on p. 2156).
– (1887/88). "On the Stability or Instability of certain Fluid Motions, II". *Proc. Lond. Math. Soc.* 19, pp. 67–74. MR: 1576885 (cit. on p. 2156).
Osborne Reynolds (1883). "An Experimental Investigation of the Circumstances Which Determine Whether the Motion of Water Shall Be Direct or Sinuous, and of the Law of Resistance in Parallel Channels." *Proceedings of the Royal Society of London* 35.224-226, pp. 84–99 (cit. on p. 2156).
Peter J. Schmid and Dan S. Henningson (2001). *Stability and transition in shear flows*. Vol. 142. Applied Mathematical Sciences. Springer-Verlag, New York, pp. xiv+556. MR: 1801992 (cit. on p. 2156).
Herbert Brian Squire (1933). "On the stability for three-dimensional disturbances of viscous fluid flow between parallel walls". In: *Proc. R. Soc. Lond. A*. Vol. 142. 847. The Royal Society, pp. 621–628 (cit. on p. 2157).
Lloyd N. Trefethen, Anne E. Trefethen, Satish C. Reddy, and Tobin A. Driscoll (1993). "Hydrodynamic stability without eigenvalues". *Science* 261.5121, pp. 578–584. MR: 1229495 (cit. on p. 2162).
J. Vanneste (2001/02). "Nonlinear dynamics of anisotropic disturbances in plane Couette flow". *SIAM J. Appl. Math.* 62.3, pp. 924–944. MR: 1897729 (cit. on p. 2162).
J Vanneste, PJ Morrison, and T Warn (1998). "Strong echo effect and nonlinear transient growth in shear flows". *Physics of Fluids* 10.6, pp. 1398–1404 (cit. on p. 2162).
Akiva M. Yaglom (2012). *Hydrodynamic instability and transition to turbulence*. Vol. 100. Fluid Mechanics and its Applications. With a foreword by Uriel Frisch and a memorial note for Yaglom by Peter Bradshaw. Springer, Dordrecht, pp. xii+600. MR: 3185102 (cit. on pp. 2156, 2159).
JH Yu and CF Driscoll (2002). "Diocotron wave echoes in a pure electron plasma". *IEEE transactions on plasma science* 30.1, pp. 24–25 (cit. on p. 2162).
JH Yu, CF Driscoll, and TM O'Neil (2005). "Phase mixing and echoes in a pure electron plasma". *Physics of plasmas* 12.5, p. 055701 (cit. on p. 2162).

Received 2018-02-25.

Jacob Bedrossian
jacob@cscamm.umd.edu
jacob@math.umd.edu

Yu Deng
yudeng@cims.nyu.edu

Nader Masmoudi
masmoudi@cims.nyu.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2185–2210)

QUANTITATIVE ESTIMATES FOR ADVECTIVE EQUATION WITH DEGENERATE ANELASTIC CONSTRAINT

Didier Bresch and Pierre-Emmanuel Jabin

Abstract

In these proceedings we are interested in quantitative estimates for advective equations with an anelastic constraint in presence of vacuum. More precisely, we derive a quantitative stability estimate and obtain the existence of renormalized solutions. Our main objective is to show the flexibility of the method introduced recently by the authors for the compressible Navier-Stokes' system. This method seems to be well adapted in general to provide regularity estimates on the density of compressible transport equations with possible vacuum state and low regularity of the transport velocity field; the advective equation with degenerate anelastic constraint considered here is another good example of that. As a final application we obtain the existence of global renormalized solution to the so-called lake equation with possibly vanishing topography.

1 Introduction

New mathematical tools allowing to encode quantitative regularity estimates for the continuity equation written in Eulerian form have been recently developed by the authors [see Bresch and Jabin [2015] and Bresch and Jabin [2017a]] to answer two longstanding problems: Global existence of weak solutions for compressible Navier–Stokes with thermodynamically unstable pressure or with anisotropic viscous stress tensor. These articles provide a new point of view regarding the weak stability procedure (and more precisely on the space compactness for the density) in compressible fluid mechanics compared to what was developed mainly by P.–L. Lions and E. Feireisl *et al.*: See for example Feireisl [2004], Feireisl and Novotný [2009], Lions [1996].

D. Bresch is partially supported by the ANR-16-CE06-0011 project FRAISE and P.–E. Jabin is partially supported by NSF Grants 1312142 and 161453 and by NSF Grant RNMS (Ki-Net) 1107444.
MSC2010: primary 35B40; secondary 35B45, 35K35, 76Y05.
Keywords: Quantitative regularity estimates, advective equation, degenerate anelastic constraint, vacuum state, stability property, lake equations, vanishing topography, renormalized solutions.

In the present work, we want to show the flexibility of the method introduced in Bresch and Jabin [2015, 2017a] by focusing on quantitative stability estimates for advective equations with a vector field satisfying a degenerate anelastic constraint (linked to a non-negative scalar function). The method itself introduces weights which solve a dual equation and allow to propagate appropriately weighted norms on the initial solution. In a second time, a control on where those weights may vanish allow to deduce global and precise quantitative regularity estimates. For a more general introduction to the method, we refer interested readers to Bresch and Jabin [2017b].

The theory of existence and uniqueness for advection equations with rough force fields is now quite extensive, and we refer among others to the seminal articles DiPerna and Lions [1989], Ambrosio [2004], and to De Lellis [2007] and Ambrosio and Crippa [2014] for a general introduction to the topic. But quantitative regularity estimates were first derived on the Lagrangian formulation by G. Crippa and C. De Lellis in Crippa and De Lellis [2008]. The main idea is to identity the "good" trajectories where the flow has some regularity and then proving that those good trajectories have a large probability, which strongly inspired the Eulerian approach that we present here. This type of Lagrangian estimate is also used for example in Bohun, Bouchut, and Crippa [2016], Bouchut and Crippa [2013], Hauray, Le Bris, and Lions [2007] and Champagnat and Jabin [2010]. Note that quantitative regularity estimates for nonlinear continuity equations at the Eulerian level have also been introduced in Belgacem and Jabin [2013], Belgacem and Jabin [2016] using a nonlocal characterization of compactness in the spirit of Bourgain, Brezis, and Mironescu [2001]. PDE's with anelastic constraints are found in many different settings and we briefly refer for instance to Klein [2005], Lipps and Hemler [1982], Durran [1989], Wilhelmson and Ogura [1972], Masmoudi [2007], Feireisl, Málek, Novotný, and Straškraba [2008] in meteorology, to Bresch and Métivier [2006], Lacave, Nguyen, and Pausader [2014], and to Levermore, Oliver, and Titi [1996] for lakes and Perrin and Zatorska [2015], to Lannes [2017] for the dynamics of congestion or floating structures, to Donatelli and Feireisl [2017] for astrophysics and to Barré, Chiron, Goudon, and Masmoudi [2015] for asymptotic regime of strong electric fields to understand the importance to study PDEs with anelastic constraints especially the advective equation. As an application, we derive a new existence result for the so-called lake equation with possibly vanishing bathymetry which could vanish. The fact that we can obtain renormalized solutions in the vorticity formulation is in particular a significant improvement compared to previous results such as in Lacave, Nguyen, and Pausader [2014].

Let us now present more specifically the problem that we consider: Let Ω be a bounded smooth domain in $\mathbb{R}^d$ with $d = 1, 2$ or 3. We study the following advective equation

$$a\,(\partial_t \phi + u \cdot \nabla \phi) = 0 \text{ in } (0,T) \times \Omega \tag{1}$$

with a velocity field u such that

$$\text{(2)} \qquad \operatorname{div}(au) = 0 \text{ in } (0,T) \times \Omega, \qquad a\, u \cdot n|_{(0,T)\times\partial\Omega} = 0.$$

where a is a given non-negative scalar function which depends only on the space variable and is continuous on $\overline{\Omega}$. The initial condition is given by

$$\text{(3)} \qquad a\phi|_{t=0} = m_0 \text{ in } \Omega.$$

To avoid assuming any regularity on a, we still need to impose additional conditions on a: There exists a measurable non-negative function $\alpha(x)$, $r > 1$ and $q > p^*$ (with as usual $1/p^* + 1/p = 1$) s.t.

$$\text{(4)} \quad \alpha(x) \le a(x),\ A(\alpha, a) = \int_\Omega \left(|\nabla \alpha^{1/p^*}(x)|^q + a(x)\left(|\log \alpha(x)| + |\nabla \log \alpha(x)|^r\right)\right) dx < \infty.$$

Of course if $a \in W^{1,p}$ with $p > 1$ and $a\,|\log a| \in L^1$ then we could just choose $\alpha = a^k$ with $k \ge 1$. But (4) is far more general as in particular it does not require any regularity on a away from its vanishing set.

An example. To illustrate the condition (4), assume that there exists a Lipschitz domain $O \subset \Omega$ s.t. $a = 0$ on O^c and on O for some exponents $k,\ l > 0$

$$C^{-1} \min((d(x, \partial O))^k, 1) \le a(x) \le C \min((d(x,\partial O))^l, 1).$$

Then by taking $\alpha = \min((d(x,\partial O))^\theta, 1)$ with $\theta > p^*$, we immediately satisfy (4).

Let us now consider a velocity field u such that (with a slight abuse of notation as $\|u\|_a$ is not a norm)

$$\text{(5)} \quad \|u\|_a := \|u\|_{L^\infty_t L^p_a} + \int_0^T \int_\Omega a(x)\, |\nabla u(t,x)| \log\bigl(e + |\nabla(u(t,x))|\bigr)\, dx\, dt < \infty,$$

with $p > 1$ fixed and where the Lebesgue space $L^q_t L^p_a$ and more generally the Sobolev space $L^q_t W^{1,p}_a$ are defined by the norms

$$\|f\|_{L^q_t L^p_a} := \left\| \left(\int_\Omega |f|^p\, a(x)\, dx \right)^{1/p} \right\|_{L^q([0,\,T])} < \infty,$$

$$\|f\|_{L^q_t W^{1,p}_a} := \left\| \left(\int_\Omega (|f|^p + |\nabla f|^p)\, a(x)\, dx \right)^{1/p} \right\|_{L^q([0,\,T])} < \infty.$$

Because we do not have direct bounds on div u or even on ∇u as a may vanish, the standard theory of renormalized solutions cannot be applied to provide regularity (compactness of the solutions) or uniqueness. Concerning the boundary conditions on the velocity field, the anelastic constraint (2) and the integrability assumption on the velocity field allow to consider velocity fields satisfying the boundary condition in (2) in a weak sense, see for instance Lacave, Nguyen, and Pausader [2014].

We propose here to extend the method introduced in Bresch and Jabin [2017b] to this degenerate PDE system (1)–(3) through an appropriate three level weights control. This helps to encode quantitative stability estimates when approaching the degenerate constraint by a non-degenerate one: a standard procedure when you want to approximate a degenerate PDE. The conclusion will be existence of renormalized solution to the advective equations with degenerate anelastic constraint, as per

Theorem 1. *We have stability and existence of renormalized solutions:*
1. (Stability) For any C^1 sequences a_ε, α_ε, u_ε and a sequence of Lipschitz open domains Ω_ε with

- *a_ε is bounded from below, $\inf_{\Omega_\varepsilon} a_\varepsilon > 0$, and we have the divergence condition*

$$div\,(a_\varepsilon\, u_\varepsilon) = 0, \tag{6}$$

- *a_ε, α_ε, u_ε satisfy (2) and (4)-(5) uniformly in ε: $\sup_\varepsilon A(\alpha_\varepsilon, a_\varepsilon) + \sup_\varepsilon \|u_\varepsilon\|_{a_\varepsilon} < \infty$,*
- *Ω_ε converges to Ω for the Hausdorff distance on sets and $\|a_\varepsilon - a\|_{L^1(\Omega_\varepsilon\cap\Omega)} \to 0$ as $\varepsilon \to 0$,*

and for any sequence of initial data ϕ^0_ε uniformly bounded in $L^\infty(\mathbb{R}^d)$ and compact in $L^1(\mathbb{R}^d)$, consider the unique Lipschitz solution ϕ_ε to

$$a_\varepsilon\,(\partial_t \phi_\varepsilon + u_\varepsilon\cdot\nabla\phi_\varepsilon) = 0, \quad in\ \Omega_\varepsilon, \tag{7}$$

with boundary condition

$$a_\varepsilon\, u_\varepsilon \cdot n = 0, \quad on\ \partial\Omega_\varepsilon. \tag{8}$$

Then ϕ_ε is compact in $L^\infty_t L^2_{a_\varepsilon}$ and converges to a renormalized solution to (1) with (2).
2. (Existence) Let ϕ_0 be in $L^\infty(\Omega)$ and (a, α, u) satisfy (2) and the bounds (4) and (5). Then there exists a renormalized solution ϕ of (1) with initial data (3).

We present a possible strategy at the end of the article to use our techniques to prove that any weak solution is a renormalized solution and thus provide uniqueness of the solution; the full argument would however go beyond the limited scope of these proceedings.

The main ingredient to prove Theorem 1, is to obtain uniform regularity estimates on ε. This is done in two steps: First introducing appropriate weights in Section 2 and then propagating regularity in the next section. We can then construct a sequence of solutions ϕ^ε for the approximate coefficients a_ε and obtain the renormalized solution as the strong limit. We conclude the manuscript by showing the existence of global renormalized equations for the lake equations and presenting also a formal derivation of the model from compressible equation from Fluid Mechanics. Since our method is based on a doubling of variable argument, we make abundant use of notations like $u^x = u(t,x)$ to keep track of the physical variable (comparing u^x and u^y for $x \neq y$) whereas the value of the time variable is usually obvious.

2 Three-level weights procedure and properties

The estimates in this part hold for general coefficients with appropriate renormalized solutions but will later be used with the approximate coefficients a_ε, α_ε and the velocity u_ε.

As in Bresch and Jabin [ibid.], we introduce auxiliary equations that will help to identify the appropriate trajectories where the flow has some regularity. In this paper, we do it in three steps to control trajectories : where α is very small, where $|u|$ is large and where oscillations in the velocity field occur. More precisely, we define w_a solution to

$$\partial_t w_a + u \cdot \nabla w_a = -\gamma \frac{|u \cdot \nabla \alpha|}{\alpha} w_a, \qquad w_a|_{t=0} = (\alpha(x))^\gamma. \tag{9}$$

The weight w_a controls which trajectories can get close to points where α (and hence a) are very small. Next we introduce w_u solution to

$$\partial_t w_u + u \cdot \nabla w_u = -w_u\, |u(t,x)|\, \frac{1 + \int_0^t |\nabla u(s,x)|\, ds}{1 + \int_0^t |u(s,x)|\, ds}, \qquad w_u|_{t=0} = 1, \tag{10}$$

which controls trajectories going near points where $|u|$ is large. Finally we define our main weight, controlling oscillations in the velocity field

$$\partial_t w + u \cdot \nabla w = -D\ \ w, \qquad w|_{t=0} = 1, \tag{11}$$

with

$$D = \lambda \Big[\frac{M\,|\nabla(\alpha\, u)|}{\alpha} + (M\,|\nabla \alpha|(x))^\theta\, |u^x|^\theta + |\alpha^x|^{-\theta^*} \Big]$$

for some constants λ, θ and θ^* (chosen later on) respectively such that $\lambda > 0$, $1/\theta^* = 1 - 1/\theta$ with $p > \theta > 1$.

Observe that for general a, α and u only satisfying (4)-(5), we are at this point incapable of ensuring that there exist renormalized solutions to Eqs (9), (10), (11); in fact this would only follow from a first application of our method.

However assuming that such solutions exist, we can easily investigate their properties, summarized in the following

Lemma 2. *Assume that* (4) *holds and that* u *satisfies* (2) *and* (5). *Then*

- *Consider* w_a *a renormalized solution to* (9). *One has that*

(12)
$$0 \le w_a(t,x) \le (\alpha(x))^\gamma \le (a(x))^\gamma,$$
$$\int_\Omega a(x)\, w_u(t,x)\, |\log w_a(t,x)|\, dx$$
$$\le C\,\gamma\, \Big((1+\|u\|_{L^1_t L^p_a})\, \|\nabla \log\alpha\|_{L^r_a(\Omega)} + \|a\, \log\alpha\|_{L^1(\Omega)}\Big).$$

- *Consider* w_u *a renormalized solution to* (10). *One has that*

(13)
$$0 \le w_u(t,x) \le \frac{1}{1+\int_0^t |u(s,x)|\, ds}, \qquad \int_\Omega a(x)\, |\log w_u(t,x)|\, dx \le C_T\, \|u\|_a.$$

- *Finally consider* w *a renormalized solution to* (11). *One has that*

$$0 \le w(t,x) \le 1,$$
$$\int_\Omega a(x)\, w_a(t,x)\, |\log w(t,x)|\, dx \le C\,T + C\, \|u\|^\theta_{L^p_a}\, \|\nabla\alpha^{1/p^*}\|^\theta_{L^q} \tag{14}$$
$$+ C \int_0^T \int_\Omega a\, |\nabla u|\, \log(e+|\nabla u|)\, dx\, dt.$$

Lemma 2 in particular shows that $w_u > 0$ a-almost everywhere, that $w_a > 0$ a w_u-almost everywhere and finally that $w > 0$ a w_a-almost everywhere; and by the previous points, $w_a > 0$ and $w > 0$ a-almost everywhere as well.

Proof. 1) **Estimates on w_u.**

1-1) *Pointwise control.* Since $w = 1$ identically at $t = 0$ and $D \ge 0$, one trivially has that $0 \le w \le 1$. The other estimates are less straightforward and we start by proving them on w_u. Define

$$\varphi(t,x) = -\log(1+\int_0^t |u(s,x)|\, ds),$$

and notice that

$$\partial_t \varphi + u \cdot \nabla \varphi = -\frac{|u(t,x)|}{1+\int_0^t |u(s,x)|\,ds} - \frac{u(t,x)\cdot \int_0^t \nabla_x u(s,x)\cdot \frac{u(s,x)}{|u(s,x)|}\,ds}{1+\int_0^t |u(s,x)|\,ds},$$

while $\varphi(t=0,x)=0$. Therefore by (10), one has that

$$\partial_t \log w_u + u\cdot\nabla_x \log w_u \le \partial_t\varphi + u\cdot\nabla\varphi.$$

By the maximum principle since $\log w_u = \varphi$ at $t=0$, we have that $\log w_u \le \varphi$ and by taking the exponential

$$w_u \le e^{\varphi} = \frac{1}{1+\int_0^t |u(s,x)|\,ds}.$$

1-2) *A* log-*control on* w_u. Using again the equation (10), and since $\mathrm{div}(a\,u)=0$, we have that

$$\frac{d}{dt}\int_\Omega a(x)\,|\log w_u(t,x)|\,dx = \int_\Omega a(x)\,|u(t,x)|\,\frac{1+\int_0^t |\nabla u(s,x)|\,ds}{1+\int_0^t |u(s,x)|\,ds}\,dx.$$

Therefore by the definition of φ

$$\int_\Omega a(x)\,|\log w_u(t_0,x)|\,dx = -\int_0^{t_0}\int_\Omega \partial_t\varphi(t,x)\,\left(a+\int_0^t a(x)\,|\nabla u(s,x)|\,ds\right)\,dx\,dt.$$

Integrating by part in time

$$\begin{aligned}\int_\Omega a(x)\,|\log w_u(t_0,x)|\,dx =& -\int_0^t\int_\Omega a\,\partial_t\varphi(t,x)\\ &+\int_0^{t_0}\int_\Omega \varphi(t,x)\,a(x)\,|\nabla u(t,x)|\,dx\,dt\\ &-\int_\Omega a(x)\,\varphi(t_0,x)\int_0^{t_0} |\nabla u(s,x)|\,ds\,dx.\end{aligned}$$

Remark that the first term reads

$$0 \le -\int_0^t\int_\Omega a\,\partial_t\varphi(t,x) \le \|u\|_{L^1_t L^1_a}.$$

Note that the second term in the right-hand side is negative. For the last term, we use the well-known convex inequality, $x\,y \le x\log(e+x)+e^y$ for $x,\ y \ge 0$ to bound

$$
\begin{aligned}
-\int_\Omega a(x)\,\varphi(t_0,x)&\int_0^{t_0} |\nabla u(s,x)|\,ds\,dx\\
&\le \int_0^{t_0}\int_\Omega a(x)\,\Big(|\nabla u(s,x)|\,\log(e+|\nabla u(s,x)|)+e^{|\varphi(t_0,x)|}\Big)\,ds\,dx\\
&\le \int_0^{t_0}\int_\Omega a(x)\,\Big(|\nabla u(s,x)|\,\log(e+|\nabla u(s,x)|)+1+\int_0^{t_0}|u(r,x)|\,dr\Big)\,ds\,dx,
\end{aligned}
$$

again by the definition of φ. Hence

$$
\begin{aligned}
&\int_\Omega a(x)\,|\log w_u(t_0,x)|\,dx\\
&\le C_T\left(\|u\|_{L^1_t L^1_a}+\int_0^{t_0}\int_\Omega a(x)\,|\nabla u(s,x)|\,\log(e+|\nabla u(s,x)|)\,dx\,ds\right).
\end{aligned}
$$

2) **Estimates on w_a.**

2.1) *Pointwise control on w_a.* We now turn to the estimate on w_a. First note that

$$
\partial_t\alpha+u\cdot\nabla\alpha=u\cdot\nabla\alpha\ge -\frac{|u\cdot\nabla\alpha|}{\alpha}\,\alpha,
$$

and therefore, just as for w_u, by the maximum principle $\log w_a \le \gamma\log\alpha$ which leads to

$$
w_a(t,x)\le(\alpha(x))^\gamma
$$

and the other inequality as $\alpha\le a$.

1-2) *A* log*-control on w_a.* We also follow the same strategy to bound $|\log w_a|$ and obtain in a straightforward manner, using Eq. (10) on w_u, that

$$
\begin{aligned}
\int_\Omega a(x)\,w_u(t_0,x)\,|\log w_a(t_0,x)|\,dx \le&\gamma\int_0^{t_0}\int_\Omega a(x)\,w_u\,|u|\,|\nabla\log\alpha|\,dx\,dt\\
&+\int_\Omega a\,|\log w_a(t=0,x)|\,dx.
\end{aligned}
$$

From the initial data on w_a, $w_a(t=0,x)=(\alpha(x))^\gamma$, we have that

$$
\int_\Omega a\,|\log w_a(t=0,x)|\,dx\le\gamma\int_\Omega a\,|\log\alpha|\,dx.
$$

Furthermore since a and α do. not dependent on time, we also have that

$$\int_0^{t_0}\int_\Omega a(x)\, w_u\, |u|\, |\nabla \log\alpha|\, dx\, dt \le \int_\Omega a(x)\, |\nabla \log \alpha| \int_0^{t_0} \frac{|u(t,x)|\, dt}{1+\int_0^t |u(s,x)|\, ds}\, dx$$
$$= \int_\Omega a(x)\, |\nabla\log\alpha| \int_0^{t_0} \partial_t \log\left(1+\int_0^t |u(s,x)|\, ds\right) dt\, dx$$
$$= \int_\Omega a(x)\, |\nabla\log\alpha|\, \log\left(1+\int_0^{t_0} |u(s,x)|\, ds\right) dx$$

By bounding the log polynomially and a Hölder estimate, we deduce that

$$\int_\Omega a(x)\, w_u(t_0,x)\, |\log w_a(t_0,x)|\, dx\, dt \le \gamma\, \|\log\alpha\|_{L^1_a} + C_\mu\, \gamma\, \|u\|_{L^1_t L^p_a}\, \|\nabla\log\alpha\|_{L^{1+\mu}_a},$$

for any $\mu > 0$. Choosing μ s.t. $1+\mu \le r$ one has

(15)
$$\int_\Omega a(x)\, w_u(t_0,x)\, |\log w_a(t_0,x)|\, dx\, dt \le \gamma\, \|\log\alpha\|_{L^1_a} + \gamma\, \|u\|_{L^1_t L^p_a}\, \|\nabla\log\alpha\|_{L^r_a}.$$

2) **Estimates on w.** The point wise estimate on w is straightforward due to the damping term and the initial data. We now turn to the last estimate on $\log w$. Following similar calculations with Eqs. (11) and (9), we have that

$$\frac{d}{dt}\int_\Omega a(x)\, w_a(t,x)\, |\log w(t,x)|\, dx \le \int_\Omega a(x)\, w_a(t,x)\, D(t,x)\, dx.$$

Since $w_a(t,x) \le (\alpha(x))^\gamma$, if $\gamma \ge \theta^*$, one has from the definition of D in (11) that

$$\frac{d}{dt}\int_\Omega a(x)\, w_a(t,x)\, |\log w(t,x)|\, dx \le$$
$$\le \int_\Omega a(x)\left(M\, |\nabla(\alpha u)| + (M|\nabla\alpha|)^\theta |u|^\theta (\alpha)^\gamma + 1\right) dx.$$

We may simply bound

$$\int_\Omega a(x)\, (M|\nabla\alpha|)^\theta\, |u|^\theta (\alpha)^\gamma\, dx \le C\, \|\nabla\alpha\|^\theta_{L^q} \|u\|^\theta_{L^p_a},$$

with $q > \theta$ and recalling the maximal function is bounded on L^q as $q > 1$. For the other term, by the standard properties of the maximal function, one has that

$$\int_0^T\int_\Omega a(x)\, M|\nabla(\alpha\, u)|(t,x)\, dx\, dt \le C\int_0^T\int_\Omega M|\nabla(\alpha\, u)|(t,x)\, dx\, dt$$
$$\le C\int_0^T \||\nabla(\alpha\, u)(t,.)|\,\|_{\mathcal{H}^1}\, dt,$$

where $\mathcal{H}^1$ is the classical Hardy space. Since $|\nabla(\alpha u)|$ is always positive and Ω is bounded, this Hardy norm reduces to a $L\ \log L$ estimate

$$\|\,|\nabla(\alpha\, u)|\,\|_{\mathcal{H}^1} \sim C\,\left(\int_\Omega |\nabla(\alpha\, u)|\,\log(e+|\nabla(\alpha\, u)|)\,dx\right).$$

This is of course slightly non-optimal as we are losing possible cancellations in $\nabla(\alpha\, u)$, but necessary here if we want to keep positive weights. Of course since $\nabla(\alpha\, u) = u\nabla\alpha + \alpha\,\nabla u$, we have for example by the properties of the log and Hölder estimates that

$$\begin{aligned}\int_\Omega |\nabla(\alpha u)|\,\log(e+|\nabla(\alpha u)|))\,dx \leq & C\int_\Omega \alpha\,|\nabla u|\,\log(e+|\nabla u|)\,dx\\ &+C\,\|u\|^\theta_{L^p_\alpha}\,\|\nabla\alpha^{1/p^*}\|^\theta_{L^q},\end{aligned}$$

where one needs $q>p^*$. Therefore since $\alpha\leq a$, we finally find that

$$\begin{aligned}\int_\Omega a(x)\,w_a(t,x)\,|\log w(t,x)|\,dx &\leq C\,T + C\,\|u\|^\theta_{L^p_a}\,\|\nabla\alpha^{1/p^*}\|^\theta_{L^q}\\ &+C\int_0^T\int_\Omega a\,|\nabla u|\,\log(e+|\nabla u|)\,dx\,dt.\end{aligned}\tag{16}$$

□

3 Compactness and quantitative regularity estimates

We consider here any renormalized solution to our main equation (1) and prove that it satisfies some quantified uniform regularity. As in the previous section those estimates will be applied for our approximate coefficients a_ε, α_ϵ as at this time we have not yet obtained renormalized solution in the general case.

3.1 Regularity conditioned by the weights. The first step is to propagate an adhoc semi-norms constructed with the weights, namely

Proposition 3. *Assume that ϕ is a renormalized solution to the transport equation in advective form (1) with constraints (2). Let us define $\overline{a}$ corresponds to a on Ω and 0 on $\mathbb{R}^d\backslash\overline{\Omega}$. Assume as well that we have renormalized solutions w_a to (9), w_u to (10) and w to (11) with λ large enough. One has that for any h and for $q>p^*$*

$$\begin{aligned}\int_{\mathbb{R}^{2d}} \overline{a}^x\,\overline{a}^y\,&\frac{|\phi(t,x)-\phi(t,y)|}{(h+|x-y|)^d}\,w_a(t,x)\,w_u(t,x)\,w(t,x)\,w_a(t,x)\,w_u(t,y)\ \,w(t,y)\,dx\,dy\\ &\leq \int_{\mathbb{R}^{2d}} \overline{a}^x\,\overline{a}^y\,\frac{|\phi^0(x)-\phi^0(y)|}{(h+|x-y|)^d}\,dx\,dy\\ &\quad + C\,|\log h|^{1/2}\,\|\phi\|_{L^\infty}\,(\|u\|_a+\|u\|^\theta_a)\,(1+\|\nabla\alpha^{1/p^*}\|_{L^q})^\theta.\end{aligned}$$

Proof. We skip the bar on a to simplify calculations. Since ϕ is a renormalized solution, one has the non-linear identity

$$a^x a^y \Big[\partial_t |\phi^x - \phi^y| + u^x \cdot \nabla_x |\phi^x - \phi^y| + u^y \cdot \nabla_y |\phi^x - \phi^y|\Big] = 0.$$

Hence

$$\begin{aligned}\partial_t (a^x a^y |\phi^x - \phi^y|\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y) \\ + a^x a^y u^x \cdot \nabla_x (|\phi^x - \phi^y|\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y) \\ + a^x a^y u^y \cdot \nabla_y (|\phi^x - \phi^y|\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y) \\ \le -a^x a^y\, (D^x + D^y)\, |\phi^x - \phi^y|\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y.\end{aligned}$$

Multiplying by $(h + |x-y|)^{-d}$ and integrating by parts yields

$$\begin{aligned}\frac{d}{dt} \int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, dx\, dy \le \\ \le d \int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^{d+1}}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\ (u(t,x) - u(t,y)) \cdot \frac{x-y}{|x-y|}\, dx\, dy \\ - \int_{\Omega^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, (D^x + D^y)\, dx\, dy.\end{aligned}$$

As usual the main issue is the commutator estimate. As ∇u is only controlled when integrated against a, this is a more delicate issue. Indeed in principle $u^x - u^y$ involves the values of ∇u between x and y whereas we only have the values of a at x and y. It is the reason why we need to introduce α, which has some regularity, and proceed with the following decomposition

$$\begin{aligned}\left|(u(t,x) - u(t,y)) \cdot \frac{x-y}{|x-y|}\right| &\le \frac{1}{\alpha^x\, \alpha^y}\, \alpha^x\, \alpha^y\, |u(t,x) - u(t,y)| \\ &\le (\alpha^x)^{-1}\, (\alpha^y)^{-1}\, |\alpha^x\, u^x - \alpha^y\, u^y|\, \frac{\alpha^x + \alpha^y}{2} \\ &\quad + (\alpha^x)^{-1}\, (\alpha^y)^{-1}\, |\alpha^x - \alpha^y|\, \frac{\alpha^x\, |u^x| + \alpha^y\, |u^y|}{2}.\end{aligned}$$

By symmetry in x and y this leads to

$$
\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^{2d}} a^x a^y \,\frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, dx\, dy
&\le d\int_{\mathbb{R}^{2d}} a^x a^y\,\frac{|\phi^x-\phi^y|}{(h+|x-y|)^{d+1}}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\
&\qquad |\alpha^x\, u^x-\alpha^y\, u^y|\,\frac{\alpha^x+\alpha^y}{\alpha^x\,\alpha^y}\, dx\, dy \\
&\quad + d\int_{\mathbb{R}^{2d}} a^x a^y\,\frac{|\phi^x-\phi^y|}{(h+|x-y|)^{d+1}}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\
&\qquad |\alpha^x-\alpha^y|\,\frac{\alpha^x\,|u^x|+\alpha^y\,|u^y|}{\alpha^x\,\alpha^y}\, dx\, dy \\
&\quad - \int_{\Omega^{2d}} a^x a^y\,\frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^x\, w_u^y\,(D^x+D^y)\, dx\, dy.
\end{aligned}
\tag{17}
$$

We now appeal to the technical lemmas that have already been used in Bresch and Jabin [2017b] to control the difference u^x-u^y.

Lemma 4. *There exists $C>0$ s.t. for any $f\in W^{1,1}(\mathbb{R}^d)$, one has*

$$|f(x)-f(y)|\le C\,|x-y|\,(D_{|x-y|}f(x)+D_{|x-y|}f(y)),$$

where we denote

$$D_h f(x)=\frac{1}{h}\int_{|z|\le h}\frac{|\nabla f(x+z)|}{|z|^{d-1}}\,dz.$$

A full proof of such well known result can for instance be found in Champagnat and Jabin [2010] in a more general setting namely $f\in BV$. Through a simple dyadic decomposition, one may also immediately deduce that

$$D_h\, f(x)\le C\, M|\nabla f|(x), \tag{18}$$

where M denotes the usual maximal operator, and thus recovering the classical bound

$$|f(x)-f(y)|\le C\,|x-y|\,(M|\nabla f|(x)+M|\nabla f|(y)). \tag{19}$$

Applying Lemma 4 to Eq. (17), we find, by symmetry in x and y that

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, dx\, dy \\
\leq C \int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\
(D_{|x-y|}(\alpha\, u)(x)+D_{|x-y|}(\alpha\, u)(y))\, \frac{dx\, dy}{\alpha^x} \\
+ C \int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\
(D_{|x-y|}\alpha(x)+D_{|x-y|}\alpha(y))\, \frac{|u^x|}{\alpha^y}\, dx\, dy \\
- \int_{\Omega^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, (D^x+D^y)\, dx\, dy.
\end{aligned}$$

We recall the definition of the penalization D^x

$$D^x = \lambda\left(\frac{M\,|\nabla(\alpha\, u)|(x)}{\alpha^x} + (M\,|\nabla\alpha|(x))^\theta\, |u^x|^\theta + |\alpha^x|^{-\theta^*}\right)$$

with $\lambda > 0$ chosen large enough. Since $v\, w \leq v^\theta + w^{\theta^*}$, one has that

$$D_{|x-y|}\alpha(x)\, \frac{|u^x|}{\alpha^y} \leq (M\,|\nabla\alpha|(x))^\theta\, |u^x|^\theta + |\alpha^y|^{-\theta^*}.$$

By the bound (18) with some symmetry in x and y, and using λ large enough, we therefore obtain that

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, dx\, dy \\
\leq C \int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\
(D_{|x-y|}(\alpha\, u)(y)-D_{|x-y|}(\alpha\, u)(x))\, \frac{dx\, dy}{\alpha^x} \\
+ C \int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\
(D_{|x-y|}\alpha(y)-D_{|x-y|}\alpha(x))\, \frac{|u^x|}{\alpha^y}\, dx\, dy.
\end{aligned} \tag{20}$$

Recalling now Lemma 2, we have that $w_a^x \le \alpha^x$. Therefore,

$$\int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y$$

$$(D_{|x-y|}(\alpha\, u)(y) - D_{|x-y|}(\alpha\, u)(x))\, \frac{dx\, dy}{\alpha^x} \le$$

$$\le C\, \|\phi\|_{L^\infty} \int_{\mathbb{R}^{2d}} \frac{|D_{|x-y|}(\alpha\, u)(y) - D_{|x-y|}(\alpha\, u)(x)|}{(h+|x-y|)^d}\, dx\, dy$$

$$\le C\, \|\phi\|_{L^\infty} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} \int_0^R \big|D_\rho(\alpha\, u)(x+\rho\, w) - D_\rho(\alpha\, u)(x)\big|\, \frac{d\rho}{h+\rho}\, d\rho\, dx\, dw,$$

by a direct change of variables to polar coordinates in $y - x$ and where R is the diameter of Ω. This leads to a square function type of estimates as by Cauchy-Schwartz

$$\int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y$$

$$(D_{|x-y|}(\alpha\, u)(y) - D_{|x-y|}(\alpha\, u)(x))\, \frac{dx\, dy}{(\alpha^x)^\theta} \le$$

$$\le C\, \|\phi\|_{L^\infty} \int_{\mathbb{S}^{d-1}} |\log h|^{1/2}$$

$$\int_{\mathbb{R}^d} \left(\int_0^R \big|D_\rho(\alpha\, u)(x+\rho\, w) - D_\rho(\alpha\, u)(x)\big|^2\, \frac{d\rho}{h+\rho}\right)^{1/2} dx\, dw.$$

We now recall the classical estimate (see for example the remark on page 159 in Stein [1993])

Lemma 5. *For any $1 < p < \infty$, any family L_ρ of kernels satisfying for some $s > 0$*

(21)
$$\int L_\rho = 0, \quad \sup_\rho\, (\|L_\rho\|_{L^1} + \rho^s\, \|L_\rho\|_{W^{s,1}}) \le C_L, \quad \sup_\rho \rho^{-s} \int |z|^s\, |L_\rho(z)|\, dz \le C_L.$$

Then there exists $C > 0$ depending only on C_L above s.t. for any f in the Hardy space $\mathcal{H}^1(\Omega)$

$$\int_{\mathbb{R}^d} \left(\int_0^1 |L_\rho \star f(x)|^2\, \frac{d\rho}{h+\rho}\right)^{1/2} dx \le C\, \|f\|_{\mathcal{H}^1}, \tag{22}$$

whereas if $f \in L^p$ with $1 < p < \infty$

$$\int_{\mathbb{R}^d} \left(\int_0^1 |L_\rho \star f(x)|^2\, \frac{d\rho}{h+\rho}\right)^{p/2} dx \le C_p\, \|f\|_{L^p}^p. \tag{23}$$

Observe that obviously

$$D_\rho f = \bar{L}_\rho \star |\nabla f|, \quad \bar{L}_\rho(x) = \frac{1}{\rho\, |x|^{d-1}}\, \mathbb{I}_{|x|\leq \rho} = \rho^{-d}\, \bar{L}(x/\rho)$$

with $\bar{L}(x) = \frac{1}{|x|^{d-1}}\, \mathbb{I}_{|x|\leq 1}$. Hence defining $L_\rho(x) = \bar{L}_\rho(x) - \bar{L}_\rho(x+\rho\, w)$, we can easily check that L_ρ satisfies the assumptions of Lemma 5. This proves that

$$\int_{\mathbb{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^{d+1}}\, w^x\, w_a^x\, w^y\, w_a^y\, (D_{|x-y|}(\alpha\, u)(y) - D_{|x-y|}(\alpha\, u)(x))\, \frac{dx\, dy}{\alpha^x} \\ \leq C\, \|\phi\|_{L^\infty}\, |\log h|^{1/2}\, \|\, |\nabla(\alpha\, u)|\, \|_{\mathcal{H}^1(\Omega)}.$$

We now follow the exact same steps as for the bound at the end of the proof of Lemma 2. Note that here it would be easier to use the cancellations in $\nabla(\alpha u)$ by being more precise in Lemma 4 and using an exact representation instead of a bound. For simplicity though, here we have kept the more direct version of Lemma 4. Hence we have that

$$\int_\Omega |\nabla(\alpha u)|\, \log(e + |\nabla(\alpha u)|)\, dx \leq C\, \int_\Omega a\, |\nabla u|\, \log(e+|\nabla u|)\, dx \\ + C\, \|u\|_{L^p_a}^\theta\, \|\nabla \alpha^{1/p^*}\|_{L^q}^\theta,$$

where again one needs $q > p^*$. This lets us conclude that

$$\begin{aligned} &\int_0^T \int_{\mathbf{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w^y\, w_a^y \\ &\qquad (D_{|x-y|}(\alpha\, u)(y) - D_{|x-y|}(\alpha\, u)(x))\, \frac{dx\, dy\, dt}{\alpha^x} \leq \qquad (24) \\ &\leq C_T\, \|\phi\|_{L^\infty}\, |\log h|^{1/2}\, [\|u\|_a + \|u\|_{L^p_a}^\theta \|\nabla \alpha^{1/p^*}\|_{L^q}^\theta] \end{aligned}$$

We apply the same strategy to the other term in the bound (20). We again start using that $w_a^y \leq \alpha^y$ to obtain that

$$\begin{aligned} &\int_0^T \int_{\mathbf{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y \\ &\qquad (D_{|x-y|}\alpha(y) - D_{|x-y|}\alpha(x))\, \frac{|u^x|\, dx\, dy\, dt}{\alpha^y} \leq \\ &\leq \|\phi\|_{L^\infty} \int_{\mathbf{R}^{2d}} \frac{|D_{|x-y|}\alpha(y) - D_{|x-y|}\alpha(x)|}{(h+|x-y|)^d} \int_0^T w_u^x\, a^x\, |u^x|\, dt\, dx\, dy, \end{aligned}$$

since α is independent of time. By Lemma 2, $w_u \le 1/(1+\int_0^t |u(s,x)|\,ds$ and hence

$$\int_0^T w_u(t,x)\,|u(t,x)|\,dt \le \int_0^T \frac{|u(t,x)|}{1+\int_0^t |u(s,x)|\,ds}\,dt$$
$$= \int_0^T \partial_t \log\left(1+\int_0^t |u(s,x)|\,ds\right)\,dt$$
$$= \log\left(1+\int_0^T |u(s,x)|\,ds\right).$$

Choose now any $\mu > 0$ and bound

$$\log\left(1+\int_0^T |u(s,x)|\,ds\right) \le C_\mu \left(1+\int_0^T |u(s,x)|\,ds\right)^{\mu/(1+\mu)},$$

so that by Hölder since $1-1/(1+\mu) = \mu/(1+\mu)$

$$\int_0^T \int_{\mathbf{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y$$
$$(D_{|x-y|}\alpha(y) - D_{|x-y|}\alpha(x))\,\frac{|u^x|\,dx\,dy\,dt}{\alpha^y} \le$$
$$\le C_\mu\, \|\phi\|_{L^\infty}\, \|u\|_{L^1_t L^1_a}\, |\log h|^{\mu/(1+\mu)}$$
$$\left(\int_{\mathbf{R}^{2d}} \frac{|D_{|x-y|}\alpha(y) - D_{|x-y|}\alpha(x)|^{1+\mu}}{(h+|x-y|)^d}\,dx\,dy\right)^{1/(1+\mu)}.$$

We can now apply Lemma 5 for $f \in L^p$, and find similarly that

$$\int_{\mathbf{R}^{2d}} \frac{|D_{|x-y|}\alpha(y) - D_{|x-y|}\alpha(x)|^{1+\mu}}{(h+|x-y|)^d}\,dx\,dy \le C_\mu\, |\log h|^{(1-\mu)/2}\, \|\nabla\alpha\|_{L^{1+\mu}}^{1+\mu}.$$

This leads to

$$\int_0^T \int_{\mathbf{R}^{2d}} a^x a^y \frac{|\phi^x-\phi^y|}{(h+|x-y|)^{d+1}}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y$$
$$(D_{|x-y|}\alpha(y) - D_{|x-y|}\alpha(x))\,\frac{|u^x|\,dx\,dy\,dt}{\alpha^y} \tag{25}$$
$$\le C_\mu\, \|\phi\|_{L^\infty}\, \|u\|_{L^1_t L^1_a}\, |\log h|^{1/2}\, \|\nabla\alpha\|_{L^{1+\mu}}.$$

Choosing μ small with $1 + \mu \le q$ and combining (25) with (24) in (20), we finally conclude that

$$\frac{d}{dt}\int_{\mathbf{R}^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, w^x\, w_a^x\, w_u^x\, w^y\, w_a^y\, w_u^y\, dx\, dy \le$$
$$\le C\,|\log h|^{1/2}\,\|\phi\|_{L^\infty}\,(\|u\|_a + \|u\|_a^\theta)\,(1+\|\nabla\alpha^{1/p^*}\|_{L^q})^\theta,$$

thus proving the proposition. □

3.2 Our explicit regularity estimate. By using a straightforward interpolation argument thanks to the previous controls obtained on the different weights w_u, w_a, w, we can now state our main result

Theorem 6. *Assume that* (a, α) *satisfy* (4) *and that* (2) *and* (5) *hold for* u. *Assume as well that we have renormalized solutions* w_a *to* (9), w_u *to* (10) *and* w *to* (11). *Consider now any renormalized solution to* (1) *and denote*

$$\|\phi^0\|_h = \frac{1}{|\log h|}\int_{\Omega^{2d}} a^x a^y \frac{|\phi^0(x)-\phi^0(y)|}{(h+|x-y|)^d}\, dx\, dy.$$

Then

$$\|\phi\|_h = \frac{1}{|\log h|}\int_{\Omega^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, dx\, dy \le \frac{C}{|\log(\|\phi^0\|_h + |\log h|^{-1/2})|^{1/2}},$$

for some constant $C > 0$ *depending only on the bounds on* $\|\alpha\|_{L^\infty(\Omega)}$, $\|u\|_a$ *and* $\|\phi\|_{L^\infty((0,T)\times\Omega)}$.

Proof. The proof relies on a appropriate decomposition of the domain playing with sets constructed using intersection of the set $\{x,\ y \mid w_u(t,x) > \eta,\ w_u(t,y) > \eta\}$ or its complementary set with the set $\{x,\ y \mid w_a(t,x) > \eta',\ w_a(t,y) > \eta'\}$ and its complementary set and with the set $\{x,\ y \mid w(t,x) > \eta'',\ w(t,y) > \eta''\}$ and its complementary. More precisely, we write

$$\|\phi\|_h = \int_{\Omega^{2d}} \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, a^x\, a^y\, dx\, dy = \sum_{i=1}^4 \int_{I_i} \frac{|\phi^x - \phi^y|}{(h+|x-y|)^d}\, a^x\, a^y\, dx\, dy$$
$$= \sum_{i=1}^4 J_j$$

with

$$I_1 = \{x,\ y \mid w_u(t,x) < \eta \text{ or } w_u(t,y) < \eta\},$$

$$I_2 = \{x,\ y \mid w_u(t,x) > \eta \text{ and } w_u(t,y) > \eta\} \cap \{x,\ y \mid w_a(t,x) < \eta' \text{ or } w_a(t,y) < \eta'\}$$

and denoting

$$I = \{x,\ y \mid w_u(t,x) > \eta \text{ and } w_u(t,y) > \eta\} \cap \{x,\ y \mid w_a(t,x) > \eta' \text{ and } w_a(t,y) > \eta'\},$$

with

$$I_3 = I \cap \{x,\ y \mid w(t,x) < \eta'' \text{ or } w(t,y) < \eta''\},$$

and

$$I_4 = I \cap \{x,\ y \mid w(t,x) > \eta'' \text{ and } w(t,y) > \eta''\}.$$

Note that it is straightforward that

$$0 \le J_4 \le \frac{1}{\eta^2 \eta'^2 \eta''^2} \int_{\Omega^{2d}} a^x\, a^y\, \frac{|\phi^x - \phi^y|}{(h + |x-y|)^d} w_a^x w_a^y w_u^x w_u^y w^x w^y \, dx\, dy.$$

Remark now that by symmetry, J_1 is bounded by

$$0 \le J_1 \le |\log h| \int_{x,\ w_u(t,x) \le \eta} a^x (K_h \star a|\phi(t,x)| + K_h \star |a\phi|)\, dx,$$

where $K_h(x) = (h + |x|)^{-d}/|\log h|$ so $\|K_h\|_{L^1} = 1$. By Hölder estimate

$$\int_{x,\ w_u(t,x) \le \eta} a^x (K_k \star a|\phi(t,x)| + K_h \star |a\phi|)\, dx \le C\|\phi\|_{L^\infty} \int_{x,\ w_u(t,x) \le \eta} a\, dx.$$

Now it suffices to note that

$$\int_{x,\ w_u(t,x) \le \eta} a\, dx \le \frac{1}{|\log \eta|} \int_{x,\ w_u(t,x) \le \eta} |\log w_u(t,x)|\, a(t,x)\, dx \le \frac{C}{|\log \eta|} \|u\|_a$$

to get an appropriate control. Similarly we get using properties of w_u and w_a

$$J_2 \le \frac{C|\log h|}{\eta |\log \eta'|} \|\phi\|_{L^\infty} \int_\Omega a^x w_u^x |\log w_a^x|$$

We end the proof with the same kind of estimate on J_3 using properties of w_a and w, namely

$$J_3 \le \frac{C|\log h|}{\eta' |\log \eta''|} \|\phi\|_{L^\infty} \int_\Omega a^x w_a^x |\log w^x|$$

Now using the bounds on $a w_u |\log w_a|$ and $a w_a |\log w|$ and the uniform bounds on u and α, and using Proposition 3 we get

$$\sup_{t\in[0,T]} \frac{1}{|\log h|} \int_{\Omega^{2d}} a^x a^y \frac{|\phi^x - \phi^y|}{(h + |x-y|)^d}\, dx\, dy \le$$
$$\le \frac{C}{\eta^2 \eta'^2 \eta''^2} \Big[\|\phi^0\|_h + |\log h|^{-1/2} \Big] + C\Big[\frac{1}{|\log \eta|} + \frac{1}{\eta |\log \eta'|} + \frac{1}{\eta' |\log \eta''|} \Big].$$

Optimizing in η, η', η'' (by choosing η in function of η' and η' in function of η'' and finally η'' in function of α and $|\log h|^{-1/2}$) we get the conclusion.

□

4 Stability and existence of renormalized solutions: Proof of Theorem 1

4.1 Stability of renormalized solutions. Assume that we have been given sequences $a_\varepsilon, \alpha_\varepsilon$ and u_ε on a set Ω_ε which satisfy the assumptions specified in Theorem 1.

Since all terms are smooth, Eq. (7) has a unique Lipschitz solution ϕ_ε for any given initial data $\phi_\varepsilon^0 \in L^\infty(\Omega_\varepsilon)$. This solution is then obviously automatically renormalized. For the same reason we also trivially have solutions w_a to Eq. (9) with α_ε and u_ε and similarly for Eqs. (10) and (11). Of course while our solutions are smooth for a fixed ε, the main point is to derive and use uniform in ε bounds to obtain appropriate limits.

First define $\bar{a}_\varepsilon = a_\varepsilon$ on Ω_ε and extended by 0 on the whole of $\mathbb{R}^d$. Proceed similarly to define at the limit $\bar{a}$. Since a_ε is uniformly in L^∞, we can replace the convergence $\|a_\varepsilon - a\|_{L^1(\Omega_\varepsilon\cap\Omega)} \to 0$ and $\Omega_\varepsilon \to \Omega$ in Hausdorff distance by the simple

$$\|\bar{a}_\varepsilon - \bar{a}\|_{L^1(\mathbb{R}^d)} \longrightarrow 0.$$

From the uniform $L^\infty_t L^p_{a_\varepsilon}$ estimate for u_ε provided by (5) and $\sup_\varepsilon \|u_\varepsilon\|_{a_\varepsilon} < \infty$, we can extract a weak limit of $\bar{a}_\varepsilon^{1/p}\, u_\varepsilon$ in the whole space $\mathbb{R}^d$ and from the strong convergence of $\bar{a}_\varepsilon$, identify the limit as $\bar{a}\, u$ for some $u \in L^\infty_t L^p_a$:

$$\bar{a}_\varepsilon^{1/p}\, u_\varepsilon \longrightarrow \bar{a}^{1/p}\, u \quad \text{in } w-*\; L^\infty_t L^p(\mathbb{R}^d), \qquad u \in L^\infty_t L^p_a,$$

while for simplicity we still denote the extracted subsequence with ε. Since $\sup_\varepsilon \|\phi_\varepsilon^0\|_{L^\infty(\mathbb{R}^d)} < \infty$ then through renormalization $\sup_\varepsilon \|\phi_\varepsilon\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^d)} < \infty$, we may also extract a converging subsequence

$$\phi_\varepsilon \longrightarrow \phi \quad \text{in } w-*\; L^\infty(\mathbb{R}_+\times\mathbb{R}^d).$$

For any $\chi \in W^{1,\infty}(\mathbb{R})$ with $\chi(0) = 0$, $\chi(\phi_\varepsilon)$ still solves (7) by the chain rule for smooth functions. Choosing any test function $\psi \in C_c^\infty(\mathbb{R}^d)$, we deduce from (7) with the divergence condition (6) and the boundary conditions (8) the weak formulation

$$\frac{d}{dt}\int_{\Omega_\varepsilon} \chi(\phi_\varepsilon)\, a_\varepsilon\, \psi(x)\, dx - \int_{\Omega_\varepsilon} \chi(\phi_\varepsilon)\, a_\varepsilon\, u_\varepsilon \cdot \nabla_x \psi\, dx = 0.$$

The previous definition of $\bar{a}_\varepsilon$ and $\bar{\phi}_\varepsilon$ actually implies that this weak formulation is equivalent to the formulation in the whole space

$$\frac{d}{dt}\int_{\mathbb{R}^d} \chi(\phi_\varepsilon)\,\bar{a}_\varepsilon\,\psi(x)\,dx - \int_{\mathbb{R}^d} \chi(\phi_\varepsilon)\,\bar{a}_\varepsilon\,u_\varepsilon\cdot\nabla_x\psi\,dx = 0, \tag{26}$$

which is much simpler to use since the domain is now fixed. In that sense (26) implies the boundary condition (8) on $\partial\Omega_\varepsilon$ if one imposes that $\bar{a}_\varepsilon = 0$ out of Ω_ε. It is straightforward to check that $\bar{a} = 0$ out of Ω at the limit. Thus to prove that ϕ is a renormalized solution to (1) with (2) on the limiting set Ω, it is now enough to pass to the limit in (26).

Let us now first prove compactness in space on $\chi(\phi_\varepsilon)$ for any smooth function χ. This is exactly where our approach proves its use: We have all required assumptions to apply Theorem 6 and deduce from the compactness of ϕ_ε^0 and a_ε that

$$\limsup_{h\to 0}\frac{1}{|\log h|}\sup_\varepsilon\sup_t\int_{\mathbb{R}^{2d}} \bar{a}_\varepsilon^x\,\bar{a}_\varepsilon^y\,\frac{|\phi_\varepsilon^x-\phi_\varepsilon^y|}{(|x-y|+h)^d}\,dxdy \longrightarrow 0. \tag{27}$$

Note now that

$$a_\varepsilon^x a_\varepsilon^y(\phi_\varepsilon^x-\phi_\varepsilon^y) = (a_\varepsilon^x\phi_\varepsilon^x - a_\varepsilon^y\phi_\varepsilon^y)(a_\varepsilon^y+a_\varepsilon^x)/2 + (a_\varepsilon^y-a_\varepsilon^x)(a_\varepsilon^y\phi_\varepsilon^y+a_\varepsilon^x\phi_\varepsilon^x)/2$$

and that $|(a_\varepsilon^x\phi_\varepsilon^x - a_\varepsilon^y\phi_\varepsilon^y)| \le C\,(a_\varepsilon^x+a_\varepsilon^y)$ then get

$$|a_\varepsilon^x\phi_\varepsilon^x - a_\varepsilon^y\phi_\varepsilon^y|^2 \le C\,(a_\varepsilon^x a_\varepsilon^y|\phi_\varepsilon^x-\phi_\varepsilon^y| + |a_\varepsilon^y - a_\varepsilon^x|).$$

and therefore using (27) and compactness on a_ε, by the Rellich criterion this implies locally in space compactness of $a_\varepsilon\phi_\varepsilon$. Using the same procedure, it is possible to prove space compactness of $a_\varepsilon\chi(\phi_\varepsilon)$. We get compactness (in space and time) on $a_\varepsilon\chi(\phi_\varepsilon)$ using the renormalized equation which provides a control on $\partial_t(a_\varepsilon\chi(\phi_\varepsilon))$ allowing to use Aubin-Lions Lemma. Thus, up to a subsequence, we deduce that $a_\varepsilon\chi(\phi_\varepsilon)$ converges almost everywhere and thus $a_\varepsilon^{1-1/p}\chi(\phi_\varepsilon)$ converges almost everywhere using the compactness on a_ε. As ϕ_ε is uniformly bounded and therefore $\chi(\phi_\varepsilon)$ also, we get compactness of $a_\varepsilon^{1-1/p}\chi(\phi_\varepsilon)$. To conclude we just have to write $\chi(\phi_\varepsilon)a_\varepsilon u_\varepsilon = a_\varepsilon^{1-1/p}\chi(\phi_\varepsilon)\,a_\varepsilon^{1/p}u_\varepsilon$ and use the weak-star convergence of $a_\varepsilon^{1/p}u_\varepsilon$ in $L_t^\infty L_x^p$ and the strong convergence of $a_\varepsilon^{1-1/p}\chi(\phi_\varepsilon)$ in $L_t^1 L_x^q$ where $1/q+1/p=1$.

4.2 Existence of renormalized solutions. To obtain existence of renormalized solutions through a stability argument, it only remains to be able construct a sequence of approximations on which we may apply the previous stability argument.

In our case, given a, α and u which satisfy (2), (4) and (5), the first question is whether we can construct smooth a_ε, α_ε and u_ε which still satisfy the previous estimates uniformly in ε and where a_ε is bounded from below on Ω.

First define $\tilde\Omega_\varepsilon = \{\alpha > \varepsilon\}$. On $\tilde\Omega_\varepsilon$, one has that $a \geq \alpha > \varepsilon$; hence by (4), α belongs to a Sobolev space on a neighborhood of $\tilde\Omega_\varepsilon$ so that the boundary of $\tilde\Omega_\varepsilon$ is Lipschitz.

Define $\tilde a_\varepsilon = a$ on $\tilde\Omega_\varepsilon$ and $a = \varepsilon$ on $\Omega \setminus \tilde\Omega_\varepsilon$. Hence $\tilde a_\varepsilon$ may be discontinuous. Define similarly $\tilde\alpha_\varepsilon = \alpha$ on $\tilde\Omega_\varepsilon$ and ε outside. By the definition of $\tilde\Omega_\varepsilon$, $\tilde\alpha_\varepsilon$ does not jump on $\partial\tilde\Omega_\varepsilon$.

Note that $\tilde\alpha_\varepsilon$ satisfies (4) uniformly in ε, *i.e.* $\sup_\varepsilon A(\tilde\alpha_\varepsilon, \tilde a_\varepsilon) < \infty$; as for example

$$\tilde a_\varepsilon \, |\nabla \log \tilde\alpha_\varepsilon|^r = a \, |\nabla \log \alpha|^r \, \mathbb{I}_{\tilde\Omega_\varepsilon}.$$

Choose now a smooth and non-negative function χ s.t. $\chi(\xi/\varepsilon)$ is a good approximation of the Heaviside function with in particular $\chi(\xi/\varepsilon) = 0$ if $\xi \leq \varepsilon$ and $\chi(\xi/\varepsilon) = 1$ if $\xi \geq 2\,\varepsilon$.

Define then $u_{\varepsilon,L} = \frac{u}{1+|u|/L}\,\chi(\alpha/\varepsilon)$. And observe that

$$\begin{aligned}\nabla u_{\varepsilon,L} &= \frac{\nabla u}{1+|u|/L}\,\chi(\alpha/\varepsilon) - \frac{u}{L\,(1+|u|/L)^2} \otimes \nabla u \cdot \frac{u}{|u|} \\ &\qquad + \frac{u}{1+|u|/L} \otimes \nabla \log \alpha \, \frac{\alpha}{\varepsilon}\,\chi'(\alpha/\varepsilon) \\ &= \frac{\nabla u}{1+|u|/L}\,\chi(\alpha/\varepsilon) + U_{\varepsilon,L} + D_{\varepsilon,L}.\end{aligned}$$

Since $\frac{\alpha}{\varepsilon}\,\chi'(\alpha/\varepsilon)$ is bounded uniformly and $\nabla \log \alpha \in L^r_a$, one has that $\|D_{\varepsilon,L}\|_{L^r_a} \leq C\,L$ for some given constant C independent of ε and L. But note that $\nabla \log \alpha$ is independent of ε and L and hence equi-integrable in L^r_a while $\frac{\alpha}{\varepsilon}\,\chi'(\alpha/\varepsilon)$ converges to 0 in L^1 as $\varepsilon \to 0$. Hence for a fixed L, $D_{\varepsilon,L} \to 0$ as $\varepsilon \to 0$ for L fixed.

Therefore we can connect L and ε and choose L_ε s.t. $\|D_{\varepsilon,L}\|_{L^r_a} \to 0$ as $\varepsilon \to 0$. By the same type of equi-integrability arguments, we can show that

$$\int_\Omega a\,|U_{\varepsilon,L_\varepsilon}|\,\log(1+|U_{\varepsilon,L_\varepsilon}|)\,dx \longrightarrow 0, \quad \text{as } \varepsilon \to 0.$$

As a consequence $u_{\varepsilon,L_\varepsilon}$ still satisfies $\sup_\varepsilon \|u_{\varepsilon,L_\varepsilon}\|_a < \infty$. Hence since $u_{\varepsilon,L_\varepsilon}$ vanishes outside of $\tilde\Omega_\varepsilon$, it satisfies $\sup_\varepsilon \|u_{\varepsilon,L_\varepsilon}\|_{\tilde a_\varepsilon} < \infty$. We still need to correct the divergence and for this we solve the following elliptic equation

$$\operatorname{div}(a\,\nabla V_\varepsilon) = -a\,\operatorname{Tr}(U_{\varepsilon,L_\varepsilon} + D_{\varepsilon,L_\varepsilon}) \quad \text{in } \tilde\Omega_\varepsilon, \qquad V_\varepsilon = 0 \text{ on } \partial\tilde\Omega_\varepsilon.$$

Since a is bounded from below in $\tilde\Omega_\varepsilon$ and $\partial\Omega_\varepsilon$ is Lipschitz, this equation is well posed and we can extend V_ε to all Ω by taking $V_\varepsilon = 0$ in $\Omega \setminus \tilde\Omega_\varepsilon$. Furthermore by the previous bounds on $D_{\varepsilon,L_\varepsilon}$ and $L_{\varepsilon,L_\varepsilon}$, we have that V_ε converges to 0 in $W^{1,q}_{\tilde a_\varepsilon}$ for some $q > 1$. In dimension 2 for instance, the energy inequality would directly give this in $H^1_{\tilde a_\varepsilon}$.

We finally define $\tilde u_\varepsilon = u_{\varepsilon,L_\varepsilon} + \nabla V_\varepsilon$. From the construction, Eq. (6) holds for $\tilde a_\varepsilon$ and $\tilde u_\varepsilon$. The bounds in (4) and (5) are also satisfied uniformly in ε: $\sup_\varepsilon A(\tilde\alpha_\varepsilon, \tilde a_\varepsilon) + \sup_\varepsilon \|\tilde u_\varepsilon\|_{\tilde a_\varepsilon} < \infty$. Finally $\tilde a_\varepsilon$, $\tilde\alpha_\varepsilon$ and $\tilde u_\varepsilon$ all converge strongly.

Of course those coefficients are not yet smooth but this last step is the easiest and we only sketch it. By standard Sobolev approximation since $\tilde{a}_\varepsilon$ is now bounded from below, one may find α_ε and $\bar{u}_\varepsilon$ in $C^\infty(\Omega)$ but close to $\tilde{\alpha}_\varepsilon$ and $\tilde{u}_\varepsilon$ in the corresponding Sobolev spaces so that (4) and (5) still hold with weight $\tilde{a}_\varepsilon$ uniformly in ε.

One then approximates $\tilde{a}_\varepsilon$ by $a_\varepsilon \in C^\infty(\Omega)$, uniformly bounded and with $a_\varepsilon \geq \varepsilon$. In addition it is possible to choose $\|\tilde{a}_\varepsilon - a_\varepsilon\|_{L^1}$ small enough to obtain the uniform bounds (4) and (5): $\sup_\varepsilon A(\alpha_\varepsilon, a_\varepsilon) + \sup_\varepsilon \|u_\varepsilon\|_{a_\varepsilon} < \infty$.

We finally correct $u_\varepsilon = \bar{u}_\varepsilon + \nabla \bar{V}_\varepsilon$ as before to satisfy the divergence condition (6).

Once those approximated coefficients are constructed, we may directly apply our stability estimates to obtain at the limit a renormalized solution ϕ.

4.3 Toward the uniqueness of weak solutions to (1). We conclude this section by briefly sketching a possible strategy to obtain the uniqueness of Eq. (1) by proving, as in the classical argument, that all weak solutions are also renormalized.

Since DiPerna and Lions [1989] this is usually performed by convolving Eq. (1) for any weak solution ϕ by some smooth kernel ρ_ε and showing that $\rho_\varepsilon(a\,\phi)$ still solves (1) with a right-hand side that is vanishing in L^1. This commutator estimate would require here that

$$\int_\Omega (u(t,x) - u(t,y)) \cdot \nabla \rho_\varepsilon(x-y)\, a(y)\, \phi(t,y)\, dy \longrightarrow 0 \quad \text{in } L^1_t L^1_x \text{as } \varepsilon \to 0. \tag{28}$$

One can then typically conclude by using Sobolev bounds on a. But since there is no $a(y)$ factor in the above integral and we only control ∇u in L^1_a, this cannot work here.

A second issue arises since (28) also usually requires a control on $\operatorname{div} u$ which is again unavailable.

Instead we would propose the following approach:

- Through a stability argument, obtain the existence of renormalized solutions to (9) and (10).
- Show that the commutator estimate (28) for $\phi = w_a\, w_u$ holds by using in particular that $w_a \leq a^\gamma$. The exact calculations here should be reminiscent of what were in essence other commutator estimates in the proof of Proposition 3.
- For any weak solution ϕ, use the previous point to prove that $\phi\, w_a\, w_u$ is also a solution to (1) with the corresponding added right-hand side from (9) and (10).
- Prove a commutator estimate like (28) but where ϕ is replaced by $\phi\, w_a\, w_u$.

There are obvious technical difficulties at each step and for this reason implementing such a strategy is beyond the limited scope of these proceedings.

5 An anelastic compressible equation coming from fluid mechanics

Let us present in this subsection a PDEs system occurring in fluid mechanics where the advective equation appears with a possible degenerate anelastic constraint. Then we will look more carefully on the two-dimensional in space lake equations where an advective equation with transport velocity satisfying the anelastic constraint appears.

i) *The anelastic constraint from compressible isentropic Euler equations.* This anelastic constraint appears when the Mach (or Froude) number tends to zero starting the compressible isentropic Euler equations with some heterogeneity F (bathymetry, stratification for instance). More precisely consider the following system

$$\partial_t \rho_\varepsilon + \mathrm{div}(\rho_\varepsilon u_\varepsilon) = 0$$

with

$$\partial_t(\rho_\varepsilon u_\varepsilon) + \mathrm{div}(\rho_\varepsilon u_\varepsilon \otimes u_\varepsilon) + \frac{\nabla p(\rho_\varepsilon)}{\varepsilon^2} = \frac{\rho_\varepsilon \nabla F}{\varepsilon^2}$$

and the pressure law $p(\rho) = c\rho^\gamma$ (with two constants $c > 0$ and $\gamma > 1$) and where F is given and depends on the space variable (it represents heterogeneities in the environment). By letting formally ε to zero we get the following limit anelastic system

$$a\,(\partial_t u + u \cdot \nabla u) + a\nabla\pi = 0, \quad \mathrm{div}(au) = 0 \quad \text{where } a = \Big(\frac{\gamma-1}{c\,\gamma}\Big)^{1/(\gamma-1)} (F)^{1/(\gamma-1)}.$$

Therefore the anelastic constraint $\mathrm{div}(au) = 0$ actually accounts for the heterogeneity.

ii) *The lake equations.* This application concerns the so-called lake equation (under the rigid lid assumption) with possible vanishing topography. The PDEs is valid on a two-dimensional bounded domain Ω (the surface of the lake). This system reads

$$a(\partial_t u + u \cdot \nabla u + \nabla p) = 0 \qquad \text{with} \qquad \mathrm{div}(au) = 0 \text{ in } (0,T) \times \Omega$$

with respectively the boundary condition and the initial data

$$au \cdot n|_{(0,T)\times\partial\Omega} = 0, \qquad au|_{t=0} = m_0 \text{ in } \Omega$$

where a denotes the bathymetry and $u = (u_1, u_2)$ is a two-dimensional vector field which corresponds to the vertically averaging of the horizontal components of the velocity field $U = (U_1, U_2, W)$ in a three dimensional basin. Note that such system has been studied by Levermore, Oliver, and Titi [1996] in the non-degenerate case and by Bresch and Métivier [2006], Lacave, Nguyen, and Pausader [2014] and Munteanu [2012] in the degenerate case.

By reducing Ω to the support of a, we may assume that the bathymetry a is strictly positive in the domain Ω and possibly vanish on the shore $\partial\Omega$. Introducing the relative vorticity $\omega_R = \operatorname{curl} u/a$ where $\operatorname{curl} u = \partial_1 u_2 - \partial_2 u_1$, we check starting from the lake equation and dividing by a inside the domain that

$$\partial_t \omega_R + u \cdot \nabla \omega_R = 0 \text{ in } (0,T)\times\Omega, \qquad \omega_R|_{t=0} = \omega_0^R = \frac{\operatorname{curl} u_0}{a} \text{ in } \Omega$$

with

$$\operatorname{div}(au) = 0, \qquad \operatorname{curl} u = a\,\omega_R, \qquad au\cdot n|_{(0,T)\times\partial\Omega} = 0.$$

Remark. The boundary condition on au may be considered in a weak form if the boundary of the domain is not Lipschitz (Ω reduced by the support of a for example).

Definition. Let (u_0, ω_R^0) be such that

$$\operatorname{div}(au_0) = 0 \text{ in } \Omega, \qquad au_0 \cdot n|_{\partial\Omega} = 0$$

and

$$\omega_R^0 \in L^\infty(\Omega), \qquad \operatorname{curl} u_0 = a\,\omega_R^0.$$

A couple (v, ω) is a global renormalized solution of the vorticity formulation of the lake equation with initial condition (v^0, ω^0) if

- $\omega_R \in L^\infty((0,T)\times\Omega)$ and $\sqrt{a}u \in L^\infty(0,T; L^2(\Omega))$
- $\operatorname{div}(au) = 0$ in $(0,T)\times\Omega$ and $a\,u\cdot n|_{(0,T)\times\partial\Omega} = 0$
- $\operatorname{curl} u = a\,\omega_R$ in the distributional sense.
- For all $\chi \in W^{1,\infty}(\mathbb{R})$ with $\chi(0) = 0$, choosing $\psi \in C_c^\infty(\Omega)$, then

$$\frac{d}{dt}\int_\Omega \chi(\omega_R)\, a\psi(x)\, dx - \int_\Omega \chi(\omega_R)\, au\cdot\nabla\psi\, dx = 0.$$

Using the stability process regarding the advective equation with anelastic constraint, we can get the following result

Theorem 7. *Let a be continuous on $\overline{\Omega}$ and strictly positive in Ω. Assume that $\nabla\sqrt{a} \in L^{2+}(\Omega)$ and that there exists $\eta > 0$ such that $1/a^\eta \in L^1(\Omega)$. Then there exists a global renormalized solution of the vorticity formulation of the lake Equation.*

Constructing an approximate sequence of global renormalized solution in the sense of the definition given above for $(a_\varepsilon, \alpha_\varepsilon)$ constructed in the paper and in the whole space $\mathbb{R}^2$ is an standard procedure since the coefficients and the domain are regular and the approximate bathymetry is far from vacuum, see for instance Levermore, Oliver, and Titi [1996], Lacave, Nguyen, and Pausader [2014], Bresch and Métivier [2006]. We get the following bounds uniform with respect to the parameter ε

$$u_\varepsilon \in L^\infty(0,T; L^2_a(\mathbb{R}^2)), \qquad \omega_R^\varepsilon \in L^\infty((0,T)\times\mathbb{R}^2). \tag{29}$$

Remark now that

$$\mathrm{curl}(a_\varepsilon u_\varepsilon) = u_\varepsilon \cdot \nabla^\perp a_\varepsilon + a_\varepsilon\, \mathrm{curl} u_\varepsilon = \sqrt{a_\varepsilon} u_\varepsilon \cdot \nabla^\perp \sqrt{a}_\varepsilon + a_\varepsilon^2\, \omega_R^\varepsilon$$

and

$$\mathrm{div}(a_\varepsilon u_\varepsilon) = 0.$$

This is the system that we will use to get regularity on $a_\varepsilon \nabla u_\varepsilon$ required in the hypothesis for the stability. Using the uniform bounds on ω_R^ε and $\sqrt{a}_\varepsilon u_\varepsilon$ and the uniform bound $\nabla\sqrt{a}_\varepsilon \in L^{2^+}(\Omega)$, we get that

$$a_\varepsilon u_\varepsilon \in L^\infty(0,T; W^{1,p}(\Omega)) \text{ for some } p > 1$$

Thus writing

$$a_\varepsilon \nabla u_\varepsilon = \nabla(a_\varepsilon u_\varepsilon) - u_\varepsilon \cdot \nabla a_\varepsilon$$

we get that, uniformly in ε,

$$a_\varepsilon \nabla u_\varepsilon \in L^\infty(0,T; L^p(\Omega)) \text{ for some } p > 1. \tag{30}$$

On the other hand,

$$\int_\Omega a_\varepsilon |\nabla u_\varepsilon|\, |\log a_\varepsilon|\, dx \le \frac{1}{\eta} \int_\Omega a_\varepsilon |\nabla u_\varepsilon| (\log(e + a_\varepsilon |\nabla u_\varepsilon|) - \log \eta) + \int_\Omega \frac{1}{a_\varepsilon^\eta}.$$

for $\eta > 0$ chosen such that $1/a_\varepsilon^\eta \in L^1(\Omega)$. By combining this with (30), we obtain a uniform bound on

$$\int_0^T \int_\Omega a_\varepsilon\, |\nabla u_\varepsilon|\, \log(e + |\nabla u_\varepsilon|)\, dx\, dt,$$

leading to the uniform bound on u_ε for the quantity $\|u\|_{a_\varepsilon}$ recalling that we already control u_ε uniformly in $L^\infty(0,T; L^2_a(\Omega))$. This allows then to use the stability procedure taking $\alpha = a_\varepsilon^k$ for any $k \ge 1$ to get the conclusion of the Theorem.

Remark. It is interesting to note that we get global renormalized solution instead of global weak solution as in Lacave, Nguyen, and Pausader [2014]. In our result we use compactness on the vorticity through quantitative regularity estimate compared to compactness on the velocity field through the stream function equation and Aubin-Lions Lemma as usually.

Remark. Let us observe that assuming a behaves $\mathrm{dist}(x, \partial\Omega)^k$ the first hypothesis in the theorem asks for $k > 1$. The second hypothesis being satisfied. Of course we can generalize for more general power k playing with parameters θ using for instance that

$$a^\theta u \in L^\infty(0, T; W^{1,p}(\Omega))$$

and also

$$a^\theta \nabla u \in L^\infty(0, T; L^p(\Omega)) \text{ for some } p > 1$$

if $a^{\theta+1/2} \in L^{2+}(\Omega)$ and $\nabla a^{\theta-1/2} \in L^{2+}(\Omega)$.

References

Luigi Ambrosio (2004). "Transport equation and Cauchy problem for BV vector fields". *Invent. Math.* 158.2, pp. 227–260. MR: 2096794 (cit. on p. 2186).

Luigi Ambrosio and Gianluca Crippa (2014). "Continuity equations and ODE flows with non-smooth velocity". *Proc. Roy. Soc. Edinburgh Sect. A* 144.6, pp. 1191–1244. MR: 3283066 (cit. on p. 2186).

Julien Barré, David Chiron, Thierry Goudon, and Nader Masmoudi (2015). "From Vlasov-Poisson and Vlasov-Poisson-Fokker-Planck systems to incompressible Euler equations: the case with finite charge". *J. Éc. polytech. Math.* 2, pp. 247–296. MR: 3426749 (cit. on p. 2186).

Fethi Ben Belgacem and P-E Jabin (Nov. 2016). "Convergence of numerical approximations to non-linear continuity equations with rough force fields". arXiv: 1611.10271 (cit. on p. 2186).

Fethi Ben Belgacem and Pierre-Emmanuel Jabin (2013). "Compactness for nonlinear continuity equations". *J. Funct. Anal.* 264.1, pp. 139–168. MR: 2995703 (cit. on p. 2186).

Anna Bohun, François Bouchut, and Gianluca Crippa (2016). "Lagrangian flows for vector fields with anisotropic regularity". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 33.6, pp. 1409–1429. MR: 3569235 (cit. on p. 2186).

François Bouchut and Gianluca Crippa (2013). "Lagrangian flows for vector fields with gradient given by a singular integral". *J. Hyperbolic Differ. Equ.* 10.2, pp. 235–282. MR: 3078074 (cit. on p. 2186).

Jean Bourgain, Haim Brezis, and Petru Mironescu (2001). "Another look at Sobolev spaces". In: *Optimal control and partial differential equations*. IOS, Amsterdam, pp. 439–455. MR: 3586796 (cit. on p. 2186).

Didier Bresch and Pierre-Emmanuel Jabin (July 2015). "Global Existence of Weak Solutions for Compresssible Navier–Stokes Equations: Thermodynamically unstable pressure and anisotropic viscous stress tensor". arXiv: 1507.04629 (cit. on pp. 2185, 2186).

– (2017a). "Global weak solutions of PDEs for compressible media: a compactness criterion to cover new physical situations". In: *Shocks, singularities and oscillations in nonlinear optics and fluid mechanics*. Vol. 17. Springer INdAM Ser. Springer, Cham, pp. 33–54. MR: 3675552 (cit. on pp. 2185, 2186).

– (2017b). *Quantitative regularity estimates for compressible transport equations* (cit. on pp. 2186, 2188, 2189, 2196).

Didier Bresch and Guy Métivier (2006). "Global existence and uniqueness for the lake equations with vanishing topography: elliptic estimates for degenerate equations". *Nonlinearity* 19.3, pp. 591–610. MR: 2209290 (cit. on pp. 2186, 2207, 2209).

– (2010). "Anelastic limits for Euler-type systems". *Appl. Math. Res. Express. AMRX* 2, pp. 119–141. MR: 2719374.

Nicolas Champagnat and Pierre-Emmanuel Jabin (2010). "Well posedness in any dimension for Hamiltonian flows with non BV force terms". *Comm. Partial Differential Equations* 35.5, pp. 786–816. MR: 2753620 (cit. on pp. 2186, 2196).

Gianluca Crippa and Camillo De Lellis (2008). "Estimates and regularity results for the DiPerna-Lions flow". *J. Reine Angew. Math.* 616, pp. 15–46. MR: 2369485 (cit. on p. 2186).

Camillo De Lellis (2007). "Notes on hyperbolic systems of conservation laws and transport equations". In: *Handbook of differential equations: evolutionary equations. Vol. III*. Handb. Differ. Equ. Elsevier/North-Holland, Amsterdam, pp. 277–382. MR: 2549371 (cit. on p. 2186).

Ron J. DiPerna and Pierre-Louis Lions (1989). "Ordinary differential equations, transport theory and Sobolev spaces". *Invent. Math.* 98.3, pp. 511–547. MR: 1022305 (cit. on pp. 2186, 2206).

Donatella Donatelli and Eduard Feireisl (2017). "An anelastic approximation arising in astrophysics". *Math. Ann.* 369.3-4, pp. 1573–1597. arXiv: 1604.05860. MR: 3713551 (cit. on p. 2186).

Dale R. Durran (1989). "Improving the anelastic approximation". *Journal of the atmospheric sciences* 46.11, pp. 1453–1461 (cit. on p. 2186).

Eduard Feireisl (2004). *Dynamics of viscous compressible fluids*. Vol. 26. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, pp. xii+212. MR: 2040667 (cit. on p. 2185).

Eduard Feireisl, Josef Málek, Antonín Novotný, and Ivan Straškraba (2008). "Anelastic approximation as a singular limit of the compressible Navier-Stokes system". *Comm. Partial Differential Equations* 33.1-3, pp. 157–176. MR: 2398223 (cit. on p. 2186).

Eduard Feireisl and Antonín Novotný (2009). *Singular limits in thermodynamics of viscous fluids*. Advances in Mathematical Fluid Mechanics. Birkhäuser Verlag, Basel, pp. xxxvi+382. MR: 2499296 (cit. on p. 2185).

Maxime Hauray, Claude Le Bris, and Pierre-Louis Lions (2007). "Deux remarques sur les flots généralisés d'équations différentielles ordinaires". *C. R. Math. Acad. Sci. Paris* 344.12, pp. 759–764. MR: 2340443 (cit. on p. 2186).

Rupert Klein (2005). "Multiple spatial scales in engineering and atmospheric low Mach number flows". *M2AN Math. Model. Numer. Anal.* 39.3, pp. 537–559. MR: 2157149 (cit. on p. 2186).

Christophe Lacave, Toan T. Nguyen, and Benoit Pausader (2014). "Topography influence on the lake equations in bounded domains". *J. Math. Fluid Mech.* 16.2, pp. 375–406. MR: 3208722 (cit. on pp. 2186, 2188, 2207, 2209, 2210).

David Lannes (2017). "On the dynamics of floating structures". *Ann. PDE* 3.1, Art. 11, 81. MR: 3645931 (cit. on p. 2186).

C. David Levermore, Marcel Oliver, and Edriss S. Titi (1996). "Global well-posedness for the lake equations". *Physica D: Nonlinear Phenomena* 98.2-4, pp. 492–509 (cit. on pp. 2186, 2207, 2209).

Pierre-Louis Lions (1996). *Mathematical topics in fluid mechanics. Vol. 1*. Vol. 3. Oxford Lecture Series in Mathematics and its Applications. Incompressible models, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, pp. xiv+237. MR: 1422251 (cit. on p. 2185).

Franik B Lipps and Richard S Hemler (1982). "A scale analysis of deep moist convection and some related numerical calculations". *Journal of the Atmospheric Sciences* 39.10, pp. 2192–2210 (cit. on p. 2186).

Nader Masmoudi (2007). "Rigorous derivation of the anelastic approximation". *J. Math. Pures Appl. (9)* 88.3, pp. 230–240. MR: 2355456 (cit. on p. 2186).

Ionuţ Munteanu (2012). "Existence of solutions for models of shallow water in a basin with a degenerate varying bottom". *J. Evol. Equ.* 12.2, pp. 393–412. MR: 2923940 (cit. on p. 2207).

A. Novotný and I. Straškraba (2004). *Introduction to the mathematical theory of compressible flow*. Vol. 27. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, pp. xx+506. MR: 2084891.

Charlotte Perrin and Ewelina Zatorska (2015). "Free/congested two-phase model from weak solutions to multi-dimensional compressible Navier-Stokes equations". *Comm. Partial Differential Equations* 40.8, pp. 1558–1589. MR: 3355504 (cit. on p. 2186).

Elias M. Stein (1993). *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Vol. 43. Princeton Mathematical Series. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, pp. xiv+695. MR: 1232192 (cit. on p. 2198).

Robert Wilhelmson and Yoshimitsu Ogura (1972). "The pressure perturbation and the numerical modeling of a cloud". *Journal of the Atmospheric Sciences* 29.7, pp. 1295–1307 (cit. on p. 2186).

Received 2017-11-16.

Didier Bresch
LAMA CNRS UMR 5127
Univ. of Savoie Mont-Blanc
Bat. Le Chablais
73376 Le Bourget du Lac France
didier.bresch@univ-smb.fr

Pierre-Emmanuel Jabin
CSCAMM and Dept. of Mathematics
Univ. of Maryland College Park MD 20742 USA
pjabin@cscamm.umd.edu

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2211–2232)

INTERFACE DYNAMICS FOR INCOMPRESSIBLE FLUIDS: SPLASH AND SPLAT SINGULARITIES

DIEGO CÓRDOBA

Abstract

In this survey I report on recent progress in the study of the dynamics of the interface in between two incompressible fluids with different characteristics. In particular I focus on the formation of Splash and Splat singularities in two different settings: Euler equations and Darcy's law.

1 Introduction

We denote the interface in between two incompressible irrotational fluids in $\mathbb{R}^2$ by $\partial\Omega(t)$

$$\partial\Omega(t) = \{z(\alpha,t) = (z_1(\alpha,t), z_2(\alpha,t)) \,|\, (\alpha,t) \in (\mathbb{R},\mathbb{R}^+)\}.$$

The interface separates the plane into two regions Ω_j, with $j = 1,2$. Each Ω_j denotes the region occupied by the two different fluids with velocities $v^j = (v_1^j, v_2^j)$, different constant densities ρ_j, different constant viscosities μ_j and pressures p^j. We assume that the interface moves with the fluid i.e.

$$(\partial_t z - v^j)\cdot(\partial_\alpha z)^\perp = 0 \quad \text{on} \quad \partial\Omega$$

Taking into account surface tension σ, we assume that the pressure satisfies $p^1 = p^2(t) - \sigma K$ at $\partial\Omega(t)$, where K is the curvature of the interface $K(z) = \frac{z_{\alpha\alpha}\cdot z_\alpha^\perp}{|z_\alpha|^3}$.

The fluids are assumed to be incompressible and irrotational ($\nabla\cdot v^i = 0$, $\nabla\times v^i = 0$ in $\Omega^i(t)$). The vorticity will be supported on the free boundary curve $z(\alpha,t)$ and it has the form $\omega(x,t) = \varpi(\alpha,t)\delta(x - z(\alpha,t))$, i.e. the vorticity is a Dirac measure on z defined by

$$\langle\nabla^\perp\cdot v,\eta\rangle = \int_{\mathbb{R}} \varpi(\alpha,t)\eta(z(\alpha,t))d\alpha,$$

We thank Tania Pernas for providing the figures for the paper. DC was partially supported by the grants MTM2014-59488-P (Spain) and SEV-2015-556.

MSC2010: 76B15.

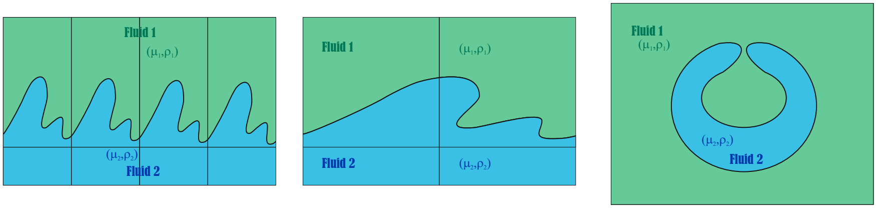


Figure 1: Interface on the real line, in a periodic domain and a closed contour.

with $\eta(x)$ being a test function.

We consider three possible scenarios:

- open curves vanishing at infinity: $\lim_{\alpha\to\infty}(z(\alpha,t)-(\alpha,0))=0$,
- periodic curves in the horizontal variable: $z(\alpha+2k\pi,t)=z(\alpha,t)+2k\pi(1,0)$.
- closed curves: $z(\alpha,t)$ is a 2π-periodic function in α.

Then $z(\alpha,t)$ evolves with a velocity field coming from the Biot-Savart law, which can be explicitly computed; it is given by the Birkhoff-Rott integral of the amplitude ϖ along the interface curve:

$$BR(z,\varpi)(\alpha,t)=\frac{1}{2\pi}PV\int\frac{(z(\alpha,t)-z(\beta,t))^{\perp}}{|z(\alpha,t)-z(\beta,t)|^2}\varpi(\beta,t)d\beta, \tag{1}$$

where PV denotes principal value.

We have

$$\begin{aligned} v^2(z(\alpha,t),t)&=BR(z,\varpi)(\alpha,t)+\frac{1}{2}\frac{\varpi(\alpha,t)}{|\partial_\alpha z(\alpha,t)|^2}\partial_\alpha z(\alpha,t),\\ v^1(z(\alpha,t),t)&=BR(z,\varpi)(\alpha,t)-\frac{1}{2}\frac{\varpi(\alpha,t)}{|\partial_\alpha z(\alpha,t)|^2}\partial_\alpha z(\alpha,t),\end{aligned} \tag{2}$$

where $v^j(z(\alpha,t),t)$ denotes the limit velocity field obtained approaching the boundary in the normal direction inside Ω^j and $BR(z,\varpi)(\alpha,t)$ is given by Equation (1). This provides us with the velocity field at the interface, from which we can subtract any term in the tangential direction without modifying the geometric evolution of the curve

$$z_t(\alpha,t)=BR(z,\varpi)(\alpha,t)+c(\alpha,t)\partial_\alpha z(\alpha,t). \tag{3}$$

A wise choice of $c(\alpha,t)$, namely:

$$c(\alpha,t) = \frac{\alpha+\pi}{2\pi}\int_{\mathbb{T}} \frac{\partial_\alpha z(\alpha,t)}{|\partial_\alpha z(\alpha,t)|^2}\cdot \partial_\alpha BR(z,\varpi)(\alpha,t)d\alpha - \\ - \int_{-\pi}^{\alpha} \frac{\partial_\alpha z(\beta,t)}{|\partial_\alpha z(\beta,t)|^2}\cdot \partial_\beta BR(z,\varpi)(\beta,t)d\beta,$$

allows us to establish the fact that the length of the tangent vector to $z(\alpha,t)$ only depends upon the variable t:

$$A(t) = |\partial_\alpha z(\alpha,t)|^2.$$

which will be of help in obtaining the needed a priori estimates.

Next, in order to close the system, we consider the following two settings:

1. *Euler equations:*

$$\rho_j \frac{\partial v^j}{\partial t} + \rho_j (v^j\cdot\nabla)v^j = -\nabla p^j - g\rho_j\, e_2 \quad \text{in} \quad \Omega_j, \tag{4a}$$

$$\nabla\cdot v^j = 0 \quad \text{and} \quad \nabla^\perp v^j = 0 \quad \text{in} \quad \Omega_j, \tag{4b}$$

$$(\partial_t z - v^j)\cdot(\partial_\alpha z)^\perp = 0 \quad \text{on} \quad \partial\Omega, \tag{4c}$$

$$p^1 - p^2 = -\sigma K \quad \text{on} \quad \partial\Omega. \tag{4d}$$

2. *Darcy's law-Muskat problem:*

$$\frac{\mu_j}{\kappa} v^j = -\nabla p^j - g\rho_j\, e_2 \quad \text{in} \quad \Omega_j, \tag{5a}$$

$$\nabla\cdot v^j = 0 \quad \text{in} \quad \Omega_j, \tag{5b}$$

$$(\partial_t z - v^j)\cdot(\partial_\alpha z)^\perp = 0 \quad \text{on} \quad \partial\Omega, \tag{5c}$$

$$p^1 - p^2 = -\sigma K \quad \text{on} \quad \partial\Omega. \tag{5d}$$

Here, $j \in \{1,2\}$, $\sigma > 0$ is the surface tension coefficient, e_2 is the second vector of a Cartesian basis, g is the acceleration due to gravity and κ the permeability of the medium.

The main goal of this survey is to review those local well-posed scenarios where initially the interface $\partial\Omega$ satisfies the chord-arc condition and later self-intersects at one point in finite time. We call such a singularity a "Splash". Since Euler equations are reversible in time, the aim is to prove a theorem of local existence starting from a Splash type singularity and choosing an initial data that opens the splash. The strategy of the proof for the free boundary (i.e. when the upper fluid is replaced by a vacuum) Euler equations enable us to extend the scenario to Splat-type singularities; that is, for those in which collapse may occur along a curve. This was first proved in Castro, D. Córdoba, Fefferman, Gancedo, and Gómez-Serrano [2013] for irrotational flows without surface tension and later extended in Castro, D. Córdoba, Fefferman, Gancedo, and Gómez-Serrano [2012] to the case of $\sigma > 0$ and for non-trivial vorticity Coutand and Shkoller [2014]. We studied similar scenarios in the presence of viscosity; free boundary Navier-Stokes equations (see Castro, D. Córdoba, Fefferman, Gancedo, and Gómez-Serrano [2015] and Coutand and Shkoller [2015]) and the one-phase Muskat problem (see Castro, D. Córdoba, Fefferman, and Gancedo [2016]). The strategy of the proof has to be different, since the equations cannot be solved backwards in time due to the presence of viscosity. We succeed in proving Splash-type singularities for the viscous free boundary problem but surprisingly, see D. Córdoba and Pernas-Castaño [2017], for the one-phase Muskat there are no Splat-type singularities.

The Splash and Splat type singularity for the free boundary has the particularity that the regularity of the interface is not lost, but the chord-arc ceases to be well defined at one point. It is very important for the proof that no fluid exists between the two curves that collapse. So the next step is to study the self-intersection of the interface in the presence of a fluid between the curves. In this sense, Fefferman, Ionescu, and Lie [2016] showed that the interface does not develop any self-intersections in finite time if the interface and the velocity on the boundary remains bounded in C^4 and C^3 respectively. These regularity spaces are not sharp. We discuss below a strategy to prove a splash-type singularity with two fluids with lower regularity.

2 Euler equations

The dynamics of the interface satisfies

$$z_t(\alpha,t) = BR(z,\varpi)(\alpha,t) + c(\alpha,t)\partial_\alpha z(\alpha,t).$$

Next, in order to close the system we apply Bernoulli's law, which leads to an equation relating the parametrization $z(\alpha,t)$ with the amplitude $\varpi(\alpha,t)$.

Let us consider an irrotational flow satisfying the Euler equations

$$\rho(v_t + v\nabla v) = -\nabla p - (0, \mathrm{g}\,\rho),$$

and the incompressibility condition $\nabla\cdot v = 0$, and let ϕ be such that $v(x,t) = \nabla\phi(x,t)$ for $x \neq z(\alpha,t)$. Then we have the expression

$$\rho(\phi_t(x,t) + \frac{1}{2}|v(x,t)|^2 + \mathrm{g}x_2) + p(x,t) = 0.$$

where

$$\rho(x_1,x_2,t) = \begin{cases} \rho_1, & x \in \Omega_1(t) \\ \rho_2, & x \in \Omega^2(t) = \mathbb{R}^2 - \Omega^1(t). \end{cases}$$

From the Biot-Savart law, for $x \neq z(\alpha,t)$, we get

$$\phi(x,t) = \frac{1}{2\pi} PV \int \arctan\Big(\frac{x_2 - z_2(\beta,t)}{x_1 - z_1(\beta,t)}\Big)\varpi(\beta,t)d\beta.$$

Let us define

$$\Pi(\alpha,t) = \phi^2(z(\alpha,t),t) - \phi^1(z(\alpha,t),t),$$

where $\phi^j(z(\alpha,t),t)$ denotes the limit obtained approaching the boundary in the normal direction inside Ω^j. It is clear that

$$\begin{aligned}\partial_\alpha \Pi(\alpha,t) &= (\nabla\phi^2(z(\alpha,t),t) - \nabla\phi^1(z(\alpha,t),t)) \cdot \partial_\alpha z(\alpha,t) \\ &= (v^2(z(\alpha,t),t) - v^1(z(\alpha,t),t)) \cdot \partial_\alpha z(\alpha,t) \\ &= \varpi(\alpha,t).\end{aligned}$$

We have

$$\begin{aligned}\phi^2(z(\alpha,t),t) &= IT(z,\varpi)(\alpha,t) + \frac{1}{2}\Pi(\alpha,t) \\ \phi^1(z(\alpha,t),t) &= IT(z,\varpi)(\alpha,t) - \frac{1}{2}\Pi(\alpha,t)\end{aligned}$$

where

$$IT(z,\varpi)(\alpha,t) = \frac{1}{2\pi} PV \int \arctan\Big(\frac{z_2(\alpha,t) - z_2(\beta,t)}{z_1(\alpha,t) - z_1(\beta,t)}\Big)\varpi(\beta,t)d\beta.$$

Then using Bernoulli's law inside each domain and taking limits approaching the common boundary, one finds

$$\rho^j(\phi^j_t(z(\alpha,t),t) + \frac{1}{2}|v^j(z(\alpha,t),t)|^2 + \mathrm{g}z_2(\alpha,t)) + p^j(z(\alpha,t),t) = 0,$$

and since

$$p^1(z(\alpha,t),t) = p^2(z(\alpha,t),t) - \sigma K,$$

yields

$$\begin{aligned}\Pi_t(\alpha,t) = &-2A_\rho\partial_t(IT(z,\varpi)(\alpha,t)) + c(\alpha,t)\varpi(\alpha,t) + A_\rho|BR(z,\varpi)(\alpha,t)|^2 \\ &+ 2A_\rho c(\alpha,t)BR(z,\varpi)(\alpha,t)\cdot\partial_\alpha z(\alpha,t) - A_\rho\frac{|\varpi(\alpha,t)|^2}{4|\partial_\alpha z(\alpha,t)|^2} - 2A_\rho \mathrm{g}z_2(\alpha,t) \\ &+ \sigma K.\end{aligned}$$

where $A_\rho = \frac{\rho_2 - \rho_1}{\rho_2 + \rho_1}$.

Then by taking a derivative, we obtain the desired formula for ϖ, which reads as follows

$$\begin{aligned}\varpi_t(\alpha,t) &= -2A_\rho \partial_t BR(z,\varpi)(\alpha,t)\cdot\partial_\alpha z(\alpha,t)\\ &\quad - A_\rho\partial_\alpha\Big(\frac{|\varpi|^2}{4|\partial_\alpha z|^2}\Big)(\alpha,t) + \partial_\alpha(c\,\varpi)(\alpha,t)\\ &\quad + 2A_\rho c(\alpha,t)\partial_\alpha BR(z,\varpi)(\alpha,t)\cdot\partial_\alpha z(\alpha,t)\\ &\quad + 2A_\rho \mathrm{g}\partial_\alpha z_2(\alpha,t) + \sigma\partial_\alpha K.\end{aligned} \tag{6}$$

In other words, we have obtained the Equations (3) and (6) for the evolution of the internal wave.

Concerning the Cauchy problem for the internal wave problem (two-fluid interface scenario); the system is locally well-posed as long as the surface tension is strictly positive (see Lannes [2013b] and reference therein).

In Fefferman, Ionescu, and Lie [2016] it is shown that Splash-type singularities cannot develop smoothly in the case of regular solutions of the two-fluid interface system. The argument of the proof is based on two steps. First they prove that for a sufficient regular velocity field there exists a critical L^∞ bound for the measure of the vorticity in the boundary, because ϖ satisfies a variant of Burgers equation. The second step is to obtain a double exponential lower bound on the minimum distance in between the curves that collapse under the assumptions of regular velocities and bounded $\varpi(\alpha,t)$. With a different approach, similar results have been obtained by Coutand and Shkoller [2016] for the vortex sheet.

2.1 Splash and Splat Singularities for the Free Boundary.

If we consider one fluid in a vacuum, a free boundary problem, then Splash and Splat type singularities can develop in finite time (see Castro, D. Córdoba, Fefferman, Gancedo, and Gómez-Serrano [2013]). Since the equations are time-reversible, the strategy of the proof is to establish a local existence theorem from the initial data that has a splash or a splat singularity.

- Theorem [Splash and Splat]
  1. Initial data exist in H^k (for k sufficiently large) such that the solution to the system (7) and (8) produces a Splash-type singularity.
  2. Initial data exist in H^k (for k sufficiently large) such that the solution to the system (7) and (8) produces a Splat-type singularity.

We can also write the dynamics of the interface in terms of the free boundary $z(\alpha,t)$ and the amplitude of the vorticity $\varpi(\alpha,t)$ as before, but with $\rho_1 = 0$ and $p_1 = 0$:

$$z_t(\alpha,t) = BR(z,\varpi)(\alpha,t) + c(\alpha,t)z_\alpha(\alpha,t), \tag{7}$$

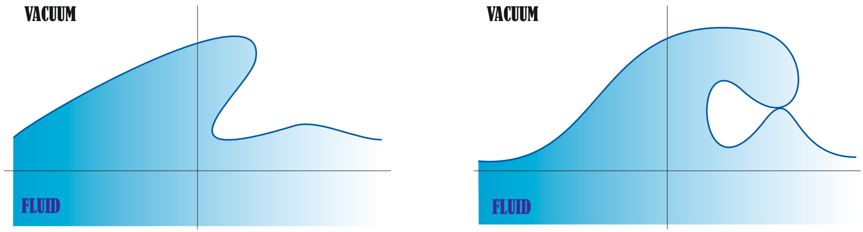


Figure 2: A Splash type singularity.

$$
\begin{aligned}
\varpi_t(\alpha,t) = & -2BR_t(z,\varpi)(\alpha,t)\cdot z_\alpha(\alpha,t) - \left(\frac{\varpi^2}{4|\partial_\alpha z|^2}\right)_\alpha(\alpha,t) + (c\varpi)_\alpha(\alpha,t) \\
& + 2c(\alpha,t)BR_\alpha(z,\varpi)(\alpha,t)\cdot z_\alpha(\alpha,t) - 2(z_2)_\alpha(\alpha,t) \\
& + \sigma\left(\frac{z_{\alpha\alpha}(\alpha,t)\cdot z_\alpha^\perp(\alpha,t)}{|z_\alpha(\alpha,t)|^3}\right)_\alpha
\end{aligned}
\tag{8}
$$

The first local existence results, for the free boundary incompressible Euler equations, are due to Nalimov [1974], Yosihara [1982] and Craig [1985] for near equilibrium initial data. Local existence for general initial data in Sobolev spaces was first achieved by Wu [1997] in 2d and in 3d Wu [1999], assuming initially the arc-chord condition. For other variations and results see Alazard, Burq, and Zuily [2011], Ambrose and Masmoudi [2005], Christodoulou and Lindblad [2000], A. Córdoba, D. Córdoba, and Gancedo [2010], Coutand and Shkoller [2007], Lannes [2005], Lannes [2013b], Lannes [2013a], Lindblad [2005], Shatah and Zeng [2008b], Shatah and Zeng [2008a], Shatah and Zeng [2011] and P. Zhang and Z. Zhang [2008].

We assume that between the curves that collapse there is no fluid, then we have no control over the growth of the amplitude of the vorticity. The amplitude of the vorticity is not even well-defined at the splash point. We use a conformal map with a complex square root, for example

$$P(w) = \left(\tan\left(\frac{w}{2}\right)\right)^{1/2}, \quad w\in\mathbb{C},$$

to keep apart the self-intersecting points, taking the branch of the square root above passing through the splash point such that its singular points (where P cannot be inverted) are located outside the fluid. The evolution equations for the curve in the new coordinate system has the following form

$$\tilde{z}(\alpha,t) = P(z(\alpha,t))$$

and the new amplitude $\tilde{\omega}$:

$$\tilde{z}_t(\alpha,t) = Q^2(\alpha,t)BR(\tilde{z},\tilde{\omega})(\alpha,t) + \tilde{c}(\alpha,t)\tilde{z}_\alpha(\alpha,t), \tag{9}$$

$$\begin{aligned}\tilde{\omega}_t(\alpha,t) =& -2BR_t(\tilde{z},\tilde{\omega})(\alpha,t)\cdot\tilde{z}_\alpha(\alpha,t) - |BR(\tilde{z},\tilde{\omega})|^2(Q^2)_\alpha(\alpha,t)\\ &-\Big(\frac{Q^2(\alpha,t)\tilde{\omega}(\alpha,t)^2}{4|\tilde{z}_\alpha(\alpha,t)|^2}\Big)_\alpha + 2\tilde{c}(\alpha,t)BR_\alpha(\tilde{z},\tilde{\omega})\cdot\tilde{z}_\alpha(\alpha,t)\\ &+(\tilde{c}(\alpha,t)\tilde{\omega}(\alpha,t))_\alpha - 2\left(P_2^{-1}(\tilde{z}(\alpha,t))\right)_\alpha\\ &+\sigma\left(\frac{Q^3}{|\tilde{z}_\alpha(\alpha,t)|^3}(\tilde{z}_\alpha^T HP_2^{-1}\tilde{z}_\alpha\nabla P_1^{-1}\cdot\tilde{z}_\alpha - \tilde{z}_\alpha^T HP_1^{-1}\tilde{z}_\alpha\nabla P_2^{-1}\cdot\tilde{z}_\alpha)\right)_\alpha\\ &+\sigma\left(Q\frac{\tilde{z}_{\alpha\alpha}(\alpha,t)\cdot\tilde{z}_\alpha^\perp(\alpha,t)}{|\tilde{z}_\alpha(\alpha,t)|^3}\right)_\alpha\end{aligned} \tag{10}$$

where

$$Q^2(\alpha,t) = \left|\frac{dP}{dw}(P^{-1}(\tilde{z}(\alpha,t)))\right|^2,$$

and HP_i^{-1} denotes the Hessian matrix of P_i^{-1}, which is the i-th ($i=\{1,2\}$) component of the transformation P^{-1}.

This map P transforms the splash into a closed curved whose chord-arc is well-defined. We select an initial velocity that immediately separates the point of collapse. In the new domain, the existence of solutions for Equations (9) and (10) is proven by using energy estimates. In order to return to the original domain and obtain solutions to (7) and (8), it is necessary to invert the map P.

The strategy of the proof enables us to extend the scenario to Splat-type singularities; that is, for those in which collapse may occur along a curve.

2.2 Splash singularities for the internal wave. In this section, we discuss the possible formation of singularities of an internal wave. If the vacuum is replaced by an incompressible fluid with low density, then low-density fluid resists the collapse of the curves. If the self-intersection still occurs, then we need to show a local existence in a lower regularity space as in Fefferman, Ionescu, and Lie [2016] (which allows the fluid inside to escape more easily) or to show that the interface losses regularity at the time of self-intersection. Either way our previous analysis for the splash does not work for the case of an internal wave and it demands new ideas.

Our program first consists in showing stationary splash singularities with two fluids. The main idea here is to perturb a one-parameter family of exact stationary to the free boundary Euler equations with no gravity. We use the implicit function theorem to construct stationary solutions with a sufficiently small ρ_1 that captures the splash point. Below

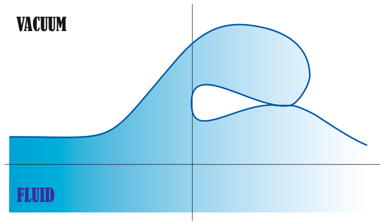


Figure 3: Splat-type singularity

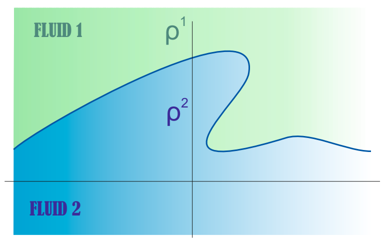


Figure 4: Two fluids: internal waves

we sketch a brief description of the construction. This scenario could lead to an understanding of the plausible low regularity dynamical splash singularity.

2.3 Stationary Splash singularity. In each domain, the fluid flow is governed by the stationary, incompressible, irrotational Euler equations; that is, the respective velocities v^j and the corresponding pressures p^j satisfy

$$\rho_j (v^j \cdot \nabla) v^j = -\nabla p^j - g\rho_j\, e_2 \quad \text{in} \quad \Omega_j, \tag{11a}$$

$$\nabla \cdot v^j = 0 \quad \text{and} \quad \nabla^\perp v^j = 0 \quad \text{in} \quad \Omega_j, \tag{11b}$$

$$v^j \cdot (\partial_\alpha z)^\perp = 0 \quad \text{on} \quad \partial\Omega, \tag{11c}$$

$$p^1 - p^2 = -\sigma K \quad \text{on} \quad \partial\Omega. \tag{11d}$$

We assume that the interface satisfies periodicity conditions

$$z_1(\alpha + 2\pi) = z_1(\alpha) + 2\pi, \quad z_2(\alpha + 2\pi) = z_2(\alpha)$$

and is symmetric with respect to the y-axis:

$$z_1(-\alpha) = -z_1(\alpha), \quad z_2(-\alpha) = z_2(\alpha).$$

In particular, we will often restrict attention to one period of Ω_k, where we assume the interface has a splash point z_*, i.e.

$$\exists \alpha_* \in (0, \pi) : \quad z_* = z(\alpha_*) = z(-\alpha_*), \quad z'(\alpha_*) = -z'(-\alpha_*),$$

dividing Ω_1 into two disjoint open sets Ω_1^i ($i = 1, 2$), having one boundary point in common; i.e.

$$\Omega_1 = \Omega_1^1 \cup \Omega_1^2, \qquad \overline{\Omega}_1^1 \cap \overline{\Omega}_1^2 = \{z_*\},$$

where Ω_-^1 is assumed to be bounded, while Ω_-^2 is unbounded. In particular, both Ω_1^j possess an outward cusp with a common tip at z_*.

We use the hodograph transform with respect to the lower fluid to fix the parametrization, for details see D. Córdoba, Enciso, and Grubic [2016a], in order to transform the free boundary problem into a problem on a fixed domain; in this case, the lower half-plane $\mathbb{C}_-$. More precisely, we use the analytic function

$$w = \phi + i\psi : \Omega_2 \longrightarrow \mathbb{C}_-$$

(which can be shown to be a conformal bijection extending to a homeomorphism up to the boundary, cf.D. Córdoba, Enciso, and Grubic [ibid.] Lemma 9.)) as an independent variable instead of $z = x + iy$. Here, ϕ is the potential of the flow, while ψ is the stream function defined implicitly via

$$v = \nabla\phi = \nabla^\perp\psi.$$

The parametrization is fixed by requiring

$$\phi(z(\alpha)) = \alpha, \quad \psi(z(\alpha)) = 0,$$

which implies a simple relation between the velocity of the lower fluid and the tangent vector on the interface

$$v^2 \cdot z_\alpha = \nabla\phi \cdot z_\alpha = 1 \quad \Rightarrow \quad 2BR(z, \omega) \cdot \partial_\alpha z + \omega = 2.$$

A stationary solution of the two-fluid system is reduced to finding 2π-periodic functions $\omega(\alpha)$ and $z(\alpha)-(\alpha,0)$ satisfying

$$2|\partial_\alpha z|^2 M(z)+\epsilon\,\omega(\omega-2)=2, \tag{12a}$$

$$2BR(z,\omega)\cdot\partial_\alpha z+\omega=2, \tag{12b}$$

$$BR(z,\omega)\cdot\partial_\alpha^\perp z=0, \tag{12c}$$

where M is given by

$$M(z)=-\frac{2\rho_2}{\rho_2-\rho_1}qK(z)-2gz_2+1+2\kappa,$$

$q:=\frac{\sigma}{\rho_2}$, $\epsilon:=\frac{2\rho_1}{\rho_2-\rho_1}$, $K(z)$ is the curvature of the interface and κ is a perturbation of the constant arising in the pure capillary wave problem.

The problem can be simplified further by writing the velocity vector in terms of polar coordinates

$$\partial_\alpha z=|\partial_\alpha z|e^{i\theta}=e^{if},\qquad f=\theta+i\tau:\mathbb{C}_-\to\mathbb{C},$$

where f is analytic in the lower half-plane and continuous up to the boundary. Since θ and τ are 2π-periodic conjugate functions, on the interface they must be related by the periodic Hilbert transform; i.e. $\tau(\alpha)=H\theta(\alpha)$ and we can therefore take $\theta(\alpha)$ as our main unknown and consider z as a function of θ via the integral operator

$$z(\alpha)=I(\theta)(\alpha):=\int_{-\pi}^{\alpha}e^{-H\theta(\alpha')+i\theta(\alpha')}d\alpha'. \tag{13}$$

As shown in D. Córdoba, Enciso, and Grubic [ibid.], the system (11)-(12) is then equivalent to the following system of equations

$$q\Big(1+\frac{\epsilon}{2}\Big)\frac{d\theta}{d\alpha}+\sinh H\theta+ge^{-H\theta}\,\operatorname{Im}I(\theta)-\kappa e^{-H\theta}-\frac{\epsilon}{4}e^{H\theta}\varpi(\varpi-2)=0, \tag{14a}$$

$$2BR(z,\varpi)\cdot\partial_\alpha z+\varpi=2, \tag{14b}$$

where $\theta(\alpha)$ and $\varpi(\alpha)$ are 2π-periodic functions, θ is odd, ϖ is even and $z:=I(\theta)$ is defined by (13) .

The above system depends on four constants q,κ,ε and g, where g represents the gravity, κ is the integration constant of the Bernoulli equation, q is related to the surface tension coefficient via $q:=\frac{\sigma}{\rho_2}$, while

$$\varepsilon:=\frac{2\rho_1}{\rho_2-\rho_1}$$

detects the presence of an upper fluid. On setting ϵ to zero, the equations decouple and we recover the capillary-gravity wave problem as studied in Akers, Ambrose, and Wright [2014]. If, in addition, we set $g = 0$, we recover the pure capillary waves problem as formulated by Levi-Civita (see e.g.Levi-Civita [1925] and Okamoto and Shōji [2001]), namely;

Pure capillary waves problem 1. *Find a 2π-periodic, analytic function $f = \theta + i\tau$ on the lower half-plane that satisfies*

$$q\frac{d\theta}{d\alpha} = -\sinh H\theta$$

on the boundary and tends to zero at infinity.

This problem admits a family of exact solutions depending on the parameter q. In fact, Crapper [1957] has shown that the family of analytic functions

$$f_A(w) := 2i \log \frac{1 + Ae^{-iw}}{1 - Ae^{-iw}}$$

has all the required properties. Parameter A depends on q via

$$q = \frac{1 + A^2}{1 - A^2}.$$

and it actually suffices to consider $A \geq 0$, since the transformation $A \mapsto -A$ corresponds to a translation $\alpha \to \alpha + \pi$. The corresponding wave profiles are given by

$$z_A(\alpha) = \alpha + \frac{4i}{1 + Ae^{-i\alpha}} - 4i.$$

where the constant has been chosen to have $z_A(\alpha) = (\alpha, 0)$ for $A = 0$. For sufficiently large values of parameter A, these solutions can no longer be represented as a graph of a function and eventually self-intersect. It is not hard to see that the curve z_A does not have self-intersections if and only if

$$A < A_0 \approx 0.45467.$$

For $A = A_0$, the curve $z_A(\alpha)$ exhibits a splash, while for A slightly larger than A_0 the curve intersects at exactly two points, and the intersection is transverse.

Since, for a sufficiently regular, non self-intersecting curve (e.g. if $A < A_0$), the operator

$$\mathfrak{A}(z)(\omega) := 2BR(z, \omega) \cdot z_\alpha$$

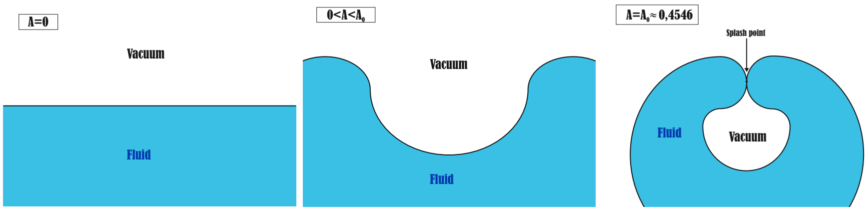


Figure 5: Interface at different values of the parameter A.

is a compact bounded linear operator from $H^1 \to H^1$ whose eigenvalues are strictly smaller than 1 (in absolute value), the operator $1 + \mathfrak{A}(z)$ is invertible (see Baker, Meiron, and Orszag [1982]). In previous work, A. Córdoba, D. Córdoba, and Gancedo [2010] and A. Córdoba, D. Córdoba, and Gancedo [2011], in order to obtain a priori estimates, in the case of a moving interface it was necessary to estimate the norms of $1 + \mathfrak{A}(z)$ in terms of the chord-arc and smoothness of z.

In particular, we can solve (14) for the corresponding vorticity ω_A by inverting $1 + \mathfrak{A}(z_A)$ and we can use the implicit function theorem on the system of equations (14) to perturb around $(\epsilon, g, \kappa) = 0$. The idea of perturbing Crapper waves to construct a more general stationary interface was introduced in Akers, Ambrose, and Wright [2014] (see also Ambrose, Strauss, and Wright [2016] and de Boeck [2014]). In D. Córdoba, Enciso, and Grubic [2016a], we used the implicit function theorem to perturb the wave profiles of the Crapper family for all values of the parameter $A < A_0$ and thereby construct solutions of (14) with the upper fluid present (i.e. $\rho_1 > 0$) arbitrarily close to the splash. We define an *η-splash curve*, for arbitrarily small $\eta > 0$, if

$$\inf_{\alpha<\beta} \frac{|z(\alpha) - z(\beta)|}{\min\{\alpha - \beta, 1\}} \leq \eta$$

.

- Theorem [Stationary almost-Splash singularities for two fluids] Let us fix the density of the second fluid $\rho_2 > 0$ and consider any $\eta > 0$. For any sufficiently small $\rho_1 \geq 0$ and g, there is some positive surface tension coefficient σ such that there exists a periodic solution to the two-fluid problem (11) for which the interface $\partial\Omega$ is an η-splash curve.

However, if the interface exhibits a splash singularity (e.g. if $A = A_0$), in which case the arc-chord condition fails and the domain Ω_1 is no longer connected, with both connecting components exhibiting an outward cusp touching at the splash point, the operator $\mathfrak{A}(z) : H^1 \to H^1$ is no longer compact and we cannot invert the operator as usual. But again, in the case of one fluid; i.e. $\epsilon = \rho_1 = 0$, we can take advantage of a suitable conformal map that opens up the splash point for $\partial\Omega = \{z(\alpha)\}$. The solution z is invertible and its inverse $\omega = \phi + i\psi$ is such that $v = \nabla\phi$ satisfies (11) with $p^1 = \rho_1 = 0$. Furthermore, by perturbing and using again the implicit function theorem, we can prove:

- Theorem [Stationary Splash singularities for one fluid]

  Let us fix the density of the second fluid $\rho_2 > 0$ and assume that $\rho_1 = 0$. Then for any small enough g, there is some positive surface tension coefficient σ such that there exists a periodic solution to the system (11) for which the interface $\partial\Omega$ has a splash singularity.

The geometry near the singular point suggests the use of an appropriate homogeneous weighted Sobolev spaces (cf. works of Maz'ya and Soloviev [2010]) with weights given in powers of the distance to the splash point which we set at the origin $z = 0$. On parametrizing the interface for simplicity around the origin as a graph, these spaces are defined as follows: Let

$$w_\beta(x) := |x|^\beta$$

be the weight function for $\beta \in \mathbb{R}$ and x in some interval $I \in \mathbb{R}$ containing the origin. Then for $k \in \mathbb{N}$

$$u \in W^k_{p,\beta} :\Longleftrightarrow w_{\beta+j}(x)\partial_x^j u \in L^p,\ j \leq k.$$

We take w_β to be a Muckenhaupt weight, in which case the Hilbert transform is bounded on the weighted Lebesgue space $\mathfrak{L}_{p,\beta} = W^0_{p,\beta}$. This is equivalent to requiring $0 < \beta + p^{-1} < 1$. The first step is to show that

- $\mathfrak{A}(z) : W^k_{p,\beta} \to W^k_{p,\beta}$ is continuous.

However, we cannot expect $1 + \mathfrak{A}(z)$ to be invertible as it stans, since $W^k_{p,\beta}$ does not take into account the order μ of the cusp or equivalently the way in which the arc-chord condition blows-up as we approach the singularity. We address this question by showing that $1 + \mathfrak{A}(z)$ actually has values in a smaller Banach space; i.e. we aim to show

- $1 + \mathfrak{A}(z) : W^1_{p,\beta} \to X_{\beta,\mu}$ continuous on a closed subspace $X_{\beta,\mu} \subset W^1_{p,\beta}$,
- $1 + \mathfrak{A}(z) : W^1_{p,\beta} \to X_{\beta,\mu}$ invertible by using conformal maps.

Finally, after adjusting the Banach space for θ, we show that we can use the implicit function theorem on (14) defined on these new spaces (see D. Córdoba, Enciso, and Grubic [2016b]).

- Theorem [Stationary Splash singularities for two fluids] Let us fix the density of the second fluid $\rho_2 > 0$. Then for any sufficiently small upper fluid density $\rho_1 \geq 0$ and g, there is some positive surface tension coefficient σ for which there exists a stationary solution two-fluid Euler equations such that the interface $\partial\Omega$ has a Splash singularity. The regularity of $\partial\Omega$ and ϖ is C^2 and C^α.

In a forthcoming paper, we are interested in applying the ideas outlined above to study the dynamics in which a non-intersecting curve self-intersects in finite-time for the full Euler equations.

Therefore, it is very appealing to try to formulate the problem within the framework of weighted Sobolev space. The main tool consists of the apriori estimates for a carefully chosen energy functional involving an H^k-norm for the curvature and an $H^{k+1/2}$-norm for the vorticity with $k \in \mathbb{N}$ sufficiently high coupled with some lower-order correction terms.

We wish to generalize this approach to the case of weighted Sobolev spaces. Our first goal is to prove a local existence theorem in the setting of weighted Sobolev spaces

Two related questions arise, which Sobolev spaces to use and how to generalize the energy functional. In the dynamic case, it seems more natural to use the non-homogeneous weighted Sobolev spaces $V^k_{p,\gamma}$ which, with respect to the power weight w_γ, are defined as follows

$$u \in V^k_{p,\gamma} :\Longleftrightarrow w_\gamma(x)\partial^j_x u \in L^p,\ j \le k.$$

Under certain conditions, these spaces are equivalent to homogeneous weighted Sobolev spaces up to a polynomial. The spaces have to be chosen in such a way that $\mathfrak{A}$ remains invertible, but they also depend on how we generalize the energy functional. A first attempt would be a naive generalization of the energy functional to spaces of the type $V^k_{p,\gamma}$ for non self-intersecting interface. For instance, this approach implies

$$\partial^k K \in \mathfrak{L}_{p,\gamma} \ \Rightarrow\ \partial^k \omega \in \mathfrak{L}_{p,\gamma-1/2} \text{ and } \Lambda^{1/2}\partial^k\omega \in \mathfrak{L}_{p,\gamma},$$

while some of the difficulties include the non-positivity of certain terms otherwise positive in the non-weighted setting. In the second phase, we would need to correct the energy functional to include the effects of the blow-up of the arc-chord when the interface self-intersects.

3 Incompressible porous media equation-Darcy's law

In this section, we consider the basic setting of two immiscible fluids in a porous media having different densities and viscosities separated by a sharp interface which is modeled by Darcy's law. To model the contact of a porous media with air, a dry region, one assumes that the media on top has zero viscosity and density. This is the so called Muskat problem (see Muskat [1934]), which has attracted a lot of attention over the last ten years. Saffman and Taylor made the observation that the one phase version (one of the fluids has zero viscosity) was also known as the Hele-Shaw cell equation (see Saffman and Taylor [1958]), which, in turn, is the zero-specific heat case of the classical one-phase Stefan problem.

Darcy's law is the following momentum equation for the velocity v

$$\frac{\mu}{\kappa}v = -\nabla p - (0, \mathrm{g}\,\rho), \tag{15}$$

where p is the pressure, μ is the dynamic viscosity, κ is the permeability of the medium, ρ is the liquid density and g is the acceleration due to gravity. Together with the incompressibility condition $\nabla \cdot v = 0$, equation (15) implies that the flow is irrotational in each domain Ω_j. So again the vorticity is a Dirac measure on z and we have an expression of the velocity, as in (2), in terms of z and the measure ϖ.

Then taking the limit in Darcy's, law we obtain

$$\begin{aligned}
&\Big(\frac{\mu^2}{\kappa}v^2(z(\alpha,t),t) - \frac{\mu^1}{\kappa}v^1(z(\alpha,t),t)\Big)\cdot\partial_\alpha z(\alpha,t) = \\
&\qquad = -(\nabla p^2(z(\alpha,t),t) - \nabla p^1(z(\alpha,t),t))\cdot\partial_\alpha z(\alpha,t) - \mathrm{g}(\rho^2-\rho^1)\,\partial_\alpha z_2(\alpha,t)\\
&\qquad = -\partial_\alpha(p^2(z(\alpha,t),t) - p^1(z(\alpha,t),t)) - \mathrm{g}(\rho^2-\rho^1)\,\partial_\alpha z_2(\alpha,t)\\
&\qquad = -\mathrm{g}(\rho^2-\rho^1)\,\partial_\alpha z_2(\alpha,t) + \sigma K_\alpha,
\end{aligned}$$

since $p^2(z(\alpha,t),t) - p^1(z(\alpha,t),t) = \sigma K$. Moreover

$$\begin{aligned}
&\frac{\mu^2+\mu^1}{2\kappa}\varpi(\alpha,t) + \frac{\mu^2-\mu^1}{\kappa}BR(z,\varpi)(\alpha,t)\cdot\partial_\alpha z(\alpha,t) = \\
&\qquad\qquad\qquad\qquad = -\mathrm{g}(\rho^2-\rho^1)\partial_\alpha z_2(\alpha,t) + \sigma K_\alpha,
\end{aligned}$$

so that

$$\begin{aligned}
\varpi(\alpha,t) &= -A_\mu 2BR(z,\varpi)(\alpha,t)\cdot\partial_\alpha z(\alpha,t)\\
&\qquad - 2\kappa\mathrm{g}\frac{\rho^2-\rho^1}{\mu^2+\mu^1}\partial_\alpha z_2(\alpha,t) + \frac{2\kappa\sigma}{\mu^2+\mu^1}K_\alpha.
\end{aligned} \tag{16}$$

where $A_\mu = \frac{\mu_2-\mu_1}{\mu_2+\mu_1}$.

For this setting, we consider here the following two scenarios; asymptotically flat at infinity or periodic curves in the horizontal variable. The local well-posedness in Sobolev spaces is guaranteed with $\sigma > 0$ (see Ambrose [2014] and reference therein). With no surface tension, the result turns out to be false for some initial data. Rayleigh and Saffman-Taylor gave a condition that must be satisfied for the linearized model in order to have a solution locally in time namely, that the normal component of the pressure gradient jump at the interface has to have a distinguished sign (see Saffman and Taylor [1958], Ambrose [2004] and A. Córdoba, D. Córdoba, and Gancedo [2011]). This is known as the Rayleigh-Taylor condition $RT(\alpha,t)$ and can be written as follows:

$$RT(\alpha,t) = \frac{\mu^2-\mu^1}{\kappa}BR(z,\varpi)(\alpha,t)\cdot\partial_\alpha z(\alpha,t) + \mathrm{g}(\rho^2-\rho^1)\partial_\alpha z_1(\alpha,t) > 0.$$

Using Hopf's lemma, the Rayleigh-Taylor condition is satisfied for the one phase Muskat; i.e. $\mu_1 = \rho_1 = 0$ (see Castro, D. Córdoba, Fefferman, and Gancedo [2016]). For the case of equal viscosities, $\mu_1 = \mu_2$, this condition holds when the more dense fluid lies below the interface.

For the Muskat problem, splash singularity cannot be developed in the case in which $\mu_1 = \mu_2$ and $\rho_1 \neq \rho_2$, for more details see Gancedo and Strain [2014]. However, the splash can be achieved with $\mu_1 = \rho_1 = 0$, i.e. the one-phase Muskat problem, where Ω_1 corresponds to the dry region (see Castro, D. Córdoba, Fefferman, and Gancedo [2016]). On the other hand, the presence of viscosity, $\mu_2 > 0$ may prevent the existence of Splat singularities (see D. Córdoba and Pernas-Castaño [2017]):

- Theorem [Splash and Non Splat for the one phase Muskat problem]

    1. Initial data exist in H^k such that the solution to the system (3) and (16) produces a Splash-type singularity with $\mu_1 = \rho_1 = 0$.
    2. There are NO Splat-type singularities for the system (3) and (16) .

Sketch of the proof of the splash: Since the velocity is divergence free and irrotational in Ω_2 there exists harmonic functions ψ and ϕ such that $u = \nabla\phi = \nabla^{\perp}\psi$. The strategy is now to prove a local existence theorem from an almost splash configuration. We apply the map P from the previous section to a new system of coordinates $(\tilde{x}, \tilde{y})$ for which the domain Ω_2 has been transformed to a non-self-intersecting domain $\tilde{\Omega}$.

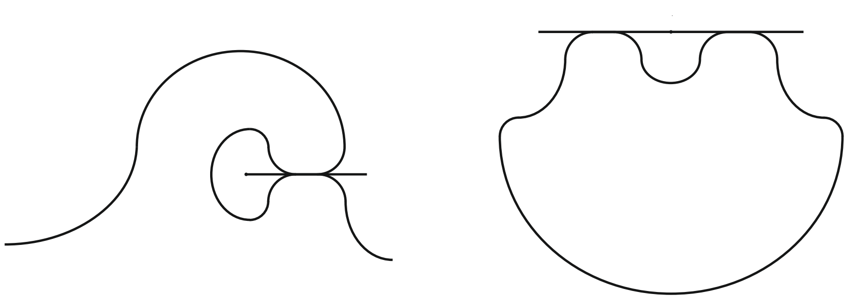


Figure 6: Domains Ω and $\tilde{\Omega} = P(\Omega)$

We define the new potential and a stream function in $\tilde{\Omega}$ by

$$\tilde{\psi}(\tilde{x}, \tilde{y}, t) \equiv \psi(P^{-1}(\tilde{x}, \tilde{y}), t), \qquad \tilde{\phi}(\tilde{x}, \tilde{y}, t) \equiv \phi(P^{-1}(\tilde{x}, \tilde{y}), t)$$

and the new velocity is defined by

$$\tilde{v}(\tilde{x}, \tilde{y}, t) \equiv \nabla\tilde{\phi}(\tilde{x}, \tilde{y}, t).$$

The boundary of the domain $\tilde{\Omega}$ takes the form $\tilde{z}(\alpha, t) = P(z(\alpha, t))$, and then we can extend our velocity to the whole space by taking $\tilde{\omega}(x, t) = \tilde{\omega}(\alpha, t)\delta(\tilde{x} - \tilde{z}(\alpha, t))$. Take $\tilde{z}(\alpha, t)$ and $\tilde{\omega}(\alpha, t)$ to be the unknowns. The Muskat equation in $\tilde{\Omega}$ without surface tension takes the form

$$\tilde{z}_t(\alpha, t) = Q^2(\alpha, t)BR(\tilde{z}, \tilde{\omega})(\alpha, t) + \tilde{c}(\alpha, t)\tilde{z}_\alpha(\alpha, t)$$

$$\tilde{\omega}(\alpha, t) = -2BR(\tilde{z}, \tilde{\omega})(\alpha, t) \cdot \tilde{z}_\alpha(\alpha, t) - 2\mathrm{g}\frac{\rho_2}{\mu_2}\partial_\alpha \left(P_2^{-1}(\tilde{z}(\alpha, t))\right).$$

where

$$Q^2(\alpha, t) = \left|\frac{dP}{dw}(z(\alpha, t))\right|^2.$$

We denote the energy E_k of the above system by

$$E_k(\tilde{z}, t) = \|\tilde{z}\|^2_{H^k}(t) + \|\mathcal{F}(\tilde{z})\|^2_{L^\infty}(t) + \frac{1}{m(Q^2\tilde{RT})(t)} + \sum_{l=0}^{4} \frac{1}{m(q^l)(t)},$$

where the chord-arc condition $\mathcal{F}(z)(\alpha, \beta, t)$ is defined as

$$\mathcal{F}(z)(\alpha, \beta, t) = \frac{|\beta|}{|z(\alpha, t) - z(\alpha - \beta, t)|} \qquad \forall\, \alpha, \beta \in (-\pi, \pi),$$

with

$$\mathcal{F}(z)(\alpha, 0, t) = \frac{1}{|\partial_\alpha z(\alpha, t)|}.$$

And

$$m(Q^2\tilde{RT})(t) = \min_{\alpha\in\mathbb{T}} Q^2(\alpha, t)\tilde{RT}(\alpha, t), \quad m(q^l)(t) = \min_{\alpha\in\mathbb{T}} |z(\alpha, t) - q^l|.$$

where q^l are the singular points of the P^{-1} conformal map and $\tilde{RT}$ is the Rayleigh-Taylor condition in the new domain. Then, the following estimate holds:

$$\frac{d}{dt}E_k(t) \le C(E_k(t))^p$$

for $k \geq 3$. The constants C and p depend only on k. These a priori estimates will lead to a local existence result for the contour equation in the tilde domain as long as the initial curve is smooth and satisfies the chord-arc condition in $\tilde{\Omega}$. So therefore we can prove a local existence for an almost splash configuration in Ω_2.

In order to finish the proof, we show the following proposition:

Proposition 3.1. *Let $x(\alpha, t)$ and $y(\alpha, t)$ be two curves which satisfy the contour equation in $\tilde{\Omega}(t)$. Then, the following estimate holds:*

$$\frac{d}{dt}\|x - y\|_{H^1}(t) \leq C(\sup_{[0,T]} E_3(x,t) + \sup_{[0,T]} E_3(y,t))^p \|x - y\|_{H^1}(t).$$

Above, $E_3(x,t)$ and $E_3(y,t)$ are given by local existence. The constants C and p are universal.

Now instead of showing local existence backwards in time from a splash point, in the transformed domain we prove local existence forward in time and show that the solutions depend stably on the initial conditions. The velocity at the splash point forces the interface to cross the branch in $\tilde{\Omega}$ due to the Rayleigh-Taylor condition. By a perturbative argument we can choose a smooth initial data close to a splash which in finite time self-intersects for the one-phase Muskat problem.

Sketch of the proof of the no-splat: The presence of viscosity implies that the equation has non-local parabolic behavior which will lead to an instant analyticity result for the curve. The basic mechanism to prove the absence of splat singularities is to notice that, at the critical time of splat formation, analyticity of the interface should still be retained. Thus, the contradiction comes from the failure of the unique continuation of an analytic function. This mechanism draws a clear line between the splash and splat scenarios since in the former case a single point touch would not have contradicted analyticity.

Below we show the main estimates that provide instant analyticity into the strip

$$S(t) = \{\alpha + i\zeta \in \mathbb{C} : \alpha \in \mathbb{T}, |\zeta| < \lambda t\},$$

where we take the derivative in time of the following norm

$$\|\tilde{z}\|^2_{H^k(S)}(t) = \|\tilde{z}\|^2_{L^2(S)}(t) + \sum_{\pm} \int_{\mathbb{T}} |\partial^k_\alpha \tilde{z}(\alpha \pm i\lambda t, t)|^2 d\alpha,$$

and obtain

$$\frac{d}{dt}\|\tilde{z}\|^2_{H^k(S)}(t) \leq exp C(\|\tilde{z}\|^2_{H^k(S)}(t) + ||\mathfrak{F}(\tilde{z})||_{L^\infty})$$

for a short time t.

Now we need to show a control of the decay of the strip of the analyticity in the new coordinate system as long as the curve remains smooth and the chord-arc condition is satisfied. Let the curve $\tilde{z}$ be initially an analytic curve in the strip

$$S(t=0) = \{\alpha + i\zeta \in \mathbb{C} : \alpha \in \mathbb{T}, |\zeta| < h(0)\}.$$

We obtain a priori estimates on the evolution of the norm $\|\tilde{z}\|^2_{H^k(S)}(t)$ on the strip

$$S(t) = \{\alpha + i\zeta \in \mathbb{C} : \alpha \in \mathbb{T}, |\zeta| < h(t)\}$$

in the following way

$$\begin{aligned}\frac{d}{dt}\sum_{\pm}\int_{\mathbb{T}} |\partial_\alpha^k \tilde{z}_j(\alpha \pm ih(t))|^2 d\alpha \leq \exp CE(t) \\ - 10h'(t)\int_{\mathbb{T}} \Lambda(\partial_\alpha^k \tilde{z}_j)(\alpha)\overline{\partial_\alpha^k \tilde{z}_j(\alpha)}d\alpha \\ + (\exp CE(t)h(t) + \frac{h'(t)}{10} + \exp CE(t))\|\Lambda^{\frac{1}{2}}\partial_\alpha^k \tilde{z}\|^2_{L^2(S)}.\end{aligned}$$

where $E(t) = (\|\tilde{z}\|^2_{H^k(S)}(t) + ||\mathfrak{F}(\tilde{z})||_{L^\infty})$ and the operator Λ is defined by $\widehat{\Lambda(f)}\xi = |\xi|\widehat{f}$. Choosing,

$$h(t) = \exp(-10\int_0^t G(r)dr)[\int_0^t -10G(r)\exp(10\int_0^r G(s)ds)dr + h(0)]$$

where $G(t) = \exp C(\|\tilde{z}\|^2_{H^k(S)}(t) + ||\mathfrak{F}(\tilde{z})||_{L^\infty})(t)$, we get the desired estimation.

Hence, we can conclude that our transformed curve $\tilde{z}$ is real-analytic into the strip $S(t)$. From the estimates this complex strip decays exponentially until a time that depends on the regularity of the curve and the arc-chord condition. Thus, by applying P^{-1}, we have that the analytic curve self-intersects along an arc; therefore, we get a contradiction.

References

Benjamin F. Akers, David M. Ambrose, and J. Douglas Wright (2014). "Gravity perturbed Crapper waves". *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 470.2161, pp. 20130526, 14. MR: 3145074 (cit. on pp. 2222, 2223).

T. Alazard, N. Burq, and C. Zuily (2011). "On the water-wave equations with surface tension". *Duke Math. J.* 158.3, pp. 413–499. MR: 2805065 (cit. on p. 2217).

David M. Ambrose (2004). "Well-posedness of two-phase Hele-Shaw flow without surface tension". *European J. Appl. Math.* 15.5, pp. 597–607. MR: 2128613 (cit. on p. 2226).

– (2014). "The zero surface tension limit of two-dimensional interfacial Darcy flow". *J. Math. Fluid Mech.* 16.1, pp. 105–143. MR: 3171344 (cit. on p. 2226).

David M. Ambrose and Nader Masmoudi (2005). "The zero surface tension limit of two-dimensional water waves". *Comm. Pure Appl. Math.* 58.10, pp. 1287–1315. MR: 2162781 (cit. on p. 2217).

David M. Ambrose, Walter A. Strauss, and J. Douglas Wright (2016). "Global bifurcation theory for periodic traveling interfacial gravity-capillary waves". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 33.4, pp. 1081–1101. arXiv: 1411.5569. MR: 3519533 (cit. on p. 2223).

Gregory R. Baker, Daniel I. Meiron, and Steven A. Orszag (1982). "Generalized vortex methods for free-surface flow problems". *J. Fluid Mech.* 123, pp. 477–501. MR: 687014 (cit. on p. 2223).

J. Thomas Beale, Thomas Y. Hou, and John S. Lowengrub (1993). "Growth rates for the linearized motion of fluid interfaces away from equilibrium". *Comm. Pure Appl. Math.* 46.9, pp. 1269–1301. MR: 1231428.

Peter de Boeck (Apr. 2014). "Existence of capillary-gravity waves that are perturbations of Crapper's waves". arXiv: 1404.6189 (cit. on p. 2223).

Angel Castro, Diego Córdoba, Charles Fefferman, and Francisco Gancedo (2016). "Splash singularities for the one-phase Muskat problem in stable regimes". *Arch. Ration. Mech. Anal.* 222.1, pp. 213–243. MR: 3519969 (cit. on pp. 2214, 2227).

Angel Castro, Diego Córdoba, Charles Fefferman, Francisco Gancedo, and Javier Gómez-Serrano (2012). "Finite time singularities for water waves with surface tension". *J. Math. Phys.* 53.11, pp. 115622, 26. MR: 3026567 (cit. on p. 2214).

– (2013). "Finite time singularities for the free boundary incompressible Euler equations". *Ann. of Math. (2)* 178.3, pp. 1061–1134. MR: 3092476 (cit. on pp. 2213, 2216).

Angel Castro, Diego Córdoba, Charles Fefferman, Francisco Gancedo, and Javier Gómez-Serrano (Apr. 2015). "Splash singularities for the free boundary Navier-Stokes equations". arXiv: 1504.02775 (cit. on p. 2214).

Demetrios Christodoulou and Hans Lindblad (2000). "On the motion of the free surface of a liquid". *Comm. Pure Appl. Math.* 53.12, pp. 1536–1602. MR: 1780703 (cit. on p. 2217).

Antonio Córdoba, Diego Córdoba, and Francisco Gancedo (2010). "Interface evolution: water waves in 2-D". *Adv. Math.* 223.1, pp. 120–173. MR: 2563213 (cit. on pp. 2217, 2223).

– (2011). "Interface evolution: the Hele-Shaw and Muskat problems". *Ann. of Math. (2)* 173.1, pp. 477–542. MR: 2753607 (cit. on pp. 2223, 2226).

Diego Córdoba, Alberto Enciso, and Nastasia Grubic (2016a). "On the existence of stationary splash singularities for the Euler equations". *Adv. Math.* 288, pp. 922–941. MR: 3436402 (cit. on pp. 2220, 2221, 2223).

Diego Córdoba, Alberto Enciso, and Nastasia Grubic (2016b). "On the existence of stationary splash singularities for the Euler equations". *Adv. Math.* 288, pp. 922–941. MR: 3436402 (cit. on p. 2224).

Diego Córdoba and Tania Pernas-Castaño (2017). "Non-splat singularity for the one-phase Muskat problem". *Trans. Amer. Math. Soc.* 369.1, pp. 711–754. MR: 3557791 (cit. on pp. 2214, 2227).

Daniel Coutand and Steve Shkoller (2007). "Well-posedness of the free-surface incompressible Euler equations with or without surface tension". *J. Amer. Math. Soc.* 20.3, pp. 829–930. MR: 2291920 (cit. on p. 2217).

– (2014). "On the finite-time splash and splat singularities for the 3-D free-surface Euler equations". *Comm. Math. Phys.* 325.1, pp. 143–183. MR: 3147437 (cit. on p. 2214).

– (May 2015). "On the splash singularity for the free-surface of a Navier-Stokes fluid". arXiv: 1505.01929 (cit. on p. 2214).

– (2016). "On the impossibility of finite-time splash singularities for vortex sheets". *Arch. Ration. Mech. Anal.* 221.2, pp. 987–1033. arXiv: 1407.1479. MR: 3488542 (cit. on p. 2216).

Walter Craig (1985). "An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits". *Comm. Partial Differential Equations* 10.8, pp. 787–1003. MR: 795808 (cit. on p. 2217).

G. D. Crapper (1957). "An exact solution for progressive capillary waves of arbitrary amplitude". *J. Fluid Mech.* 2, pp. 532–540. MR: 0091075 (cit. on p. 2222).

Charles Fefferman, Alexandru D. Ionescu, and Victor Lie (2016). "On the absence of splash singularities in the case of two-fluid interfaces". *Duke Math. J.* 165.3, pp. 417–462. MR: 3466160 (cit. on pp. 2214, 2216, 2218).

Francisco Gancedo and Robert M. Strain (2014). "Absence of splash singularities for surface quasi-geostrophic sharp fronts and the Muskat problem". *Proc. Natl. Acad. Sci. USA* 111.2, pp. 635–639. MR: 3181769 (cit. on p. 2227).

David Lannes (2005). "Well-posedness of the water-waves equations". *J. Amer. Math. Soc.* 18.3, pp. 605–654. MR: 2138139 (cit. on p. 2217).

– (2013a). "A stability criterion for two-fluid interfaces and applications". *Arch. Ration. Mech. Anal.* 208.2, pp. 481–567. MR: 3035985 (cit. on p. 2217).

– (2013b). *The water waves problem*. Vol. 188. Mathematical Surveys and Monographs. Mathematical analysis and asymptotics. American Mathematical Society, Providence, RI, pp. xx+321. MR: 3060183 (cit. on pp. 2216, 2217).

T. Levi-Civita (1925). "Détermination rigoureuse des ondes permanentes d'ampleur finie". *Math. Ann.* 93.1, pp. 264–314. MR: 1512238 (cit. on p. 2222).

Hans Lindblad (2005). "Well-posedness for the motion of an incompressible liquid with free surface boundary". *Ann. of Math. (2)* 162.1, pp. 109–194. MR: 2178961 (cit. on p. 2217).

Vladimir Maz'ya and Alexander Soloviev (2010). *Boundary integral equations on contours with peaks*. Vol. 196. Springer Science & Business Media (cit. on p. 2224).

Morris Muskat (1934). "Two fluid systems in porous media. The encroachment of water into an oil sand". *Physics* 5.9, pp. 250–264 (cit. on p. 2225).

V. I. Nalimov (1974). "The Cauchy-Poisson problem". *Dinamika Splošn. Sredy* Vyp. 18 Dinamika Židkost. so Svobod. Granicami, pp. 104–210, 254. MR: 0609882 (cit. on p. 2217).

Hisashi Okamoto and Mayumi Shōji (2001). *The mathematical theory of permanent progressive water-waves*. Vol. 20. Advanced Series in Nonlinear Dynamics. World Scientific Publishing Co., Inc., River Edge, NJ, pp. xii+229. MR: 1869386 (cit. on p. 2222).

P. G. Saffman and Geoffrey Taylor (1958). "The penetration of a fluid into a porous medium or Hele-Shaw cell containing a more viscous liquid". *Proc. Roy. Soc. London. Ser. A* 245, 312–329. (2 plates). MR: 0097227 (cit. on pp. 2225, 2226).

Jalal Shatah and Chongchun Zeng (2008a). "A priori estimates for fluid interface problems". *Comm. Pure Appl. Math.* 61.6, pp. 848–876. MR: 2400608 (cit. on p. 2217).

– (2008b). "Geometry and a priori estimates for free boundary problems of the Euler equation". *Comm. Pure Appl. Math.* 61.5, pp. 698–744. MR: 2388661 (cit. on p. 2217).

– (2011). "Local well-posedness for fluid interface problems". *Arch. Ration. Mech. Anal.* 199.2, pp. 653–705. MR: 2763036 (cit. on p. 2217).

Sijue Wu (1997). "Well-posedness in Sobolev spaces of the full water wave problem in 2-D". *Invent. Math.* 130.1, pp. 39–72. MR: 1471885 (cit. on p. 2217).

– (1999). "Well-posedness in Sobolev spaces of the full water wave problem in 3-D". *J. Amer. Math. Soc.* 12.2, pp. 445–495. MR: 1641609 (cit. on p. 2217).

Hideaki Yosihara (1982). "Gravity waves on the free surface of an incompressible perfect fluid of finite depth". *Publ. Res. Inst. Math. Sci.* 18.1, pp. 49–96. MR: 660822 (cit. on p. 2217).

Ping Zhang and Zhifei Zhang (2008). "On the free boundary problem of three-dimensional incompressible Euler equations". *Comm. Pure Appl. Math.* 61.7, pp. 877–940. MR: 2410409 (cit. on p. 2217).

Received 2017-11-30.

Diego Córdoba
Instituto de Ciencias Matemáticas
Consejo Superior de Investigaciones Científicas
C/ Nicolas Cabrera, 13-15, 28049 Madrid, Spain
dcg@icmat.es

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2233–2258)

ON THE STRUCTURE OF MEASURES CONSTRAINED BY LINEAR PDES

Guido De Philippis and Filip Rindler

Abstract

The aim of this note is to present some recent results on the structure of the singular part of measures satisfying a PDE constraint and to describe some applications.

1 Introduction

We describe recent advances obtained by the authors and collaborators concerning the structure of singularities in measures satisfying a linear PDE constraint. Besides its own theoretical interest, understanding the structure of singularities of PDE-constrained measures turns out to have several (sometimes surprising) applications in the calculus of variations, geometric measure theory, and metric geometry.

Let $\mathcal{A}$ be a k'th-order linear constant-coefficient differential operator acting on $\mathbb{R}^N$-valued functions, i.e.

$$\mathcal{A}u := \sum_{|\alpha|\le k} A_\alpha \partial^\alpha u, \qquad u \in \mathrm{C}^\infty(\Omega;\mathbb{R}^N),$$

where $A_\alpha \in \mathbb{R}^n \otimes \mathbb{R}^N$ are linear maps from $\mathbb{R}^N$ to $\mathbb{R}^n$ and $\partial^\alpha = \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}$ for every multindex $\alpha = (\alpha_1, \dots, \alpha_d) \in (\mathbb{N} \cup \{0\})^d$.

The starting point of the investigation is the following:

Question 1.1. *Let $\mu \in \mathfrak{M}(\Omega, \mathbb{R}^N)$ be an $\mathbb{R}^N$-valued Radon measure on an open set $\Omega \subset \mathbb{R}^d$ and let μ be $\mathcal{A}$-free, i.e. μ solves the system of linear PDEs*

(1-1) $$\mathcal{A}\mu := \sum_{|\alpha|\le k} A_\alpha \partial^\alpha \mu = 0 \qquad \textit{in the sense of distributions.}$$

G. D. P. is supported by the MIUR SIR-grant "Geometric Variational Problems" (RBSI14RVEZ). F. R. acknowledges the support from an EPSRC Research Fellowship on "Singularities in Nonlinear PDEs" (EP/L018934/1).

MSC2010: primary 35D30; secondary 28B05, 42B37, 49Q20, 53C23.

What can be said about the singular part[1] *of* μ*?*

In answering the above question a prominent role is played by the *wave cone* associated with the differential operator $\mathcal{A}$:

$$\Lambda_{\mathcal{A}} := \bigcup_{|\xi|=1} \ker \mathbb{A}^k(\xi) \subset \mathbb{R}^N \qquad \text{with} \qquad \mathbb{A}^k(\xi) = (2\pi \mathrm{i})^k \sum_{|\alpha|=k} A_\alpha \xi^\alpha,$$

where we have set $\xi^\alpha := \xi_1^{\alpha_1} \cdots \xi_d^{\alpha_d}$.

Roughly speaking, $\Lambda_{\mathcal{A}}$ contains all the amplitudes along which the system (1-1) is *not elliptic*. Indeed if we assume that $\mathcal{A}$ is homogeneous, $\mathcal{A} = \sum_{|\alpha|=k} A_\alpha \partial^\alpha$, then it is immediate to see that $\lambda \in \mathbb{R}^N$ belongs to $\Lambda_{\mathcal{A}}$ if and only if there exists a non-zero $\xi \in \mathbb{R}^d \setminus \{0\}$ such that $\lambda h(x \cdot \xi)$ is $\mathcal{A}$-free for all smooth functions $h\colon \mathbb{R} \to \mathbb{R}$. In other words, "one-dimensional" oscillations and concentrations are possible only if the amplitude (direction) belongs to the wave cone. For this reason the wave cone plays a crucial role in the compensated compactness theory for sequences of $\mathcal{A}$-free maps, see Murat [1978, 1979], Tartar [1979], Tartar [1983], and DiPerna [1985], and in convex integration theory, see for instance Chiodaroli, De Lellis, and Kreml [2015], Chiodaroli, Feireisl, Kreml, and Wiedemann [2017], De Lellis and Székelyhidi [2009, 2012], De Lellis and Székelyhidi [2013], Székelyhidi and Wiedemann [2012], and Isett [2016] and the references cited therein. However, all these references are concerned with oscillations only, not with concentrations.

Since the singular part of a measure can be thought of as containing "condensed" concentrations, it is quite natural to conjecture that $|\mu|^s$-almost everywhere the polar vector $\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}$ belongs to $\Lambda_{\mathcal{A}}$. This is indeed the case and the main result of De Philippis and Rindler [2016]:

Theorem 1.2. *Let $\Omega \subset \mathbb{R}^d$ be an open set, let $\mathcal{A}$ be a k'th-order linear constant-coefficient differential operator as above, and let $\mu \in \mathfrak{M}(\Omega; \mathbb{R}^N)$ be an $\mathcal{A}$-free Radon measure on Ω with values in $\mathbb{R}^N$. Then,*

$$\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) \in \Lambda_{\mathcal{A}} \qquad \textit{for } |\mu|^s\textit{-a.e. } x \in \Omega.$$

[1] If not specified, the terms "singular" and "absolutely continuous" always refer to the Lebesgue measure. We also recall that, thanks to the Radon-Nikodym theorem, a vector-valued measure μ can be written as

$$\mu = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}\mathrm{d}|\mu| = g\,\mathcal{L}^d + \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}\mathrm{d}|\mu|^s$$

where $|\mu|$ is the total variation measure, $g \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ and $\mathcal{L}^d$ is the Lebesgue measure.

Remark 1.3. Theorem 1.2 is also valid in the situation

$$\mathcal{A}\mu = \sigma \qquad \text{for some } \sigma \in \mathcal{M}(\Omega; \mathbb{R}^n). \tag{1-2}$$

This can be reduced to the setting of Theorem 1.2 by defining

$$\tilde{\mu} := (\mu, \sigma) \in \mathcal{M}(\mathbb{R}^d; \mathbb{R}^{N+n})$$

and $\tilde{\mathcal{A}}$ (with an additional 0'th-order term) such that (1-2) is equivalent to $\tilde{\mathcal{A}}\tilde{\mu} = 0$. It is easy to check that $\Lambda_{\tilde{\mathcal{A}}} = \Lambda_{\mathcal{A}} \times \mathbb{R}^n$ and that for $|\mu|$-almost every point $\frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}$ is proportional to $\frac{\mathrm{d}\mu}{\mathrm{d}|\tilde{\mu}|}$.

One interesting feature of Theorem 1.2 is that it gives information about the directional structure of μ at singular points (the "shape of singularities"). Indeed, it is not hard to check that for all "elementary" $\mathcal{A}$-free measures of the form

$$\mu = \lambda\nu, \qquad \text{where} \qquad \lambda \in \Lambda_{\mathcal{A}},\ \nu \in \mathcal{M}^+(\mathbb{R}^d), \tag{1-3}$$

the scalar measure ν is necessarily translation invariant along directions that are orthogonal to the *characteristic set*

$$\Xi(\lambda) := \bigl\{\xi \in \mathbb{R}^d \,:\, \lambda \in \ker \mathbb{A}(\xi)\bigr\}.$$

Note that $\Xi(\lambda)$ turns out to be a subspace of $\mathbb{R}^d$ whenever $\mathcal{A}$ is a first-order operator. In this case, the translation invariance of ν in the directions orthogonal to $\Xi(\lambda)$ is actually the best information one can get from (1-3).

In the case of operators of order $k > 1$, due to the lack of linearity of the map $\xi \mapsto \mathbb{A}^k(\xi)$ for $k > 1$, the structure of elementary $\mathcal{A}$-free measures is more complicated and not yet fully understood.

In the next sections we will describe some applications of Theorem 1.2 to the following problems:

- The description of the singular part of derivatives of BV- and BD-maps.
- Lower semicontinuity for integral functionals defined on measures.
- Characterization of generalized gradient Young measures.
- The study of the sharpness of the Rademacher's theorem .
- Cheeger's conjecture on Lipschitz differentiability spaces.

In Section 7 we will sketch the proof of Theorem 1.2.

2 Structure of singular derivatives

Let $f : \mathbb{R}^\ell \otimes \mathbb{R}^d \to [0, \infty)$ be a linear growth integrand with $f(A) \sim |A|$ for $|A|$ large. Consider the following variational problem:

$$\int_\Omega f(\nabla u) \, \mathrm{d}x \to \min, \qquad u \in \mathrm{C}^1(\Omega; \mathbb{R}^\ell) \text{ with given boundary conditions.}$$

It is well known that in order to apply the Direct Method of the calculus of variations one has to relax the above problem to a setting where it is possible to obtain both compactness of minimizing sequences and lower semicontinuity of the functional with respect to some topology, usually the weak(*) topology in some function space. Due to the linear growth of the integrand the only easily available estimate on a minimizing sequence (u_k) is an a-priori bound on the L^1-norm of their derivativesx:

$$\sup_k \int_\Omega |\nabla u_k| \, \mathrm{d}x < \infty.$$

It is then quite natural to relax the functional to the space $\mathrm{BV}(\Omega, \mathbb{R}^\ell)$ of functions of *bounded variation*, i.e. those functions $u \in \mathrm{L}^1(\Omega; \mathbb{R}^\ell)$ whose distributional gradient is a matrix-valued Radon measure. A fine understanding of the possible behavior of measures arising as derivatives is then fundamental to study the weak* lower semicontinuity of the functional as well as its relaxation to the space BV.

In this respect, in Ambrosio and Giorgi [1988] Ambrosio and De Giorgi proposed the following conjecture:

Question 2.1. *Is the singular part of the derivative of a function $u \in \mathrm{BV}(\Omega; \mathbb{R}^\ell)$, which is usually denoted by $D^s u$, always of rank one? Namely, is it true that*

$$\frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}(x) = a(x) \otimes b(x)$$

for $|D^s u|$-a.e. x and some $a(x) \in \mathbb{R}^\ell$, $b(x) \in \mathbb{R}^d$?

Their conjecture was motivated by the fact that this structure is trivially true for the so-called *jump part* of $D^s u$ (which is always of the form $[u] \otimes n \, \mathcal{H}^{d-1} \llcorner J$, where J is the $\mathcal{H}^{d-1}$-rectifiable jump set, n is a normal on J, and $[u]$ is the jump height in direction n); see Ambrosio, Fusco, and Pallara [2000, Chapter 3] for a complete reference concerning functions of bounded variations.

A positive answer to the above question was given by Alberti in Alberti [1993] with his celebrated *rank-one theorem*. It was recognized quickly that this result has a central place in the calculus of variations and importance well beyond, in particular because it implies that locally all singularities in BV-functions are necessarily *one-directional*. Indeed, after a blow-up (i.e. magnification) procedure at $|D^s u|$-almost every point, the blow-up limit measure depends only on a single direction and is translation-invariant with respect to all orthogonal directions. This is not surprising for jumps, but it is a strong assertion about all other singularities in the *Cantor part* of $D^s u$, i.e., the remainder of $D^s u$ after subtracting the jump part.

While Alberti's original proof is geometric in nature, one can also interpret the theorem as a result about singularities in PDEs: BV-derivatives Du satisfy the PDE

$$\operatorname{curl} Du = 0 \qquad \text{in the sense of distributions,}$$

which can be written with a linear constant-coefficient PDE operator $\mathcal{A} := \sum_{j=1}^{d} A_j \partial_j$ as $\mathcal{A}\mu = 0$ in the sense of distributions.

Besides its intrinsic theoretical interest, the rank-one theorem also has many applications in the theory of BV-functions, for instance for lower semicontinuity and relaxation Ambrosio and Dal Maso [1992], Fonseca and Müller [1993], and Kristensen and Rindler [2010], integral representation theorems Bouchitté, Fonseca, and Mascarenhas [1998], Young measure theory Kristensen and Rindler [2012] and Rindler [2014], and the study of continuity equations with BV-vector fields Ambrosio [2004]. We refer to Ambrosio, Fusco, and Pallara [2000, Chapter 5] for further history.

At the the end of this section we will see that Alberti's rank-one theorem is a straightforward consequence of Theorem 1.2. Let us also mention that recently a very short proof of the Alberti rank-one theorem has been given by Massaccesi and Vittone in Massaccesi and Vittone [2016].

In problems arising in the theory of geometrically-linear elasto-plasticity Suquet [1978], Suquet [1979], and Temam and Strang [1980/81] one often needs to consider a larger space of functions than the space of functions of bounded variations. Indeed, in this setting energies usually only depend on the *symmetric part* of the gradient and one has to consider the following type of variational problem:

$$\int_\Omega f(\mathcal{E}u)\,\mathrm{d}x \to \min, \qquad u \in \mathrm{C}^1(\Omega;\mathbb{R}^d) \text{ with given boundary conditions,}$$

where $\mathcal{E}u = (\nabla u + \nabla u^T)/2 \in (\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}}$ is the symmetric gradient ($(\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}}$ being canonically isomorphic to the space of symmetric $(d \times d)$-matrices) and f is a linear-growth integrand with $f(A) \sim |A|$ for $|A|$ large. In this case, for a minimizing sequence (u_k) one can only obtain that

$$\sup_k \int_\Omega |\mathcal{E}u_k|\,\mathrm{d}x < \infty$$

and, due to the failure of Korn's inequality in L^1 Ornstein [1962], Conti, Faraco, and Maggi [2005], and Kirchheim and Kristensen [2016], this is not enough to ensure that $\sup_k \int |\nabla u_k| \, dx < \infty$. One then introduces the space $BD(\Omega)$ of functions of *bounded deformation*, i.e. those functions $u \in L^1(\Omega; \mathbb{R}^d)$ such that the symmetrized distributional derivative exists as a Radon measure, i.e.,

$$Eu := \frac{1}{2}(Du + Du^T) \in \mathfrak{M}(\Omega; (\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}}),$$

see Ambrosio, Coscia, and Dal Maso [1997], Temam [1983], and Temam and Strang [1980/81]. Clearly, $BV(\Omega, \mathbb{R}^d) \subset BD(\Omega)$ and the inclusion is strict Conti, Faraco, and Maggi [2005] and Ornstein [1962]. Note that for $u \in BV(\Omega; \mathbb{R}^d)$ as a consequence of Alberti's rank-one theorem one has

$$\frac{dE^s u}{d|E^s u|}(x) = a(x) \odot b(x),$$

where $a \odot b = (a \otimes b + b \otimes a)/2$ is the symmetrized tensor product. One is then naturally led to the following conjecture:

Question 2.2. *Is it true that for every function $u \in BD(\Omega)$ it holds that*

$$\frac{dE^s u}{d|E^s u|}(x) = a(x) \odot b(x)$$

for $|E^s u|$-a.e. x and some $a(x), b(x) \in \mathbb{R}^d$?

Again, besides its theoretical interest, it has been well known that a positive answer of the above question would have several applications to the study of lower semicontinuity and relaxation of functionals defined on BD, see the next section, as well as in establishing the absence of a singular part for minimizers, see for instance Francfort, Giacomini, and Marigo [2015, Remark 4.8].

Let us conclude this section by showing how both a positive answer to Question 2.2 and a new proof of Alberti's rank-one theorem can easily be obtained by applying Theorem 1.2 to suitable differential operators:

Theorem 2.3. *Let $\Omega \subset \mathbb{R}^d$ be open. Then:*

(i) If $u \in BV(\Omega; \mathbb{R}^\ell)$, then

$$\frac{dD^s u}{d|D^s u|}(x) = a(x) \otimes b(x) \qquad \text{for } |D^s u|\text{-a.e. } x \text{ and some } a(x) \in \mathbb{R}^\ell,\ b(x) \in \mathbb{R}^d.$$

(ii) If $u \in \mathrm{BD}(\Omega)$, then

$$\frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x) = a(x) \odot b(x) \qquad \text{for } |E^s u|\text{-a.e. } x \text{ and some } a(x), b(x) \in \mathbb{R}^d.$$

Proof. Observe that $\mu = Du$ is curl-free,

$$0 = \operatorname{curl} \mu = \left(\partial_i \mu_j^k - \partial_j \mu_i^k\right)_{i,j=1,\dots,d;\, k=1,\dots,\ell}.$$

Then, assertion (i) above follows from

$$\Lambda_{\mathrm{curl}} = \left\{ a \otimes \xi \,:\, a \in \mathbb{R}^\ell,\ \xi \in \mathbb{R}^d \setminus \{0\} \right\},$$

which can be proved by an easy computation.

In the same way, if $\mu = Eu$, then μ satisfies the *Saint-Venant compatibility conditions*,

$$0 = \operatorname{curl}\operatorname{curl} \mu := \left(\sum_{i=1}^d \partial_{ik} \mu_i^j + \partial_{ij} \mu_i^k - \partial_{jk} \mu_i^i - \partial_{ii} \mu_j^k \right)_{j,k=1,\dots,d}.$$

It is now a direct computation to check that

$$\Lambda_{\mathrm{curl}\,\mathrm{curl}} = \left\{ a \odot \xi \,:\, a \in \mathbb{R}^d,\ \xi \in \mathbb{R}^d \setminus \{0\} \right\}.$$

This shows assertion (ii) above. □

3 Functionals on measures

The theory of integral functionals with linear-growth integrands defined on vector-valued measures satisfying PDE constraints is central to many questions of the calculus of variations. In particular, their relaxation and lower semicontinuity properties have attracted a lot of attention, see for instance Ambrosio and Dal Maso [1992], Fonseca and Müller [1993, 1999], Fonseca, Leoni, and Müller [2004], Kristensen and Rindler [2010], Rindler [2011], and Baía, Chermisi, Matias, and Santos [2013]. Based on Theorem 1.2 one can unify and extend many of these results.

Concretely, let $\Omega \subset \mathbb{R}^d$ be an open and bounded set and consider the functional

$$\mathfrak{F}[\mu] := \int_\Omega f\left(x, \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^d}(x)\right) \mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}\mu^s}{\mathrm{d}|\mu|^s}(x)\right) \mathrm{d}|\mu|^s(x), \tag{3-1}$$

defined for finite vector Radon measures $\mu \in \mathfrak{M}(\Omega; \mathbb{R}^N)$ with values in $\mathbb{R}^N$ and satisfying

$$\mathcal{A}\mu = 0 \qquad \text{in the sense of distributions.}$$

Here, $f\colon \Omega \times \mathbb{R}^N \to [0,\infty)$ is a Borel integrand that has *linear growth at infinity*, i.e.,

$$|f(x,A)| \le M(1+|A|) \qquad \text{for all } (x,A) \in \Omega \times \mathbb{R}^N.$$

We also assume that the *strong recession function* of f exists, which is defined as

$$f^\infty(x,A) := \lim_{\substack{x' \to x \\ A' \to A \\ t \to \infty}} \frac{f(x',tA')}{t}, \qquad (x,A) \in \overline{\Omega} \times \mathbb{R}^N. \tag{3-2}$$

The (weak*) lower semicontinuity properties of $\mathfrak{F}$ depend on (generalized) *convexity* properties of the integrand in its second variable. For this, we need the following definition: A Borel function $h\colon \mathbb{R}^N \to \mathbb{R}$ is called $\mathcal{A}^k$*-quasiconvex* ($\mathcal{A}^k = \sum_{|\alpha|=k} A_\alpha \partial^\alpha$ being the principal part of $\mathcal{A}$) if

$$h(F) \le \int_Q h(F + w(y))\,\mathrm{d}y$$

for all $F \in \mathbb{R}^N$ and all Q-periodic $w \in \mathrm{C}^\infty(Q;\mathbb{R}^N)$ such that $\mathcal{A}^k w = 0$ and $\int_Q w\,\mathrm{d}y = 0$, where $Q := (0,1)^d$ is the open unit cube in $\mathbb{R}^d$; see Fonseca and Müller [1999] for more on this class of integrands. For $\mathcal{A} = \operatorname{curl}$ this notion is equivalent to the classical *quasiconvexity* as introduced by Morrey [1952].

It has been known for a long time that $\mathcal{A}^k$-quasiconvexity of $f(x,\bullet)$ is a necessary condition for the sequential weak* lower semicontinuity of $\mathfrak{F}$ on $\mathcal{A}$-free measures. As for the sufficiency, we can now prove the following general lower semicontinuity theorem, which is taken from Arroyo-Rabasa, Philippis, and Rindler [2017] (where also more general results can be found):

Theorem 3.1. *Let $f\colon \Omega \times \mathbb{R}^N \to [0,\infty)$ be a continuous integrand with linear growth at infinity such that f is uniformly Lipschitz in its second argument, f^∞ exists as in (3-2), and $f(x,\bullet)$ is $\mathcal{A}^k$-quasiconvex for all $x \in \Omega$. Further assume that there exists a modulus of continuity $\omega\colon [0,\infty) \to [0,\infty)$ (increasing, continuous, $\omega(0) = 0$) such that*

$$|f(x,A) - f(y,A)| \le \omega(|x-y|)(1+|A|) \qquad \textit{for all } x,y \in \Omega,\ A \in \mathbb{R}^N. \tag{3-3}$$

Then, the functional $\mathfrak{F}$ is sequentially weakly lower semicontinuous on the space*

$$\mathcal{M}(\Omega;\mathbb{R}^N) \cap \ker \mathcal{A} := \bigl\{ \mu \in \mathcal{M}(\Omega;\mathbb{R}^N) \;:\; \mathcal{A}\mu = 0 \bigr\}.$$

Remark 3.2. As special cases of Theorem 3.1 we get, among others, the following well-known results:

(i) For $\mathcal{A} = \operatorname{curl}$, one obtains BV-lower semicontinuity results in the spirit of Ambrosio and Dal Maso [1992] and Fonseca and Müller [1993].

(ii) For $\mathcal{A} = \operatorname{curl}\operatorname{curl}$, the second order operator expressing the Saint-Venant compatibility conditions, we re-prove the lower semicontinuity and relaxation theorem in the space of functions of bounded deformation (BD) from Rindler [2011].

(iii) For first-order operators $\mathcal{A}$, a similar result was proved in Baía, Chermisi, Matias, and Santos [2013].

The proof of Theorem 3.1 essentially follows by combining Theorem 1.2 with the main theorem of Kirchheim and Kristensen [2016], which establishes that the restriction of f^∞ to the linear space spanned by the wave cone is in fact *convex* at all points of $\Lambda_{\mathcal{A}}$ (in the sense that a supporting hyperplane exists). In this way we gain classical convexity for f^∞ at singular points, which can be exploited via the theory of generalized Young measures developed in DiPerna and Majda [1987], Alibert and Bouchitté [1997], and Kristensen and Rindler [2012] and also briefly discussed in the next section.

One can also show relaxation results, where f is not assumed to be $\mathcal{A}^k$-quasiconvex in the second argument and the task becomes to compute the largest weakly* lower semicontinuous functional below $\mathfrak{F}$; see Arroyo-Rabasa, Philippis, and Rindler [2017] for more details.

4 Characterization of generalized Young measures

Young measures quantitatively describe the asymptotic oscillations in L^p-weakly converging sequences. They were introduced in Young [1937, 1942a,b] and later developed into an important tool in modern PDE theory and the calculus of variations in Tartar [1979], Tartar [1983], Ball [1989], and Ball and James [1987] and many other works. In order to deal with concentration effects as well, DiPerna & Majda extended the framework to so-called "generalized" Young measures, see DiPerna and Majda [1987], Alibert and Bouchitté [1997], Kružík and Roubíček [1997], Fonseca, Müller, and Pedregal [1998], Sychev [1999], and Kristensen and Rindler [2012]. In the following we will refer also to these objects simply as "Young measures". We recall some basic theory, for which proofs and examples can be found in Alibert and Bouchitté [1997], Kristensen and Rindler [2012], and Rindler [2011].

Let again $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. For $f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N)$ we define

$$\mathbf{E}(\Omega;\mathbb{R}^N) := \left\{ f \in \mathrm{C}(\overline{\Omega} \times \mathbb{R}^N) \;:\; f^\infty \text{ exists in the sense (3-2)} \right\}.$$

A *(generalized) Young measure* $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N) \subset \mathbf{E}(\Omega;\mathbb{R}^N)^*$ on the open set $\Omega \subset \mathbb{R}^d$ with values in $\mathbb{R}^N$ is a triple $\nu = (\nu_x, \lambda_\nu, \nu_x^\infty)$ consisting of

(i) a parametrized family of probability measures $(\nu_x)_{x\in\Omega} \subset \mathfrak{M}_1(\mathbb{R}^N)$, called the *oscillation measure*;

(ii) a positive finite measure $\lambda_\nu \in \mathfrak{M}_+(\overline{\Omega})$, called the *concentration measure*; and

(iii) a parametrized family of probability measures $(\nu_x^\infty)_{x\in\overline{\Omega}} \subset \mathfrak{M}_1(\mathbb{S}^{N-1})$, called the *concentration-direction measure*,

for which we require that

(iv) the map $x \mapsto \nu_x$ is *weakly* measurable* with respect to $\mathcal{L}^d$, i.e. the function $x \mapsto \langle f(x,\bullet), \nu_x\rangle$ is $\mathcal{L}^d$-measurable for all bounded Borel functions $f : \Omega \times \mathbb{R}^N \to \mathbb{R}$,

(v) the map $x \mapsto \nu_x^\infty$ is weakly* measurable with respect to λ_ν, and

(vi) $x \mapsto \langle |\bullet|, \nu_x\rangle \in \mathrm{L}^1(\Omega)$.

The *duality pairing* between $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ and $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ is given as

$$\begin{aligned}\big\langle f,\nu\big\rangle &:= \int_\Omega \big\langle f(x,\bullet),\nu_x\big\rangle \,\mathrm{d}x + \int_{\overline{\Omega}} \big\langle f^\infty(x,\bullet),\nu_x^\infty\big\rangle \,\mathrm{d}\lambda_\nu(x)\\ &:= \int_\Omega \int_{\mathbb{R}^N} f(x,A)\,\mathrm{d}\nu_x(A)\,\mathrm{d}x + \int_{\overline{\Omega}} \int_{\partial\mathbb{B}^N} f^\infty(x,A)\,\mathrm{d}\nu_x^\infty(A)\,\mathrm{d}\lambda_\nu(x).\end{aligned}$$

If $(\gamma_j) \subset \mathfrak{M}(\overline{\Omega};\mathbb{R}^N)$ is a sequence of Radon measures with $\sup_j |\gamma_j|(\overline{\Omega}) < \infty$, then we say that the sequence (γ_j) *generates* a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$, in symbols $\gamma_j \xrightarrow{\mathbf{Y}} \nu$, if for all $f \in \mathbf{E}(\Omega;\mathbb{R}^N)$ it holds that

$$\begin{aligned}&f\left(x,\frac{\mathrm{d}\gamma_j}{\mathrm{d}\mathcal{L}^d}(x)\right)\mathcal{L}^d \llcorner \Omega + f^\infty\left(x,\frac{\mathrm{d}\gamma_j^s}{\mathrm{d}|\gamma_j^s|}(x)\right)|\gamma_j^s|(\mathrm{d}x)\\ &\quad\overset{*}{\rightharpoonup} \big\langle f(x,\bullet),\nu_x\big\rangle\mathcal{L}^d \llcorner \Omega + \big\langle f^\infty(x,\bullet),\nu_x^\infty\big\rangle \lambda_\nu(\mathrm{d}x) \qquad \text{in } \mathfrak{M}(\overline{\Omega}).\end{aligned}$$

Here, γ_j^s is the singular part of γ_j with respect to Lebesgue measure.

It can be shown that if $(\gamma_j) \subset \mathfrak{M}(\overline{\Omega};\mathbb{R}^N)$ is a sequence of measures with $\sup_j |\gamma_j|(\overline{\Omega}) < \infty$ as above, then there exists a subsequence (not relabeled) and a Young measure $\nu \in \mathbf{Y}(\Omega;\mathbb{R}^N)$ such that $\gamma_j \xrightarrow{\mathbf{Y}} \nu$, see Kristensen and Rindler [2012].

When considering generating sequences (γ_j) as above that satisfy a differential constraint like curl-freeness (i.e. the generating sequence is a sequence of *gradients*), the following question arises:

Question 4.1. *Can one characterize the class of Young measures generated by sequences satisfying some (linear) PDE constraint?*

In applications, such results provide valuable information on the allowed oscillations and concentrations that are possible under this differential constraint, which usually constitutes a strong restriction. Characterization theorems are of particular use in the relaxation of minimization problems for non-convex integral functionals, where one passes from a functional defined on functions to one defined on Young measures. A characterization theorem then allows one to restrict the class of Young measures over which to minimize. This strategy is explained in detail (for classical Young measures) in Pedregal [1997].

The first general classification results are due to Kinderlehrer and Pedregal [1991, 1994], who characterized classical *gradient* Young measures, i.e. those generated by gradients of $\mathrm{W}^{1,p}$-bounded sequences, $1 < p \leq \infty$. Their theorems put such gradient Young measures in duality with quasiconvex functions. For generalized Young measures the corresponding result was proved in Fonseca, Müller, and Pedregal [1998] (also see Kałamajska and Kružík [2008]) and numerous other characterization results in the spirit of the Kinderlehrer–Pedregal theorems have since appeared, see for instance Kružík and Roubíček [1996], Fonseca and Müller [1999], Fonseca and Kružík [2010], and Benešová and Kružík [2016].

The characterization of generalized BV-Young measures, i.e. those ν generated by a sequence (Du_j) of the BV-derivatives of maps $u_j \in \mathrm{BV}(\Omega;\mathbb{R}^\ell))$ was first achieved in Kristensen and Rindler [2012]. A different, "local" proof was given in Rindler [2014], another improvement is in Kirchheim and Kristensen [2016, Theorem 6.2]. All of these arguments crucially use Alberti's rank-one theorem.

The most interesting case beyond BV is again the case of functions of bounded deformation (BD), which were introduced above: In plasticity theory Suquet [1978], Suquet [1979], and Temam and Strang [1980/81], one often deals with sequences of uniformly L^1-bounded symmetric gradients $\mathcal{E}u_j := (\nabla u_j + \nabla u_j^T)/2$. In order to understand the asymptotic oscillations and concentrations in such sequences $(\mathcal{E}u_j)$ one needs to characterize the (generalized) Young measures ν generated by them. We call such ν *BD-Young measures* and write $\nu \in \mathbf{BDY}(\Omega)$, since all BD-functions can be reached as weak* limits of sequences (u_j) as above.

In this situation the following result can be shown, see De Philippis and Rindler [2017]:

Theorem 4.2. *Let $\nu \in \mathbf{Y}(\Omega;(\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}})$ be a (generalized) Young measure. Then, ν is a BD-Young measure, $\nu \in \mathbf{BDY}(\Omega)$, if and only if there exists $u \in \mathrm{BD}(\Omega)$ with*

$$\langle \mathrm{id}, \nu_x \rangle \mathcal{L}^d_x + \langle \mathrm{id}, \nu_x^\infty \rangle (\lambda_\nu \llcorner \Omega)(\mathrm{d}x) = Eu$$

and for all symmetric-quasiconvex $h \in \mathrm{C}((\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}})$ with linear growth at infinity, the Jensen-type inequality

$$h\left(\langle \mathrm{id}, \nu_x \rangle + \langle \mathrm{id}, \nu_x^\infty \rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x) \right) \leq \langle h, \nu_x \rangle + \langle h^\#, \nu_x^\infty \rangle \frac{\mathrm{d}\lambda_\nu}{\mathrm{d}\mathcal{L}^d}(x).$$

holds at $\mathcal{L}^d$-almost every $x \in \Omega$, where $h^\#$ is defined via

$$h^\#(A) := \limsup_{\substack{A' \to A \\ t \to \infty}} \frac{h(tA')}{t}, \qquad A \in (\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}}.$$

One application of this result (in the spirit of Young's original work Young [1942a,b, 1980]) is the following: For a suitable integrand $f : \Omega \times (\mathbb{R}^d \otimes \mathbb{R}^d)_{\mathrm{sym}} \to \mathbb{R}$, the minimum principle

$$\langle f, \nu \rangle \to \min, \quad \nu \in \mathbf{BDY}(\Omega). \tag{4-1}$$

can be seen as the *extension-relaxation* of the minimum principle

$$\int_\Omega f(x, \mathcal{E}u(x)) \,\mathrm{d}x + \int_\Omega f^\infty\left(x, \frac{\mathrm{d}E^s u}{\mathrm{d}|E^s u|}(x)\right) \mathrm{d}|E^s u| \to \min \tag{4-2}$$

over $u \in \mathrm{BD}(\Omega)$. The point is that (4-2) may not be solvable if f is not symmetric-quasiconvex, whereas (4-1) always has a solution. In this situation, Theorem 4.2 then gives (abstract) restrictions on the Young measures to be considered in (4-1). Another type of relaxation involving the symmetric-quasiconvex envelope of f is investigated in Arroyo-Rabasa, Philippis, and Rindler [2017] within the framework of general linear PDE side-constraints.

The necessity part of Theorem 4.2 follows from a lower semicontinuity or relaxation theorem like the one in Rindler [2011]. For the sufficiency part (which is quite involved), one first characterizes so-called *tangent Young measures*, which are localized versions of Young measures. There are two types: regular and singular tangent Young measures, depending on whether regular (Lebesgue measure-like) effects or singular effects dominate around the blow-up point. We stress that the argument crucially rests on the BD-analogue of Alberti's rank-one theorem, see Theorem 2.3 (ii). Technically, in one of the proof steps to establish Theorem 4.2 we need to create "artificial concentrations" by compressing symmetric gradients in one direction. This is only possible if we know precisely what these singularities look like.

A characterization results for Young measures under general linear PDE constrainsts is currently not available (there is a partial result in the work Baía, Matias, and Santos [2013], but limited to first-order operators and needing additional technical assumptions). The reason is that currently not enough is known about the directional structure of $\mathcal{A}$-free measures at singular points.

5 The converse of Rademacher's theorem

Rademacher's theorem asserts that a Lipschitz function $f \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^\ell)$ is diffferentiable $\mathfrak{L}^d$-almost everywhere. A natural question, which has attracted considerable attention, is to understand how sharp this result is. The following questions have been folklore in the area for a while:

Question 5.1 (Strong converse of Rademacher's theorem). *Given a Lebesgue null set $E \subset \mathbb{R}^d$ is it possible to find some $\ell \geq 1$ and a Lipschitz function $f \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^\ell)$ such that f is not differentiable in any point of E?*

Question 5.2 (Weak converse of Rademacher's theorem). *Let $\nu \in \mathfrak{M}_+(\mathbb{R}^d)$ be a positive Radon measure such that every Lipschitz function is differentiable ν-almost everywhere. Is ν necessarily absolutely continuous with respect to $\mathfrak{L}^d$?*

Clearly, a positive answer to Question 5.1 implies a positive answer to Question 5.2. Let us also stress that in answering Question 5.1, an important role is played by the dimension ℓ of the target set, see point (ii) below, while this does not have any influence on Question 5.2, see Alberti and Marchese [2016]. We refer to Alberti, Csörnyei, and Preiss [2005, 2010b] and Alberti and Marchese [2016] for a detailed account on the history of these problems and here we simply record the following facts:

(i) For $d = 1$ a positive answer to Question 5.1 is due to Zahorski [1946].

(ii) For $d \geq 2$ there exists a null set E such that every Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^\ell$ with $\ell < d$ is differentiable in at least one point of E. This is was proved by Preiss in Preiss [1990] for $d = 2$ and later extended by Preiss and Speight in Preiss and Speight [2015] to every dimension.

(iii) For $d = 2$ a positive answer to Question 5.1 has been given by Alberti, Csörnyei and Preiss as a consequence of their deep result concerning the structure of null sets in the plane Alberti, Csörnyei, and Preiss [2005, 2010b,a]. Namely, they show that for every null set $E \subset \mathbb{R}^2$ there exists a Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that f is not differentiable at any point of E.

(iv) For $d \geq 2$ an extension of the result described in point (iii) above, i.e. that for every null set $E \subset \mathbb{R}^d$ there exists a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^d$ such that f is not differentiable at any point of E, has been announced in 2011 by Jones [2011].

Let us now show how Question 5.2 is related to Question 1.1. In Alberti and Marchese [2016, Theorem 1.1] Alberti & Marchese have shown the following result:

Theorem 5.3 (Alberti–Marchese). *Let $\nu \in \mathfrak{M}_+(\mathbb{R}^d)$ be a positive Radon measure. Then, there exists a vector space-valued ν-measurable map $V(\nu, x)$ (the* decomposability bundle of ν*) such that:*

(i) Every Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$ is differentiable in the directions of $V(\nu, x)$ at ν-almost every x.

(ii) There exists a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$ such that for ν-almost every x and every $v \notin V(\nu, x)$ the derivative of f at x in the direction of v does not exist.

Thanks to the above theorem, Question 5.2 is then equivalent to the following:

Question 5.4. *Let $\nu \in \mathfrak{M}_+(\mathbb{R}^d)$ be a positive Radon measure such that $V(\nu, x) = \mathbb{R}^d$ for ν-almost every x. Is ν absolutely continuous with respect to $\mathcal{L}^d$?*

The link between the above question and Theorem 1.2 is due to the following result, again due to Alberti & Marchese, see Alberti and Marchese [2016, Corollary 6.5] and De Philippis and Rindler [2016, Lemma 3.1][2].

Lemma 5.5. *Let $\nu \in \mathfrak{M}_+(\mathbb{R}^d)$ be a positive Radon measure. Then the following are equivalent:*

1. *The decomposability bundle of ν is of full dimension, i.e. $V(\nu, x) = \mathbb{R}^d$ for ν-almost every x.*
2. *There exist $\mathbb{R}^d$-valued measures $\mu_1, \ldots, \mu_d \in \mathfrak{M}(\mathbb{R}^d; \mathbb{R}^d)$ with measure-valued divergences $\operatorname{div} \mu_i \in \mathfrak{M}(\mathbb{R}^d; \mathbb{R})$ such that $\nu \ll |\mu_i|$ for $1 = 1, \ldots, d$ and*[3]

$$\text{span}\left\{\frac{\mathrm{d}\mu_1}{\mathrm{d}|\mu_1|}(x), \ldots, \frac{\mathrm{d}\mu_d}{\mathrm{d}|\mu_d|}(x)\right\} = \mathbb{R}^d \qquad \text{for } \nu\text{-a.e. } x. \tag{5-1}$$

With the above lemma at hand, a positive answer to Question 5.4 (and thus to Question 5.2) follows from Theorem 1.2 in a straightforward fashion:

Theorem 5.6. *Let $\nu \in \mathfrak{M}_+(\mathbb{R}^d)$ be a positive Radon measure such that every Lipschitz function is differentiable ν-almost everywhere. Then, ν is absolutely continuous with respect to $\mathcal{L}^d$.*

[2]In the cited references the results are stated in terms of normal currents. By the trivial identifications of the space of normal currents with the space of measure-valued vector fields whose divergence is a measure it is immediate to see that they are equivalent to our Lemma 5.5

[3]Note that since $\nu \ll |\mu_i|$ for all $i = 1, \ldots, d$, in item (ii) above all the Radon-Nikodym derivatives $\frac{\mathrm{d}\mu_i}{\mathrm{d}|\mu_i|}$ $i = 1, \ldots, d$ exist for ν-a.e. x.

Proof. Let ν be a measure such that $V(\nu, x) = \mathbb{R}^d$ for ν-almost every x an let μ_i be the measures provided by Lemma 5.5 (ii). Let us consider the matrix-valued measure

$$\boldsymbol{\mu} := \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix} \in \mathfrak{M}(\mathbb{R}^d; \mathbb{R}^d \otimes \mathbb{R}^d).$$

Note that $\operatorname{div} \boldsymbol{\mu} \in \mathfrak{M}(\mathbb{R}^d; \mathbb{R}^d)$, where div is the row-wise divergence operator. Since, by direct computation,

$$\Lambda_{\mathrm{div}} = \{ M \in \mathbb{R}^d \otimes \mathbb{R}^d \, : \, \operatorname{rank} M \leq d - 1 \},$$

Theorem 1.2 and Remark 1.3 imply that $\operatorname{rank}\left(\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}|\boldsymbol{\mu}|}\right) \leq d-1$ for $|\boldsymbol{\mu}|^s$-almost every point. Hence, by (5-1), ν is singular with respect to $|\boldsymbol{\mu}|^s$. On the other hand, since $\nu \ll |\mu_i|$ for all $i = 1, \ldots, d$, we get $\nu^s \ll |\boldsymbol{\mu}|^s$. Hence, we conclude $\nu^s = 0$, as desired. □

Let us conclude this section by remarking that the weak converse of Rademacher's theorem, i.e. a positive answer to Question 5.2, has some consequences for the structure of Ambrosio–Kirchheim metric currents Ambrosio and Kirchheim [2000], see the work of Schioppa [2016b]. In particular, it allows one to prove the top-dimensional case of the flat chain conjecture proposed by Ambrosio and Kirchheim in Ambrosio and Kirchheim [2000].

6 Cheeger's conjecture

Among the many applications of Rademacher's theorem, it allows one to pass from "non-infintesimal" information (the existence of certain Lipschitz maps) to infinitesimal information. For instance, one can easily establish the following fact:

There is no bi-Lipschitz map $f : \mathbb{R}^d \to \mathbb{R}^\ell$ if $d \neq \ell$.

Indeed, if this were the case, Rademacher's theorem would imply (at a differentiability point) the existence of a bijective *linear* map from $\mathbb{R}^d$ to $\mathbb{R}^\ell$ with $d \neq \ell$.

While the above statement is an immediate consequence of the theorem on the invariance of dimension (asserting that there are no bijective continuous maps from $\mathbb{R}^d$ to $\mathbb{R}^\ell$ if $d \neq \ell$), the point here is that the almost everywhere result allows to pass from a non-linear statement (the existence of a bi-Lipschitz map) to a linear one, whose rigidity can be proved by elementary methods.

This line of thought has been adopted in the study of rigidity of several metric structures. For instance, the natural generalization of Rademacher's theorem in the context of Carnot groups, which was established by Pansu in Pansu [1989], allows one to show that there are no bi-Lipscitz embeddings of a Carnot group into $\mathbb{R}^\ell$ if the former is non-commutative.

The fact that (a suitable notion of) differentiability of Lipschitz functions allows to develop a first-order calculus on metric spaces and in turn to obtain non-embedding results has been recognised by Cheeger in his seminal paper Cheeger [1999] and later studied by several authors.

Let us briefly introduce the theory of Cheeger as it has been axiomatized by Keith in Keith [2004]. Note that it is natural to generalize Rademacher's theorem to the setting of *metric measure spaces* since we need to talk about Lipschitz functions (a metric concept) and almost everywhere (a measure-theoretic concept).

Let (X, ρ, μ) be a *metric measure space*, that is, (X, ρ) is a separable, complete metric space and $\mu \in \mathfrak{M}_+(X)$ is a positive Radon measure on X. We call a pair (U, φ) such that $U \subset X$ is a Borel set and $\varphi\colon X \to \mathbb{R}^d$ is Lipschitz, a *d-dimensional chart*, or simply a *d-chart*. A function $f : X \to \mathbb{R}$ is said to be *differentiable with respect to a d-chart* (U, φ) at $x_0 \in U$ if there exists a unique (co-)vector $df(x_0) \in \mathbb{R}^d$ such that

$$\limsup_{x \to x_0} \frac{|f(x) - f(x_0) - df(x_0) \cdot (\varphi(x) - \varphi(x_0))|}{\rho(x, x_0)} = 0. \tag{6-1}$$

Definition 6.1. *A metric measure space (X, ρ, μ) is a* Lipschitz differentiability space *if there exists a countable family of $d(i)$-charts (U_i, φ_i) ($i \in \mathbb{N}$) such that $X = \bigcup_i U_i$ and any Lipschitz map $f : X \to \mathbb{R}$ is differentiable with respect to every (U_i, φ_i) at μ-almost every point $x_0 \in U_i$.*

In this terminology, the main result of Cheeger [1999] asserts that every doubling metric measure space (X, ρ, μ) satisfying a Poincaré inequality is a Lipschitz differentiability space.

In the same paper, Cheeger conjectured that the push-forward of the reference measure μ under every chart φ_i has to be absolutely continuous with respect to the Lebesgue measure, see Cheeger [ibid., Conjecture 4.63]:

Question 6.2. *For every d-chart (U, φ) in a Lipschitz differentiability space, does it hold that $\varphi_\#(\mu \llcorner U) \ll \mathcal{L}^d$?*

Some consequences of this fact concerning the existence of bi-Lipschitz embeddings of X into some $\mathbb{R}^N$ are detailed in Cheeger [ibid., Section 14], also see Cheeger and Kleiner [2006, 2009].

Let us assume that $(X, \rho, \mu) = (\mathbb{R}^d, \rho_{\mathcal{E}}, \nu)$ with $\rho_{\mathcal{E}}$ the Euclidean distance and ν a positive Radon measure, is a Lipschitz differentiability space when equipped with the (single) identity chart (note that it follows a-posteriori from the validity of Cheeger's conjecture that no mapping into a higher-dimensional space can be a chart in a Lipschitz differentiability structure of $\mathbb{R}^d$). In this case the validity of Cheeger's conjecture reduces to the validity of the (weak) converse of Rademacher's theorem, which we stated above in Theorem 5.6.

One can also prove the assertion of Cheeger's conjecture directly. Indeed, from the work of Bate [2015], and Alberti and Marchese [2016] an analogue of Lemma 5.5 for $\varphi_\#(\mu \llcorner U)$ in place of μ follows, see also Schioppa [2016a,b]. This allows one to conclude as in the proof of Theorem 5.6 to get:

Theorem 6.3. *Let (X, ρ, μ) be a Lipschitz differentiability space and let (U, φ) be a d-dimensional chart. Then, $\varphi_\#(\mu \llcorner U) \ll \mathcal{L}^d$.*

The details can be found in De Philippis, Marchese, and Rindler [2017].

We conclude this section by mentioning that the weak converse of Rademacher's theorem also has some consequences concerning the structure of measures on metric measure spaces with Ricci curvature bounded from below, see Kell and Mondino [2016] and Gigli and Pasqualetto [2016].

7 Sketch of the proof of Theorem 1.2

In this section we shall give some details concerning the proof of Theorem 1.2. For simplicity we will only consider the case in which $\mathcal{A}$ is a *first-order homogeneous* operator, namely we will assume that μ satisfies

$$\mathcal{A}\mu = \sum_{j=1}^{d} A_j \partial_j \mu = 0 \qquad \text{in the sense of distributions.}$$

Note that in this case we have

$$\Lambda_{\mathcal{A}} = \bigcup_{|\xi|=1} \ker \mathbb{A}(\xi), \qquad \mathbb{A}(\xi) = \mathbb{A}^1(\xi) = 2\pi \mathrm{i} \sum_{j=1}^{d} A_j \xi_j .$$

Let

$$E := \left\{ x \in \Omega \;:\; \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) \notin \Lambda_{\mathcal{A}} \right\},$$

and let us assume by contradiction that $|\mu|^s(E) > 0$.

Employing a fundamental technique of geometric measure theory, one can "zoom in" around a generic point of E. Indeed, one can show that for $|\mu|^s$-almost every point $x_0 \in E$

there exists a sequence of radii $r_k \downarrow 0$ such that

$$\operatorname*{w^*\text{-}lim}_{k\to\infty} \frac{(T^{x_0,r_k})_\# \mu}{|\mu|(B_{r_k}(x_0))} = \operatorname*{w^*\text{-}lim}_{k\to\infty} \frac{(T^{x_0,r_j})_\# \mu^s}{|\mu|^s(B_{r_k}(x_0))} = P_0 \nu,$$

where $T^{x,r} \colon \mathbb{R}^d \to \mathbb{R}^d$ is the dilation map $T^{x,r}(y) := (y-x)/r$, $T^{x,r}_\#$ denotes the push-forward operator[4], $\nu \in \mathrm{Tan}(x_0, |\mu|) = \mathrm{Tan}(x_0, |\mu|^s)$ is a non-zero tangent measure in the sense of Preiss [1987],

$$\lambda_0 = \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) \notin \Lambda_{\mathfrak{A}},$$

and the limit is to be understood in the weak* topology of Radon measures (i.e. in duality with compactly supported continuous functions). Moreover, one easily checks that

$$\sum_{j=1}^d A_j \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0)\, \partial_j \nu = 0 \qquad \text{in the sense of distributions.}$$

By taking the Fourier transform of the above equation, we get

$$[\mathbb{A}(\xi)\lambda_0]\, \hat{\nu}(\xi) = 0, \qquad \xi \in \mathbb{R}^d.$$

where $\hat{\nu}(\xi)$ is the Fourier transform of ν in the sense of distributions (actually, ν does not need to be a tempered distribution, hence some care is needed, see below for more details). Having assumed that $\lambda_0 \notin \Lambda_{\mathfrak{A}}$, i.e. that

$$\mathbb{A}(\xi)\lambda_0 \neq 0 \quad \text{for all } \xi \neq 0,$$

this implies $\operatorname{supp} \hat{\nu} = \{0\}$ and thus $\nu \ll \mathcal{L}^d$. The latter fact, however, is not by itself a contradiction to $\nu \in \mathrm{Tan}(x_0, |\mu|^s)$. Indeed, Preiss [ibid.] provided an example of a purely singular measure that has only multiples of Lebesgue measure as tangents (we also refer to O'Neil [1995] for a measure that has *every* measure as a tangent at almost every point).

The above reasoning provides a sort of *rigidity* property for $\mathfrak{A}$-measures: If, for a constant polar vector $\lambda_0 \notin \Lambda_{\mathfrak{A}}$ and a measure $\nu \in \mathfrak{M}_+(\mathbb{R}^d)$, the measure $\lambda_0 \nu$ is $\mathfrak{A}$-free, then necessarily $\nu \ll \mathcal{L}^d$. However, as we commented above, this is not enough to conclude. In order to prove the theorem we need to strengthen this rigidity property (absolutely continuity of the measures $\lambda_0 \nu$ with $\lambda_0 \notin \Lambda_{\mathfrak{A}}$) to a stability property which can be roughly stated as follows:

$\mathfrak{A}$-free measures μ with $|\mu|\big(\{x : \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) \in \Lambda_{\mathfrak{A}}\}\big) \ll 1$ have small singular part.

[4] That is, for any measure σ and Borel set B, $[(T^{x,r})_\# \sigma](B) := \sigma(x + rB)$

In this respect note that since $\lambda_0 \notin \Lambda_{\mathcal{A}}$ implies that $\mathbb{A}(\xi)\lambda_0 \neq 0$ for $\xi \neq 0$, one can hope for some sort of "elliptic regularization" that forces not only $\nu \ll \mathcal{L}^d$ but also

$$\mu_k := \frac{(T^{x_0,r_k})_\# \mu}{|\mu|^s(B_{r_k}(x_0))} \ll \mathcal{L}^d,$$

at least for small r_k. This is actually the case: Inspired by Allard's strong constancy lemma in Allard [1986], we can show that the ellipticity of the system at the limit (i.e. that $\mathbb{A}(\xi)\lambda_0 \neq 0$) improves the weak* convergence of (μ_k) to convergence in the total variation norm, i.e.

$$|\mu_k - \lambda_0 \nu|(B_{1/2}) \to 0. \tag{7-1}$$

Since the singular part of μ_k is asymptotically predominant around x_0, see (7-2) below, this latter fact implies that

$$|\mu_k^s - \lambda_0 \nu|(B_{1/2}) \to 0,$$

which easily gives a contradiction to $\nu \ll \mathcal{L}^d$ and concludes the proof.

Let us briefly sketch how (7-1) is obtained. For $\chi \in \mathcal{D}(B_1)$, $0 \leq \chi \leq 1$, consider the measures $\lambda_0 \chi \nu_k$, where

$$\nu_k := \frac{(T^{x_0,r_k})_\# |\mu|^s}{|\mu|^s(B_{r_k}(x_0))},$$

and note that, since we can assume that for the chosen x_0 it holds that

$$\frac{|\mu|^a(B_{r_k}(x_0))}{|\mu|^s(B_{r_k}(x_0))} \to 0, \qquad \fint_{B_{r_k}(x_0)} \left| \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x) - \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}(x_0) \right| \mathrm{d}|\mu|^s(x) \to 0, \tag{7-2}$$

we have that

$$|\lambda_0 \chi \nu_k - \chi \mu_k|(\mathbb{R}^d) \leq |\lambda_0 \nu_k - \mu_k|(B_1) \to 0. \tag{7-3}$$

Using the $\mathcal{A}$-freeness of μ_k (which trivially follows from the one of μ) we can derive an equation for $\chi \nu_k$:

$$\sum_{j=1}^d A_j \lambda_0 \partial_j (\chi \nu_k) = \sum_{j=1}^d A_j \partial_j (\lambda_0 \chi \nu_k - \chi \mu_k) + \sum_{j=1}^d A_j \mu_k \partial_j \chi. \tag{7-4}$$

Since we are essentially dealing with a-priori estimates, in the following we treat measures as if they were smooth L^1-functions; this can be achieved by a sufficiently fast regularization, see De Philippis and Rindler [2016] for more details.

Taking the Fourier transform of equation (7-4) (note that we are working with compactly supported functions) we obtain

$$\mathbb{A}(\xi)\lambda_0\widehat{\chi\nu_k}(\xi) = \mathbb{A}(\xi)\widehat{V_k}(\xi) + \widehat{R_k}(\xi), \tag{7-5}$$

where

$$V_k := \lambda_0\chi\nu_k - \chi\mu_k \qquad \text{satisfies} \qquad |V_k|(\mathbb{R}^d) \to 0 \tag{7-6}$$

and

$$R_k := \sum_{j=1}^{d} A_j\mu_k\partial_j\chi \qquad \text{satisfies} \qquad \sup_k |R_k|(\mathbb{R}^d) \leq C. \tag{7-7}$$

Scalar multiplying (7-5) by $\overline{\mathbb{A}(\xi)P_0}$, adding $\widehat{\chi\nu_k}$ to both sides and rearranging the terms, we arrive to

$$\begin{aligned}\widehat{\chi\nu_k}(\xi) &= \frac{\overline{\mathbb{A}(\xi)\lambda_0}\mathbb{A}(\xi)\widehat{V_k}(\xi)}{1+|\mathbb{A}(\xi)\lambda_0|^2} + \frac{\overline{\mathbb{A}(\xi)\lambda_0}\cdot\widehat{R_k}(\xi)}{1+|\mathbb{A}(\xi)\lambda_0|^2} + \frac{\widehat{\chi\nu_k}(\xi)}{1+|\mathbb{A}(\xi)\lambda_0|^2} \\ &=: T_0(V_k) + T_1(R_k) + T_2(\chi\nu_k),\end{aligned} \tag{7-8}$$

where

$$\begin{aligned}T_0[V] &= \mathfrak{F}^{-1}\big[m_0(\xi)\hat{V}(\xi)\big], \\ T_1[R] &= \mathfrak{F}^{-1}\big[m_1(\xi)(1+4\pi^2|\xi|^2)^{-1/2}\hat{R}(\xi)\big], \\ T_2[u] &= \mathfrak{F}^{-1}\big[m_2(\xi)(1+4\pi^2|\xi|^2)^{-1}\hat{u}(\xi)\big],\end{aligned}$$

and we have set

$$\begin{aligned}m_0(\xi) &:= (1+|\mathbb{A}(\xi)\lambda_0|^2)^{-1}\overline{\mathbb{A}(\xi)\lambda_0}\mathbb{A}(\xi), \\ m_1(\xi) &:= (1+|\mathbb{A}(\xi)\lambda_0|^2)^{-1}(1+4\pi^2|\xi|^2)^{1/2}\overline{\mathbb{A}(\xi)\lambda_0}, \\ m_2(\xi) &:= (1+|\mathbb{A}(\xi)\lambda_0|^2)^{-1}(1+4\pi^2|\xi|^2).\end{aligned}$$

We now note that since $\lambda_0 \notin \Lambda_{\mathfrak{A}}$, by homogeneity there exists $c > 0$ such that $|\mathbb{A}(\xi)\lambda_0| \geq c|\xi|$ (this is the ellipticity condition we mentioned at the beginning). Hence, the symbols m_i, $i = 1, 2, 3$, satisfy the assumptions of the Hörmander–Mihlin multiplier theorem Grafakos [2014, Theorem 5.2.7], i.e. there exists constants $K_\beta > 0$ such that

$$|\partial^\beta m_i(\xi)| \leq K_\beta|\xi|^{-|\beta|} \qquad \text{for all } \beta \in \mathbb{N}^d.$$

This implies that T_0 is a bounded operator from L^1 to $\mathrm{L}^{1,\infty}$ and thus, thanks to (7-6), we get

$$\|T_0(V_k)\|_{\mathrm{L}^{1,\infty}} \leq C|V_k|(\mathbb{R}^d) \to 0. \tag{7-9}$$

Moreover,

$$\langle T_0(V_k), \varphi\rangle = \langle V_k, T_0^*(\varphi)\rangle \to 0 \qquad \text{for every } \varphi \in \mathfrak{D}(\mathbb{R}^d). \tag{7-10}$$

where T_0^* is the adjoint operator of T_0. We also observe

$$T_1 = Q_{m_1} \circ (\mathrm{Id} - \Delta)^{-1/2} \qquad \text{and} \qquad T_2 = Q_{m_2} \circ (\mathrm{Id} - \Delta)^{-1},$$

where Q_{m_1} and Q_{m_2} are the Fourier multipliers operators associated with the symbols m_1 and m_2, respectively. In particular, again by the Hörmander–Mihlin multiplier theorem, these operators are bounded from L^p to L^p for every $p \in (1, \infty)$. Moreover, $(\mathrm{Id}-\Delta)^{-s/2}$ is a compact operator from[5] $\mathrm{L}^1_c(B_1)$ to L^q for some $q = q(d, s) > 1$, see for instance De Philippis and Rindler [2016, Lemma 2.1]. In conclusion, by (7-7) and $\sup_k |\chi\nu_k|(\mathbb{R}^d) \leq C$,

$$\{T_1(R_k) + T_2(\chi\nu_k)\}_{k\in\mathbb{N}} \qquad \text{is pre-compact in } \mathrm{L}^1(B_1). \tag{7-11}$$

Hence, combining equation (7-8) with (7-9), (7-10) and (7-11) implies that

$$\chi\nu_k = u_k + w_k,$$

where $u_k \to 0$ in $\mathrm{L}^{1,\infty}$, $u_k \overset{*}{\rightharpoonup} 0$ in the sense of distributions and (w_k) is pre-compact in $\mathrm{L}^1(B_1)$. Since $\chi\nu_k \geq 0$,

$$u_k^- := \max\{-u_k, 0\} \leq |w_k|,$$

so that the sequence (u_k^-) is pre-compact in $\mathrm{L}^1(B_1)$. Since $u_k \to 0$ in $\mathrm{L}^{1,\infty}$, Vitali's convergence theorem implies that $u_k^- \to 0$ in $\mathrm{L}^1(B_1)$ which, combined with $u_k \overset{*}{\rightharpoonup} 0$, easily yields that $u_k \to 0$ in $\mathrm{L}^1(B_1)$, see De Philippis and Rindler [ibid., Lemma 2.2]. In conclusion, $(\chi\nu_k)$ is pre-compact in $\mathrm{L}^1(B_1)$. Together with (7-1) and the weak* convergence of μ_k to $\lambda_0\nu$, this implies

$$|\mu_k - \lambda_0\nu|(B_{1/2}) \to 0,$$

which concludes the proof.

[5]Here we denote by $\mathrm{L}^1_c(B_1)$ the space of L^1-functions vanishing outside B_1.

References

Giovanni Alberti (1993). "Rank one property for derivatives of functions with bounded variation". *Proc. Roy. Soc. Edinburgh Sect. A* 123.2, pp. 239–274. MR: 1215412 (cit. on p. 2236).

Giovanni Alberti, Marianna Csörnyei, and David Preiss (2005). "Structure of null sets in the plane and applications". In: *European Congress of Mathematics*. Eur. Math. Soc., Zürich, pp. 3–22. MR: 2185733 (cit. on p. 2245).

– (2010a). "Differentiability of Lipschitz functions, structure of null sets, and other problems". In: *Proceedings of the International Congress of Mathematicians. Volume III*. Hindustan Book Agency, New Delhi, pp. 1379–1394. MR: 2827846 (cit. on p. 2245).

– (2010b). "Differentiability of Lipschitz functions, structure of null sets, and other problems". In: *Proceedings of the International Congress of Mathematicians. Volume III*. Hindustan Book Agency, New Delhi, pp. 1379–1394. MR: 2827846 (cit. on p. 2245).

Giovanni Alberti and Andrea Marchese (2016). "On the differentiability of Lipschitz functions with respect to measures in the Euclidean space". *Geom. Funct. Anal.* 26.1, pp. 1–66. MR: 3494485 (cit. on pp. 2245, 2246, 2249).

J. J. Alibert and G. Bouchitté (1997). "Non-uniform integrability and generalized Young measures". *J. Convex Anal.* 4.1, pp. 129–147. MR: 1459885 (cit. on p. 2241).

William K. Allard (1986). "An integrality theorem and a regularity theorem for surfaces whose first variation with respect to a parametric elliptic integrand is controlled". In: *Geometric measure theory and the calculus of variations (Arcata, Calif., 1984)*. Vol. 44. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 1–28. MR: 840267 (cit. on p. 2251).

L. Ambrosio and E. De Giorgi (1988). "Un nuovo tipo di funzionale del calcolo delle variazioni". *Atti Acc. Naz. dei Lincei, Rend. Cl. Sc. Fis. Mat. Natur.* 82, pp. 199–210 (cit. on p. 2236).

Luigi Ambrosio (2004). "Transport equation and Cauchy problem for BV vector fields". *Invent. Math.* 158.2, pp. 227–260. MR: 2096794 (cit. on p. 2237).

Luigi Ambrosio, Alessandra Coscia, and Gianni Dal Maso (1997). "Fine properties of functions with bounded deformation". *Arch. Rational Mech. Anal.* 139.3, pp. 201–238. MR: 1480240 (cit. on p. 2238).

Luigi Ambrosio and Gianni Dal Maso (1992). "On the relaxation in $\mathrm{BV}(\Omega; \mathbf{R}^m)$ of quasi-convex integrals". *J. Funct. Anal.* 109.1, pp. 76–97. MR: 1183605 (cit. on pp. 2237, 2239, 2240).

Luigi Ambrosio, Nicola Fusco, and Diego Pallara (2000). *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, pp. xviii+434. MR: 1857292 (cit. on pp. 2236, 2237).

Luigi Ambrosio and Bernd Kirchheim (2000). "Currents in metric spaces". *Acta Math.* 185.1, pp. 1–80. MR: 1794185 (cit. on p. 2247).

Adolfo Arroyo-Rabasa, Guido De Philippis, and Filip Rindler (Jan. 2017). "Lower semicontinuity and relaxation of linear-growth integral functionals under PDE constraints". arXiv: 1701.02230 (cit. on pp. 2240, 2241, 2244).

Margarida Baía, Milena Chermisi, José Matias, and Pedro M. Santos (2013). "Lower semicontinuity and relaxation of signed functionals with linear growth in the context of $\mathcal{A}$-quasiconvexity". *Calc. Var. Partial Differential Equations* 47.3-4, pp. 465–498. MR: 3070552 (cit. on pp. 2239, 2241).

Margarida Baía, José Matias, and Pedro M. Santos (2013). "Characterization of generalized Young measures in the $\mathcal{A}$-quasiconvexity context". *Indiana Univ. Math. J.* 62.2, pp. 487–521. MR: 3158518 (cit. on p. 2244).

J. M. Ball (1989). "A version of the fundamental theorem for Young measures". In: *PDEs and continuum models of phase transitions (Nice, 1988)*. Vol. 344. Lecture Notes in Phys. Springer, Berlin, pp. 207–215. MR: 1036070 (cit. on p. 2241).

J. M. Ball and R. D. James (1987). "Fine phase mixtures as minimizers of energy". *Arch. Rational Mech. Anal.* 100.1, pp. 13–52. MR: 906132 (cit. on p. 2241).

David Bate (2015). "Structure of measures in Lipschitz differentiability spaces". *J. Amer. Math. Soc.* 28.2, pp. 421–482. MR: 3300699 (cit. on p. 2249).

Barbora Benešová and Martin Kružík (Jan. 2016). "Weak lower semicontinuity of integral functionals and applications". arXiv: 1601.00390 (cit. on p. 2243).

Guy Bouchitté, Irene Fonseca, and Luisa Mascarenhas (1998). "A global method for relaxation". *Arch. Rational Mech. Anal.* 145.1, pp. 51–98. MR: 1656477 (cit. on p. 2237).

J. Cheeger (1999). "Differentiability of Lipschitz functions on metric measure spaces". *Geom. Funct. Anal.* 9.3, pp. 428–517. MR: 1708448 (cit. on p. 2248).

Jeff Cheeger and Bruce Kleiner (2006). "On the differentiability of Lipschitz maps from metric measure spaces to Banach spaces". In: *Inspired by S. S. Chern*. Vol. 11. Nankai Tracts Math. World Sci. Publ., Hackensack, NJ, pp. 129–152. MR: 2313333 (cit. on p. 2248).

– (2009). "Differentiability of Lipschitz maps from metric measure spaces to Banach spaces with the Radon-Nikodým property". *Geom. Funct. Anal.* 19.4, pp. 1017–1028. MR: 2570313 (cit. on p. 2248).

Elisabetta Chiodaroli, Camillo De Lellis, and Ondřej Kreml (2015). "Global ill-posedness of the isentropic system of gas dynamics". *Comm. Pure Appl. Math.* 68.7, pp. 1157–1190. MR: 3352460 (cit. on p. 2234).

Elisabetta Chiodaroli, Eduard Feireisl, Ondřej Kreml, and Emil Wiedemann (2017). "$\mathfrak{A}$-free rigidity and applications to the compressible Euler system". *Ann. Mat. Pura Appl. (4)* 196.4, pp. 1557–1572. MR: 3673680 (cit. on p. 2234).

Sergio Conti, Daniel Faraco, and Francesco Maggi (2005). "A new approach to counterexamples to L^1 estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions". *Arch. Ration. Mech. Anal.* 175.2, pp. 287–300. MR: 2118479 (cit. on p. 2238).

Camillo De Lellis and László Székelyhidi (2013). "Continuous dissipative Euler flows and a conjecture of Onsager". In: *European Congress of Mathematics*. Eur. Math. Soc., Zürich, pp. 13–29. MR: 3469113 (cit. on p. 2234).

Camillo De Lellis and László Székelyhidi Jr. (2009). "The Euler equations as a differential inclusion". *Ann. of Math. (2)* 170.3, pp. 1417–1436. MR: 2600877 (cit. on p. 2234).

– (2012). "The h-principle and the equations of fluid dynamics". *Bull. Amer. Math. Soc. (N.S.)* 49.3, pp. 347–375. MR: 2917063 (cit. on p. 2234).

Guido De Philippis, Andrea Marchese, and Filip Rindler (2017). "On a conjecture of Cheeger". In: *Measure theory in non-smooth spaces*. Partial Differ. Equ. Meas. Theory. De Gruyter Open, Warsaw, pp. 145–155. MR: 3701738 (cit. on p. 2249).

Guido De Philippis and Filip Rindler (2016). "On the structure of $\mathfrak{A}$-free measures and applications". *Ann. of Math. (2)* 184.3, pp. 1017–1039. MR: 3549629 (cit. on pp. 2234, 2246, 2251, 2253).

– (2017). "Characterization of generalized Young measures generated by symmetric gradients". *Arch. Ration. Mech. Anal.* 224.3, pp. 1087–1125. MR: 3621818 (cit. on p. 2243).

Ronald J. DiPerna (1985). "Compensated compactness and general systems of conservation laws". *Trans. Amer. Math. Soc.* 292.2, pp. 383–420. MR: 808729 (cit. on p. 2234).

Ronald J. DiPerna and Andrew J. Majda (1987). "Oscillations and concentrations in weak solutions of the incompressible fluid equations". *Comm. Math. Phys.* 108.4, pp. 667–689. MR: 877643 (cit. on p. 2241).

Irene Fonseca and Martin Kružík (2010). "Oscillations and concentrations generated by $\mathcal{A}$-free mappings and weak lower semicontinuity of integral functionals". *ESAIM Control Optim. Calc. Var.* 16.2, pp. 472–502. MR: 2654203 (cit. on p. 2243).

Irene Fonseca, Giovanni Leoni, and Stefan Müller (2004). "$\mathcal{A}$-quasiconvexity: weak-star convergence and the gap". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 21.2, pp. 209–236. MR: 2021666 (cit. on p. 2239).

Irene Fonseca and Stefan Müller (1993). "Relaxation of quasiconvex functionals in $\mathrm{BV}(\Omega, \mathbf{R}^p)$ for integrands $f(x, u, \nabla u)$". *Arch. Rational Mech. Anal.* 123.1, pp. 1–49. MR: 1218685 (cit. on pp. 2237, 2239, 2240).

– (1999). "$\mathcal{A}$-quasiconvexity, lower semicontinuity, and Young measures". *SIAM J. Math. Anal.* 30.6, pp. 1355–1390. MR: 1718306 (cit. on pp. 2239, 2240, 2243).

Irene Fonseca, Stefan Müller, and Pablo Pedregal (1998). "Analysis of concentration and oscillation effects generated by gradients". *SIAM J. Math. Anal.* 29.3, pp. 736–756. MR: 1617712 (cit. on pp. 2241, 2243).

Gilles A. Francfort, Alessandro Giacomini, and Jean-Jacques Marigo (2015). "The taming of plastic slips in von Mises elasto-plasticity". *Interfaces Free Bound.* 17.4, pp. 497–516. MR: 3450737 (cit. on p. 2238).

Nicola Gigli and Enrico Pasqualetto (July 2016). "Behaviour of the reference measure on RCD spaces under charts". arXiv: 1607.05188 (cit. on p. 2249).

Loukas Grafakos (2014). *Classical Fourier analysis*. Third. Vol. 249. Graduate Texts in Mathematics. Springer, New York, pp. xviii+638. MR: 3243734 (cit. on p. 2252).

Philip Isett (Aug. 2016). "A Proof of Onsager's Conjecture". arXiv: 1608.08301 (cit. on p. 2234).

P. Jones (2011). *Product formulas for measures and applications to analysis and geometry*. Talk given at the conference *Geometric and algebraic structures in mathematics*, Stony Brook University, May 2011 (cit. on p. 2245).

Agnieszka Kałamajska and Martin Kružík (2008). "Oscillations and concentrations in sequences of gradients". *ESAIM Control Optim. Calc. Var.* 14.1, pp. 71–104. MR: 2375752 (cit. on p. 2243).

Stephen Keith (2004). "A differentiable structure for metric measure spaces". *Adv. Math.* 183.2, pp. 271–315. MR: 2041901 (cit. on p. 2248).

Martin Kell and Andrea Mondino (July 2016). "On the volume measure of non-smooth spaces with Ricci curvature bounded below". To appear *Ann. Sc. Norm. Super. Pisa Cl. Sci. (2016)*. arXiv: 1607.02036 (cit. on p. 2249).

David Kinderlehrer and Pablo Pedregal (1991). "Characterizations of Young measures generated by gradients". *Arch. Ration. Mech. Anal.* 115.4, pp. 329–365. MR: 1120852 (cit. on p. 2243).

– (1994). "Gradient Young measures generated by sequences in Sobolev spaces". *J. Geom. Anal.* 4.1, pp. 59–90. MR: 1274138 (cit. on p. 2243).

Bernd Kirchheim and Jan Kristensen (2016). "On rank one convex functions that are homogeneous of degree one". *Arch. Ration. Mech. Anal.* 221.1, pp. 527–558. MR: 3483901 (cit. on pp. 2238, 2241, 2243).

Jan Kristensen and Filip Rindler (2010). "Relaxation of signed integral functionals in BV". *Calc. Var. Partial Differential Equations* 37.1-2, pp. 29–62. MR: 2564396 (cit. on pp. 2237, 2239).

– (2012). "Erratum to: Characterization of generalized gradient Young measures generated by sequences in $W^{1,1}$ and BV [MR2660519]". *Arch. Ration. Mech. Anal.* 203.2, pp. 693–700. MR: 2885572 (cit. on pp. 2237, 2241–2243).

Martin Kružík and Tomáš Roubíček (1996). "Explicit characterization of L^p-Young measures". *J. Math. Anal. Appl.* 198.3, pp. 830–843. MR: 1377827 (cit. on p. 2243).

– (1997). "On the measures of DiPerna and Majda". *Math. Bohem.* 122.4, pp. 383–399. MR: 1489400 (cit. on p. 2241).

Annalisa Massaccesi and Davide Vittone (Jan. 2016). "An elementary proof of the rank-one theorem for BV functions". To appear in *J. Eur. Math. Soc. (2016)*. arXiv: 1601.02903 (cit. on p. 2237).

Charles B. Morrey Jr. (1952). "Quasi-convexity and the lower semicontinuity of multiple integrals". *Pacific J. Math.* 2, pp. 25–53. MR: 0054865 (cit. on p. 2240).

François Murat (1978). "Compacité par compensation". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 5.3, pp. 489–507. MR: 506997 (cit. on p. 2234).

– (1979). "Compacité par compensation. II". In: *Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis Rome, 1978*. Bologna: Pitagora Editrice, pp. 245–256 (cit. on p. 2234).

Toby O'Neil (1995). "A measure with a large set of tangent measures". *Proc. Amer. Math. Soc.* 123.7, pp. 2217–2220. MR: 1264826 (cit. on p. 2250).

Donald Ornstein (1962). "A non-equality for differential operators in the L_1 norm". *Arch. Rational Mech. Anal.* 11, pp. 40–49. MR: 0149331 (cit. on p. 2238).

Pierre Pansu (1989). "Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un". *Ann. of Math. (2)* 129.1, pp. 1–60. MR: 979599 (cit. on p. 2248).

Pablo Pedregal (1997). *Parametrized measures and variational principles*. Vol. 30. Prog. in Nonlinear Differential Equations and their Applications. Birkhäuser Verlag, Basel, pp. xii+212. MR: 1452107 (cit. on p. 2243).

D. Preiss (1990). "Differentiability of Lipschitz functions on Banach spaces". *J. Funct. Anal.* 91.2, pp. 312–345. MR: 1058975 (cit. on p. 2245).

David Preiss (1987). "Geometry of measures in $\mathbf{R}^n$: distribution, rectifiability, and densities". *Ann. of Math. (2)* 125.3, pp. 537–643. MR: 890162 (cit. on p. 2250).

David Preiss and Gareth Speight (2015). "Differentiability of Lipschitz functions in Lebesgue null sets". *Invent. Math.* 199.2, pp. 517–559. MR: 3302120 (cit. on p. 2245).

Filip Rindler (2011). "Lower semicontinuity for integral functionals in the space of functions of bounded deformation via rigidity and Young measures". *Arch. Ration. Mech. Anal.* 202.1, pp. 63–113. MR: 2835863 (cit. on pp. 2239, 2241, 2244).

– (2014). "A local proof for the characterization of Young measures generated by sequences in BV". *J. Funct. Anal.* 266.11, pp. 6335–6371. MR: 3192455 (cit. on pp. 2237, 2243).

Andrea Schioppa (2016a). "Derivations and Alberti representations". *Adv. Math.* 293, pp. 436–528. MR: 3474327 (cit. on p. 2249).

– (2016b). "Metric currents and Alberti representations". *Journal Funct. Anal.* 271.11, pp. 3007–3081. MR: 3554700 (cit. on pp. 2247, 2249).

Pierre-M. Suquet (1979). "Un espace fonctionnel pour les équations de la plasticité". *Ann. Fac. Sci. Toulouse Math. (5)* 1.1, pp. 77–87. MR: 533600 (cit. on pp. 2237, 2243).

Pierre-Marie Suquet (1978). "Existence et régularité des solutions des équations de la plasticité". *C. R. Acad. Sci. Paris Sér. A-B* 286.24, A1201–A1204. MR: 501114 (cit. on pp. 2237, 2243).

M. A. Sychev (1999). "A new approach to Young measure theory, relaxation and convergence in energy". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 16.6, pp. 773–812. MR: 1720517 (cit. on p. 2241).

László Székelyhidi and Emil Wiedemann (2012). "Young measures generated by ideal incompressible fluid flows". *Arch. Ration. Mech. Anal.* 206.1, pp. 333–366. MR: 2968597 (cit. on p. 2234).

L. Tartar (1979). "Compensated compactness and applications to partial differential equations". In: *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*. Vol. 39. Res. Notes in Math. Pitman, Boston, Mass.-London, pp. 136–212. MR: 584398 (cit. on pp. 2234, 2241).

Luc Tartar (1983). "The compensated compactness method applied to systems of conservation laws". In: *Systems of nonlinear partial differential equations (Oxford, 1982)*. Vol. 111. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Reidel, Dordrecht, pp. 263–285. MR: 725524 (cit. on pp. 2234, 2241).

Roger Temam (1983). *Problèmes mathématiques en plasticité*. Vol. 12. Méthodes Mathématiques de l'Informatique. Gauthier-Villars, Montrouge, pp. vii+353. MR: 711964 (cit. on p. 2238).

Roger Temam and Gilbert Strang (1980/81). "Functions of bounded deformation". *Arch. Rational Mech. Anal.* 75.1, pp. 7–21. MR: 592100 (cit. on pp. 2237, 2238, 2243).

L. C. Young (1937). "Generalized curves and the existence of an attained absolute minimum in the calculus of variations". *C. R. Soc. Sci. Lett. Varsovie, Cl. III* 30, pp. 212–234 (cit. on p. 2241).

– (1942a). "Generalized surfaces in the calculus of variations". *Ann. of Math. (2)* 43, pp. 84–103. MR: 0006023 (cit. on pp. 2241, 2244).

– (1942b). "Generalized surfaces in the calculus of variations. II". *Ann. of Math. (2)* 43, pp. 530–544. MR: 0006832 (cit. on pp. 2241, 2244).

– (1980). *Lectures on the Calculus of Variations and Optimal Control Theory*. 2nd ed. Chelsea (cit. on p. 2244).

Zygmunt Zahorski (1946). "Sur l'ensemble des points de non-dérivabilité d'une fonction continue". *Bull. Soc. Math. France* 74, pp. 147–178. MR: 0022592 (cit. on p. 2245).

Received 2017-11-29.

Guido De Philippis
SISSA, Via Bonomea 265, 34136 Trieste, Italy
guido.dephilippis@sissa.it

Filip Rindler

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
F.Rindler@warwick.ac.uk

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2259–2278)

LONG TIME EXISTENCE RESULTS FOR SOLUTIONS OF WATER WAVES EQUATIONS

JEAN-MARC DELORT

Abstract

We present in this talk various results, obtained during the last years by several authors, about the problem of long time existence of solutions of water waves and related equations, with initial data that are small, smooth, and decaying at infinity. After recalling some facts about local existence theory, we shall focus mainly on global existence theorems for gravity waves equations proved by Ionescu–Pusateri, Alazard–Delort and Ifrim–Tataru. We shall describe some of the ideas of the proofs of these theorems, and mention as well related results.

1 The water waves equations

Consider an incompressible and irrotational fluid, of constant density equal to one, in a vertical gravity field of intensity g. Assume that at time t, the domain occupied by the fluid is

$$\Omega_t = \{(x,z) \in \mathbb{R}^d \times \mathbb{R}; -H_0 < z < \eta(t,x)\},$$

where $\eta(t,\cdot) : \mathbb{R}^d \to \mathbb{R}$ is such that $\inf_{x\in\mathbb{R}^d} \eta(t,x) > -H_0$, and either $H_0 \in]0,+\infty[$ (for a fluid of finite depth) or $H_0 = +\infty$ (for an infinite depth fluid).

The velocity U in the fluid solves in $\Omega = \{(t,x); x \in \Omega_t\}$ the incompressible Euler equations

$$\begin{aligned} \partial_t U + U \cdot \nabla_{x,z} U &= -\nabla_{x,z} p - g e_z \\ \operatorname{div} U &= 0 \end{aligned} \tag{1}$$

where e_z is the vertical unit vector and p the pressure inside the fluid. Moreover, the normal velocity at the bottom satisfies $U \cdot e_z|_{z=-H_0} = 0$ (when $H_0 < +\infty$) or $U \to 0$

From joint work with Thomas Alazard. Partially supported by the ANR project 13-BS01-0010-02 “Analyse asymptotique des équations aux dérivées partielles d’évolution”.
MSC2010: primary 76B15; secondary 35Q35, 35B40, 35S50.

when z goes to $-\infty$ (if $H_0 = -\infty$). Finally, the free surface is driven by the velocity of the fluid at each point of the interface $z = \eta(t,x)$, which is translated by

$$\partial_t \eta(t,x) = \sqrt{1 + |\nabla_x \eta(t,x)|^2}\, U(t,x,z) \cdot n|_{z=\eta(t,x)}, \tag{2}$$

where n is the unit outward normal vector to Ω_t at $(x, \eta(t,x))$. Moreover, the pressure above the fluid is equal to the constant atmospheric pressure, that we may take equal to zero. At the interface $z = \eta(t,x)$, the pressure of the fluid will be given by

$$p|_{z=\eta(t,x)} = -\kappa \operatorname{div}\left(\frac{\nabla_x \eta}{\sqrt{1 + |\nabla_x \eta|^2}}\right), \tag{3}$$

where the constant $\kappa \geq 0$ is the surface tension. Since, as the fluid is also assumed to be irrotational, $\operatorname{curl} U = 0$, one may express the velocity U from a potential Φ by $U = \nabla_{(x,z)}\Phi$. The incompressibility implies $\Delta_{(x,z)}\Phi = 0$ and the Euler equation (1) allows one to write an equation for Φ:

$$\partial_t \Phi + \frac{1}{2}|\nabla_{(x,z)}\Phi|^2 + gz = -p. \tag{4}$$

Moreover, one has the boundary condition at the bottom

$$\begin{aligned} &\partial_z \Phi|_{z=-H_0} = 0 \text{ (for finite depth)} \\ &\nabla_{(x,z)}\Phi \to 0 \text{ if } z \to -\infty \text{ (for infinite depth)} \end{aligned} \tag{5}$$

and, expressing $U = \nabla_{(x,z)}\Phi$ in (2), one gets

$$\partial_t \eta(t,x) = \sqrt{1 + |\nabla_x \eta(t,x)|^2}\, \partial_n \Phi(t,x,z) \text{ on } z = \eta(t,x), \tag{6}$$

denoting by ∂_n the outwards normal unit derivative at the free interface. The Craig-Sulem-Zakharov formulation of the water waves system, given in Zakharov [1968] and Craig and Sulem [1993] (see also the book of Lannes [2013b]) is obtained expressing in (4), (6), the potential Φ from its boundary data. More precisely, denote by ψ the restriction of Φ to the interface $z = \eta(t,x)$. Then Φ solves the elliptic boundary values problem

$$\begin{aligned} \Delta_{(x,z)}\Phi &= 0 \\ \Phi|_{z=\eta(t,x)} &= \psi \\ \partial_z \Phi|_{z=-H_0} &= 0 \end{aligned} \tag{7}$$

(or, for the last condition, $\nabla_{(x,z)}\Phi \to 0$ when $z \to -\infty$ in the case of infinite depth). One denotes by $G(\eta)\psi$ the Dirichlet-Neuman operator defined by

$$G(\eta)\psi = \sqrt{1 + |\nabla_x \eta|^2}\, \partial_n \Phi|_{z=\eta(t,x)}, \tag{8}$$

where Φ solves (7). Plugging this information in (4) restricted to $z = \eta(t,x)$, (6), and using (3), one obtains for (η, ψ) the system

$$
\begin{aligned}
&\partial_t \eta = G(\eta)\psi \\
&\partial_t \psi = -g\eta - \frac{1}{2}|\nabla_x \psi|^2 + \frac{\big(G(\eta)\psi + \nabla_x \eta \cdot \nabla_x \psi\big)^2}{2(1+|\nabla_x \eta|^2)} + \kappa \operatorname{div}\left(\frac{\nabla_x \eta}{\sqrt{1+|\nabla_x \eta|^2}}\right).
\end{aligned} \tag{9}
$$

This is the system we intend to study below, in the case of pure gravity water waves, i.e. when $g > 0$ and $\kappa = 0$.

2 The question of local existence

The question of existence of local in time solutions for system (9) (when $\kappa = 0$ and the fluid depth is infinite) with data in Sobolev spaces remained open for a long time, and was fully answered in 1997 by Sijue Wu in the seminal paper Wu [1997] when x belongs to $\mathbb{R}$ and in Wu [1999] when x is in $\mathbb{R}^2$. As in the subsequent sections we shall be interested mainly in the one dimensional problem, we assume for the rest of this section that the space variable x belongs to $\mathbb{R}$. The difficulty in order to prove local existence may be seen in the following way: if one writes (9) under the form

$$
\partial_t \begin{bmatrix} \eta \\ \psi \end{bmatrix} = A(\eta, \psi) \begin{bmatrix} \eta \\ \psi \end{bmatrix}, \tag{10}
$$

where A is a pseudo-differential operator with coefficients with limited smoothness, defined by

$$
A(\eta, \psi)\begin{bmatrix} \dot\eta \\ \dot\psi \end{bmatrix} = \frac{1}{2\pi} \int e^{ix\xi} M(x,\xi) \begin{bmatrix} \hat{\dot\eta}(t,\xi) \\ \hat{\dot\psi}(t,\xi) \end{bmatrix} d\xi, \tag{11}
$$

then the matrix symbol $M(x,\xi)$ (that depends on (η, ψ)) has eigenvalues whose real part may go to infinity with $|\xi|$. This instability prevents one from getting energy inequalities.

A way to overcome this difficulty, and to prove local existence for a restricted class of energy data, has been introduced by Nalimov [1974] (for infinite depth fluids) and Yosihara [1982] for finite depth ones. See also the work of Craig [1985]. The local existence of solutions for *arbitrary* Sobolev initial data has been established for infinite depth fluids by Wu [1997, 1999]. Actually, her work is not limited to an interface given by a graph $z = \eta(t,x)$, but allows upper boundaries for Ω_t given by any non self-intersecting smooth curve. The method used by the above authors was relying on the Lagrangian formulation of the water waves system, and has been at the origin of a lot of works concerning related models (like for instance the capillary-gravity wave equations, where κ in (9) is non zero) with finite or infinite depth, both for localized or unlocalized initial data. We cite in particular results of local existence of Ambrose [2003], Ambrose and Masmoudi [2005], Coutand and Shkoller [2007]. At the same time, a more geometric approach to study free boundary value problems has been developed by Christodoulou and Lindblad [2000], Lindblad [2005] and in a series of papers of Shatah and Zeng [2008b,a, 2011].

On the other hand, Lannes [2005] introduced an Eulerian approach to local existence, expressing the problem in terms of a "good unknown" ω instead of ψ. Such a "good unknown" had been introduced in the framework of free boundary problems by Alinhac [1986, 1988]. For water waves equations (in any dimension, and with eventually a bottom), Lannes showed that the system, written in terms of (η, ω), is a quasi-linear hyperbolic equation, for which Sobolev energy estimates are available and provide local existence of solutions.

This new unknown was later implemented by Alazard and Métivier [2009] in a paradifferential framework. Let us describe how it may be defined for problem (9) in one space dimension. Recall that the paraproduct $T_a b$ Bony [1981] of a bounded function a and a tempered distribution b may be defined by

$$\widehat{T_a b}(\xi) = \int_{\xi_1+\xi_2=\xi} \chi(\xi_1, \xi_2)\hat{a}(\xi_1)\hat{b}(\xi_2)\, d\xi_1 d\xi_2,$$

where χ is a smooth function satisfying

$$|\partial_{\xi_1}^{\alpha_1}\partial_{\xi_2}^{\alpha_2}\chi(\xi_1, \xi_2)| \leq C(1 + |\xi_1| + |\xi_2|)^{-\alpha_1-\alpha_2},$$

supported for $|\xi_1| \leq (1 + |\xi_2|)/10$, equal to one for $|\xi_1| \leq (1 + |\xi_2|)/100$ for instance. Then if a is in L^∞ and b is in a Sobolev space H^s, the paraproduct $T_a b$ belongs to H^s, whatever the value of s. Let us introduce:

Definition 1. *For η in $H^s(\mathbb{R})$, ψ such that $|D_x|^{\frac{1}{2}}\psi \in H^s(\mathbb{R})$, with s large enough, set*

$$B(\eta)\psi \overset{def}{=} \frac{G(\eta)\psi + (\partial_x\eta)(\partial_x\psi)}{1 + (\partial_x\eta)^2}. \tag{12}$$

One defines the good unknown ω by

$$\omega = \psi - T_{B(\eta)\psi}\eta. \tag{13}$$

Using this good unknown, and a paralinearization of the Dirichlet-Neuman operator due to Alazard and Métivier [2009], Alazard, Burq, and Zuily [2014, 2016, 2011a] and Alazard, Burq, and Zuily [2011b] proved local existence theorems for (9) (with or without surface tension) under weaker regularity assumptions on the Cauchy data than in previous works.

Finally, let us mention a last approach to local existence theory, encompassing in some way the Lagrangian formulation and the use of a good unknown, proposed by Hunter, Ifrim, and Tataru [2016], that relies on the introduction of new quantities defined as boundary values of holomorphic functions.

3 Global existence with small decaying data

Once local existence of solutions to system (9) is established, it is natural to ask the question of long time existence for small smooth enough initial data that have some decay at infinity. Up to the last but one section, we discuss this problem for (9) in infinite depth, when the surface tension κ is equal to zero, and space dimension d is equal to one or two, postponing to the last section references to other models (finite depth, presence of surface tension terms, etc).

One checks easily that the solution of (9) with $\kappa = 0$ linearized on the zero solution, with decaying initial data, has L^∞ norm that is $O(t^{-\frac{d}{2}})$ when t goes to infinity, in d space dimension, because of the dispersive effect. The first breakthrough concerning long time existence of solutions is due to Wu [2009] who proved that, in one space dimension, for smooth decaying Cauchy data of small size ϵ, the solution exists over a time interval of length $e^{c/\epsilon}$ for some positive constant c. For *two space dimensions*, the stronger decay rate of solutions of the linearized equation makes expect better results. Actually Wu [2011] and Germain, Masmoudi, and Shatah [2012b] proved that then solutions are global if the data are smooth, small, and decaying enough. Moreover, there is scattering Germain, Masmoudi, and Shatah [ibid.], i.e. the solutions of the nonlinear problem have the same asymptotics as solutions of the linearized equation on the zero state when time goes to infinity.

The main result we report on here concerns, *in one space dimension*, global existence of solutions for small, smooth, decaying Cauchy data, and modified scattering. This result has been obtained independently by Ionescu and Pusateri [2015b], using a combination of the Lagrangian and the Eulerian formulations of the equations, and by Alazard and Delort [2015a,b], through the Eulerian formulation and the good unknown introduced above. An alternative approach, based on the "holomorphic coordinates" of Hunter, Ifrim and Tataru, has also been proposed by Ifrim and Tataru [2016].

We state below the result of Alazard and Delort [2015a]. We compare it next with the statements of Ionescu and Pusateri [2015b] and Ifrim and Tataru [2016].

Theorem 2. *Fix $\gamma \in \mathbb{R} - \frac{1}{2}\mathbb{N}$ a large enough number, s, s_1 in $\mathbb{N}$ such that $s_1 \geq \frac{s}{2} + \gamma$ and $s - s_1$ is large enough. There is $\epsilon_0 > 0$ and for any $\epsilon \in]0, \epsilon_0[$, any couple (η_0, ψ_0) of* *functions satisfying for any integer* $0 \le p \le s_1$,

$$(14) \qquad \begin{aligned} &(x\partial_x)^p \eta_0 \in H^{s-p}(\mathbb{R}), \quad |D_x|^{\frac{1}{2}}(x\partial_x)^p \psi_0 \in H^{s-p-\frac{1}{2}}(\mathbb{R}) \\ &\qquad |D_x|^{\frac{1}{2}}(x\partial_x)^p\big[\psi_0 - T_{B(\eta_0)\psi_0}\eta_0\big] \in H^{s-p}(\mathbb{R}), \end{aligned}$$

with the norms in the above spaces smaller or equal to one, the evolution problem (9) (in one space dimension, with $\kappa = 0$ *and infinite depth) with initial data* $\psi|_{t=1} = \epsilon\psi_0$, $\eta|_{t=1} = \epsilon\eta_0$, *has a unique solution, continuous with values in the set of functions satisfying (14), defined on the whole interval* $[1, +\infty[$. *Moreover, if we define* $u = |D_x|^{\frac{1}{2}}\psi + i\eta$, *we have the following asymptotics when t goes to infinity:*

$$(15) \qquad u(t,x) = \frac{\epsilon}{\sqrt{t}}\alpha_\epsilon\Big(\frac{x}{t}\Big)\exp\bigg[\frac{it}{4|x/t|} + \frac{i\epsilon^2}{64}\frac{|\alpha_\epsilon(x/t)|^2}{|x/t|^5}\log t\bigg] + \epsilon t^{-\frac{1}{2}-\theta}\rho(t,x),$$

where $(\alpha_\epsilon)_{\epsilon\in]0,1]}$ *is a bounded family of functions of* $C(\mathbb{R}) \cap L^\infty(\mathbb{R})$, θ *is positive and* $\|\rho(t,\cdot)\|_{L^\infty}$ *is bounded for* $t \ge 1$.

Remark: The asymptotics (15) show that the global solution displays a modified scattering, where the phase of oscillation of linear solutions is modulated by an extra logarithmic term. This term, that becomes significant for times t such that $\log t \gg \epsilon^{-2}$, is responsible of the new difficulties arising in one dimensional problems versus two dimensional ones.

Let us compare the above result with the ones of Ionescu and Pusateri [2015b] and Ifrim and Tataru [2016].

In the first of these references, a similar result of global existence is obtained under a weaker decay condition on the initial data, namely conditions of the form (14) have to be imposed only for $0 \le p \le 1$. Moreover, one assumes the smallness of a "Z-norm", that controls $\||\xi|^\beta \hat{u}_0(\xi)\mathbb{1}_{|\xi|\le 1}\|_{L^\infty}$, for some $\beta > 0$, where $u_0 = |D_x|^{\frac{1}{2}}\psi_0 + i\eta_0$. This norm plays a key role in the proof of the optimal decay of the solution and of the modified scattering.

The result of Ifrim and Tataru is expressed from slightly different unknowns than (η, ψ) above. Actually, these authors introduce Z, a parameterization of the boundary, chosen in such a way that it is the boundary value of some holomorphic function in the fluid domain, and Q, the boundary value of another holomorphic function in the fluid, whose real part coincides with the velocity potential at the boundary. They assume that a Sobolev norm of $\big(Z(t,\alpha) - \alpha\big)|_{t=0}$ and $Q(t,\alpha)|_{t=0}$, involving essentially at most six derivatives, is small and that an H^1 norm of the action of $x\partial_x$ on an expression of these quantities is also small at the initial time. They obtain then global existence and modified scattering.

The proofs of the results of global existence of Alazard and Delort [2015a], Ifrim and Tataru [2016], and Ionescu and Pusateri [2015b] might differ in their technical details, but the difficulties that have to be overcome are essentially the same. In the rest of that report, we shall try to describe them in a non technical way, using most of the time some simplified toy models instead of the full equation (9). We shall use the formulation of the equation in terms of η and the good unknown ω, following Alazard and Delort [2015a,b], but will make also frequent references to the works Ionescu and Pusateri [2015b] and Ifrim and Tataru [2016].

4 Quadratic terms. Normal forms

Consider first a model equation of the form

$$(16)\qquad \begin{aligned}\big(D_t - p(D_x)\big)u &= N(u)\\ u|_{t=1} &= \epsilon u_0\end{aligned}$$

where $D_t = \frac{1}{i}\frac{\partial}{\partial t}$, $D_x = \frac{1}{i}\frac{\partial}{\partial x}$, $p(\xi)$ is a real valued elliptic Fourier multiplier and $N(u)$ is a nonlinearity vanishing at least at order two at zero. The water wave system, linearized at the zero state, may be expressed in terms of $u = |D_x|^{\frac{1}{2}}\psi + i\eta$, as $(D_t - |D_x|^{\frac{1}{2}})u = 0$, so that (16) is a model for that system if we take $p(\xi) = |\xi|^{\frac{1}{2}}$.

Assume first that N is semi-linear and at least cubic at zero, in that sense that it satisfies an estimate

$$(17)\qquad \|N(u)\|_{H^s} \leq C\|u\|_{L^\infty}^2\|u\|_{H^s}$$

for any $s > 0$ for instance. Then the Sobolev energy inequality associated to (16) writes

$$(18)\qquad \|u(t,\cdot)\|_{H^s} \leq \|u(1,\cdot)\|_{H^s} + C\int_1^t \|u(\tau,\cdot)\|_{L^\infty}^2\|u(\tau,\cdot)\|_{H^s}\,d\tau.$$

If one assumes that, in addition, one has some a priori estimate for $\|u(t,\cdot)\|_{L^\infty} \leq B\epsilon t^{-\frac{1}{2}}$ (as the one we expect according to (15)), we deduce by Gronwall inequality a bound

$$(19)\qquad \|u(t,\cdot)\|_{H^s} \leq C\|u(0,\cdot)\|_{H^s} t^{CB^2\epsilon^2},$$

i.e. a control of the Sobolev norm that is not uniform, but given in terms of an arbitrary small power of t (if ϵ is small). As we shall see, a bound of this type will be sufficient for our goals.

On the other hand, in the system (9) we are really interested in, the nonlinearity is *quadratic*, and not cubic. For the toy model (16), this would mean assuming, instead of (17), $\|N(u)\|_{H^s} \leq C\|u\|_{L^\infty}\|u\|_{H^s}$, so that (19) would be replaced by

$$\|u(t,\cdot)\|_{H^s} \leq C\|u(0,\cdot)\|_{H^s}\exp[CB\epsilon\sqrt{t}]$$

which is useless if one wants to study solutions on time intervals of length larger than ϵ^{-2}. The way to actually obtain estimates of the form (19), including for a quadratic nonlinearity, is well known in the semi-linear case: this is the normal forms method of Shatah [1985], that allows to reduce a quadratic nonlinearity to a cubic one. (We refer also to the more recent developments of that method introduced in the work of Germain, Masmoudi, and Shatah [2012a]. See also the Bourbaki seminar of Lannes [2013a] and references therein.) For quadratic nonlinearities, $N(u) = u^2$ for instance, the idea of the method is to look for a new unknown $\phi = u + E(u,u)$, where $E(u,u)$ is a quadratic expression of the form

$$E(u,u) = \frac{1}{2\pi} \int e^{ix(\xi_1+\xi_2)} m(\xi_1,\xi_2)\hat{u}(\xi_1)\hat{u}(\xi_2)\, d\xi_1 d\xi_2, \tag{20}$$

chosen in such a way that

$$(D_t - p(D_x))\phi = \tilde{N}(\phi) \tag{21}$$

for a new nonlinearity $\tilde{N}$ vanishing at *third* order at zero. A direct computation using (16) shows that one has to take

$$m(\xi_1,\xi_2) = \big(p(\xi_1) + p(\xi_2) - p(\xi_1+\xi_2)\big)^{-1} \tag{22}$$

in order to achieve that. The transformation $u \to \phi$ will then be bounded on H^s spaces if, for large frequencies ξ_1, ξ_2, one has a bound of the form

$$|m(\xi_1,\xi_2)| \leq C \min(|\xi_1|, |\xi_2|)^{N_0} \tag{23}$$

for some fixed N_0. In that way, if s is large enough relatively to N_0, $\|E(u,u)\|_{H^s} \leq C\|u\|_{H^s}^2$, which shows that $u \to \phi$ is a diffeomorphism from a neighborhood of zero to its image. For the toy model (16) with $p(\xi) = |\xi|^{\frac{1}{2}}$, it is easy to see that an estimate of the form (23) holds for large frequencies. For small ones, a degeneracy happens, but, in the case of the water waves system (9), it will be compensated by the fact that in the nonlinearity, operators whose symbols vanish at $\xi = 0$ act on u.

In the case of the water waves system (9), one would like to perform as well a similar normal forms method in order to eliminate quadratic terms. The difficulty is that, (9) being quasi-linear, (16) is not a convincing model for it, as the nonlinearity there depends only on u, and not on first order derivatives of u. On the other hand, if one replaces $N(u)$ by a quadratic term of the form uD_xu, and tries to eliminate it looking for a new unknown $\phi = u + E(u,u)$, one would have to express E by (20), but with a symbol $m(\xi_1,\xi_2)$ given by

$$m(\xi_1,\xi_2) = \big(p(\xi_1) + p(\xi_2) - p(\xi_1+\xi_2)\big)^{-1}\xi_2$$

which loses one derivative, so that $\|\phi\|_{H^s}$ is only estimated from $\|u\|_{H^{s+1}}$. A way to circumvent that difficulty, that appears already in the work of Ozawa, Tsutaya, and Tsutsumi [1997], is to try to combine the normal forms construction with the idea used to get *quasi-linear* energy inequalities. In order to explain that, we have to make appeal to a more accurate model than (16). We have seen in Section 2 that, in order to avoid losses of derivatives in energy estimates, it is convenient to write the water waves equation in terms of the unknowns $(\eta, \omega = \psi - T_{B(\eta)\psi}\eta)$. More precisely, if one introduces $U = \begin{bmatrix} \eta + T_\alpha \eta \\ |D_x|^{\frac{1}{2}}\omega \end{bmatrix}$, where α is some explicit function of $u = |D_x|^{\frac{1}{2}}\psi + i\eta$, vanishing at $u = 0$, one may write system (9) under the form

$$\partial_t U + \begin{bmatrix} 0 & -|D_x|^{\frac{1}{2}} \\ |D_x|^{\frac{1}{2}} & 0 \end{bmatrix} U + Q(u)U + S(u)U + C(u)U = G \tag{24}$$

where we used the following notation:

• The right hand side G is a semi-linear cubic term. This means that it will satisfy for $s \gg \rho \gg 1$ estimates of the following form:

$$\|G\|_{H^s} \le C(\|u\|_{C^\rho})\|u\|_{C^\rho}^2\|U\|_{H^s}. \tag{25}$$

Such a term satisfies thus bounds of the form (17) (with the L^∞ norm replaced by a Hölder norm). If we write the Sobolev energy inequality associated to (24), and forget the contributions of Q, S, C, we would thus get

$$\|U(t,\cdot)\|_{H^s} \le \|U(t,\cdot)\|_{H^s} + C\int_1^t \|u(\tau,\cdot)\|_{C^\rho}^2\|U(\tau,\cdot)\|_{H^s}\,d\tau. \tag{26}$$

Combined with an a priori bound $\|u(\tau,\cdot)\|_{C^\rho} = O(\epsilon t^{-\frac{1}{2}})$, this would give for $\|U(t,\cdot)\|_{H^s}$ an estimate of the form (19).

• The term $C(u)U$ is a cubic contribution given in terms of a paradifferential operator of order one $C(u)$: this means that

$$C(u)v = \frac{1}{2\pi}\int e^{ix\xi}c(u;x,\xi)\hat{v}(\xi)\,d\xi \tag{27}$$

where $u \to c(u;x,\xi)$ is a quadratic map with values in the set of functions of (x,ξ) satisfying bounds of the form

$$|\partial_\xi^\beta c(u;x,\xi)| \le C(\|u\|_{C^\rho})\|u\|_{C^\rho}^2\langle\xi\rangle^{1-\beta} \tag{28}$$

for some fixed $\rho > 0$ independent of β, and that moreover the Fourier transform with respect to x of $x \to c(u;x,\xi)$, that we denote by $\hat{c}(u;\eta,\xi)$, is supported for $|\eta| \ll |\xi|$. The paraproduct recalled in Section 3 is a special case of the above definition, corresponding to the case when $c(u;x,\xi)$ does not depend on ξ. When one computes the time derivative

$$\partial_t \|U(t,\cdot)\|_{L^2}^2 = 2\Re\langle \partial_t U, U\rangle, \tag{29}$$

the contribution to the right hand side coming from the $C(u)U$ term in (24) may be written as

$$\langle (C(u)+C(u)^*)U, U\rangle. \tag{30}$$

Because of the explicit form of the operator $C(u)$, one may check that $C(u)+C(u)^*$ is of order zero. (This reflects the fact that the equation (24) one reduced to using the good unknown is hyperbolic). Taking into account that operators of order zero are bounded on L^2, and that $u \to C(u)$ vanishes at order two at zero, one gets for (30) an upper bound in $\|u\|_{C^\rho}^2\|U\|_{L^2}$. Since the same reasoning may be done replacing in (29) L^2 norms by Sobolev ones, we see that the term $C(u)U$ in (24) would also generate in an energy inequality a contribution bounded from above by the right hand side of (26).

• The terms $Q(u)U$ and $S(u)U$ are quadratic contributions, with $Q(u)$ a paradifferential operator of order one, linear in u, and $S(u)$ a smoothing operator. These contributions have to be eliminated by normal forms. As $S(u)U$ is a semi-linear term, it may be eliminated essentially using a quadratic correction of the form (20), up to some technical issues that we do not discuss here. On the other hand, $Q(u)U$ is a quasi-linear contribution, and as we have seen above, a brutal elimination could give rise to an unbounded normal forms transformation. But again, as in (30), only the operator $Q(u)+Q(u)^*$ plays a role in an energy inequality, and by the hyperbolic structure of equation (24), such an operator is of order zero instead of one. Consequently, trying to eliminate only this term from the right hand side of (24), one may construct, including in this quasi-linear framework, a change of unknown $U \to \phi = U + E(u)U$, where E is *bounded* on H^s, and such that ϕ, and thus U, will obey an estimate of the form (26). One has thus reduced morally to a cubic nonlinearity.

We may summarize this in the following statement:

Proposition 3. *There is a bounded linear map $U \to \phi = U + E(u)U$, going from a neighborhood of zero in H^s to a neighborhood of zero in H^s, when u is in a ball of C^ρ and $1 \ll \rho \ll s$, that transforms equation (24) for U into*

$$\partial_t\phi + \begin{bmatrix} 0 & -|D_x|^{\frac12} \\ |D_x|^{\frac12} & 0\end{bmatrix}\phi + L(u)\phi + C(u)\phi = \Gamma, \tag{31}$$

where $C(u)$ is as in (24), Γ is a cubic semi-linear term, and $L(u)$ is linear in u and satisfies $\Re\langle L(u)\phi,\phi\rangle_{H^s} = 0$.

The normal forms method outlined above does not eliminate the whole quadratic part of the nonlinearity, but only those terms in it that give nonzero contributions to the energy. Because of that, one might think that it should be possible to perform the normal form procedure on the Sobolev energy itself instead of the equation. Such an approach has been performed by Hunter, Ifrim, Tataru, and Wong [2015] and used by Hunter, Ifrim, and Tataru [2016] in order to give another proof of the almost global existence result of Wu [2009]. They applied next the same method, that does not require the paralinearization of the equation, in their proof of the global existence result Ifrim and Tataru [2016]. Notice that similar ideas, (combined with an a priori paralinearization) are used in Delort [2009, 2012, 2015] for quasi-linear Klein-Gordon equations on some compact manifolds. Such an approach is particularly convenient when one studies a Hamiltonian system and wants to keep track of the Hamiltonian structure all along a normal forms reduction procedure.

On the other hand, the elimination of the contributions of the quadratic part of the nonlinearity to the energy, as a first step towards the proof of a global existence result, is performed by Ionescu and Pusateri [2015b] using the transformation in Lagrangian coordinates of Wu [2009], Totz and Wu [2012].

5 Global existence: bootstrap procedure

We shall discuss from now on the proof of the global existence result of Theorem 2 on a model equation. If $p(\xi)$ is some elliptic Fourier multiplier, consider

(32) $$(D_t - p(D_x))u = N(u)$$

with $N(u)$ a cubic nonlinearity of the form

(33) $$N(u) = \alpha_3 u^3 + \alpha_1 |u|^2 u + \alpha_{-1}|u|^2 \bar{u} + \alpha_{-3}\bar{u}^3,$$

where α_j are complex numbers, with α_1 real. Of course, equation (32) is a simplification of the real system we are studying, but it is a convincing prototype of the problem after the normal forms procedure of the preceding section has been performed in order to reduce to a cubic nonlinearity. The fact that it is semi-linear instead of quasilinear just brings some inessential technical simplifications at this level. Let us introduce the Klainerman vector field

(34) $$Z = tD_t + 2xD_x$$

that satisfies when $p(\xi) = |\xi|^{\frac{1}{2}}$

$$[D_t - p(D_x), Z] = D_t - p(D_x)$$

so that Zu solves

$$(D_t - p(D_x))(Zu) = N(u) + ZN(u). \tag{35}$$

The key of global existence (and modified scattering) is the proof of the following bootstrap assertion:

Proposition 4. *There are positive constants A, B, K, s_0 and $\epsilon_0 \in]0,1]$ such that, for any $s \geq s_0$, any $\epsilon \in]0, \epsilon_0[$, any u_0 in $H^s(\mathbb{R})$ satisfying*

$$\|u_0\|_{H^s} + \|x\partial_x u_0\|_{L^2} \leq \epsilon,$$

for any solution u of (32) with initial u_0 that exists over some interval $I = [1, T]$, and satisfies for any t in I,

$$\begin{aligned} &(A) \qquad \|u(t,\cdot)\|_{H^s} + \|Zu(t,\cdot)\|_{L^2} \leq A\epsilon t^{K\epsilon^2} \\ &(B) \qquad \|u(t,\cdot)\|_{L^\infty} \leq \frac{B\epsilon}{\sqrt{t}}, \end{aligned} \tag{36}$$

then, for t in the same interval I, one has actually

$$\begin{aligned} &(A') \qquad \|u(t,\cdot)\|_{H^s} + \|Zu(t,\cdot)\|_{L^2} \leq \frac{A}{2}\epsilon t^{K\epsilon^2} \\ &(B') \qquad \|u(t,\cdot)\|_{L^\infty} \leq \frac{B\epsilon}{2\sqrt{t}}. \end{aligned} \tag{37}$$

Remarks: • In the water waves system we are interested in, the quasi-linear nature of the problem makes that one has to control some derivatives of Zu in L^2 and of u in L^∞, i.e. one has to replace in (A), (A'), $\|Zu(t,\cdot)\|_{L^2}$ by $\|Zu(t,\cdot)\|_{H^\sigma}$ for some σ satisfying $1 \ll \sigma \ll s$, and in (B), (B'), $\|u(t,\cdot)\|_{L^\infty}$ by $\|u(t,\cdot)\|_{C^\rho}$ for some ρ with $1 \ll \rho \ll s$. This does not bring any essential new difficulty.

• In the statement of Theorem 2, we were assuming that the initial data admitted the action of a large number of iterates of $(x\partial_x)$, which would correspond in the model (32) above to make act a large number of vector fields Z. This was due to the fact that in Alazard and Delort [2015a] some non optimal choice was made in the proof of L^∞ estimates. On the other hand, in the work of Ionescu and Pusateri [2015b] and of Ifrim and Tataru [2016], only one vector field has to be used. Below, inspired by Ifrim and Tataru [2016, 2015a], we shall adopt an optimal point of view that allows one to use only one vector field in the analysis of model (32), following the method of Alazard and Delort [2015a].

Proposition 4 implies immediately the global existence result in Theorem 2 when combined with local existence theory.

The fact that (A) and (B) imply (A') is essentially trivial for the model equation (32). Actually, writing the energy inequality for (32) and (35), one gets

$$\|u(t,\cdot)\|_{H^s} + \|Zu(t,\cdot)\|_{L^2} \leq C\Big[\|u(1,\cdot)\|_{H^s} + \|Zu(1,\cdot)\|_{L^2} + \int_1^t \big[\|N(u(\tau,\cdot))\|_{H^s} + \|ZN(u(\tau,\cdot))\|_{L^2}\big]\,d\tau\Big]. \tag{38}$$

As N is cubic in $(u,\bar{u})$, the right hand side of (38) is bounded from above by

$$C\Big[\|u(1,\cdot)\|_{H^s} + \|Zu(1,\cdot)\|_{L^2} + \int_1^t \|u(\tau,\cdot)\|_{L^\infty}^2\big[\|u(\tau,\cdot)\|_{H^s} + \|Zu(\tau,\cdot)\|_{L^2}\big]\,d\tau\Big]. \tag{39}$$

Plugging (36) into (39), choosing A large enough and then ϵ_0 small enough in function of A, B, one deduces estimate (A').

Of course, in the case of system (9) (with $d = 1$, infinite depth and $\kappa = 0$), the proof of the corresponding inequality is much more technical, as one has to cope with the difficulties explained in Sections 2 and 4 in the case of Sobolev energy inequalities. Estimates in L^2 for Zu instead of u are performed in a similar way, using the new unknown ω and a normal form in order to get rid of the quadratic part of the nonlinearity.

The remaining step, in order to complete the proof of Theorem 2, is to show that (B') holds for solutions of the equation under assumptions (A) and (B).

6 Optimal decay estimates

The key point in order to prove the enhanced decay estimate (B') from (A) and (B), both in the case of the simplified model (32) or for the true water waves equation, is to derive from the PDE an ODE whose analysis will provide the wanted L^∞ bounds, as well as the asymptotics of the solution.

Several approaches have been used by different authors in order to do so. Ionescu and Pusateri [2015b] work in Fourier space in order to get an ODE for the Fourier transform of the solution. Ifrim and Tataru [2016] use a wave packets description of the solution, for which they obtain an ODE, working thus in phase-space variables. The approach in Alazard and Delort [2015a] relies on the rewriting of the PDE under study in a semi-classical framework, with a Planck constant $h = \frac{1}{t}$, so that the ODE one looks for is obtained as the semi-classical limit of the quantum problem provided by the PDE. This is the method we present below, blending the approach of Alazard and Delort [ibid.] (which was not optimal regarding to phase-space decomposition) with some of the ideas of Ifrim and Tataru [2016, 2015a]. Let us introduce:

Definition 5. *Let $\delta \in [0, \frac{1}{2}]$, $m \in \mathbb{R}$. We denote by $S^m_\delta(\mathbb{R} \times \mathbb{R})$ the space of smooth functions $(h, x, \xi) \to a_h(x, \xi)$, defined for h in $]0, 1]$, (x, ξ) in $\mathbb{R} \times \mathbb{R}$, satisfying estimates*

$$|(h\partial_h)^\gamma \partial_x^\alpha \partial_\xi^\beta a_h(x, \xi)| \le C h^{-\delta(\alpha+\beta)} \langle \xi \rangle^m. \tag{40}$$

If a is in $S^m_\delta(\mathbb{R} \times \mathbb{R})$, we define its Weyl-quantization by the formula

$$\mathrm{Op}_h^{\mathrm{W}}(a)v = \frac{1}{2\pi h} \int e^{i(x-y)\frac{\xi}{h}} a\Big(\frac{x+y}{2}, \xi\Big) v(y)\, dy d\xi \tag{41}$$

for v in $\mathcal{S}(\mathbb{R})$. We denote by $H^s_{\mathrm{sc}}(\mathbb{R})$ the space of families of functions $v = (v_h)_{h \in]0,1]}$ such that, if we define

$$\|v_h\|_{H^s_h} = \|\mathrm{Op}_h^{\mathrm{W}}(\langle \xi \rangle^s) v_h\|_{L^2} = \|\langle h D_x \rangle^s v_h\|_{L^2}, \tag{42}$$

one has

$$\|v\|_{H^s_{\mathrm{sc}}} \overset{\mathrm{def}}{=} \sup_{h \in]0,1]} \|v_h\|_{H^s_h} < +\infty.$$

Then $\mathrm{Op}_h^{\mathrm{W}}(a)$ acts from H^s_{sc} to H^{s-m}_{sc}. Consider now a solution u to equation (32) and define a new function $v(t, x)$, related to u through

$$u(t, x) = \frac{1}{\sqrt{t}} v\Big(t, \frac{x}{t}\Big). \tag{43}$$

Set $h = \frac{1}{t}$. Then v solves the equation

$$\big(D_t - \mathrm{Op}_h^{\mathrm{W}}(x\xi + p(\xi))\big)v = h[\alpha_3 v^3 + \alpha_1 |v|^2 v + \alpha_{-1} |v|^2 \bar{v} + \alpha_{-3} \bar{v}^3]. \tag{44}$$

Let us introduce the set

$$\Lambda = \{(x, \xi) \in \mathbb{R}^2; x + p'(\xi) = 0\}. \tag{45}$$

The basic idea is that this set carries the most important part of the solution, so that one may deduce an ODE from (44) restricting the symbol $x\xi + p(\xi)$ to Λ, and showing that the error one generates in that way decays like an integrable power of t when t goes to infinity.

A key point is that, in the case $p(\xi) = |\xi|^{\frac{1}{2}}$ corresponding to the linearized water waves in our model (32), Λ is a graph: there is a smooth function $\varphi : \mathbb{R}^* \to \mathbb{R}$, given by $\varphi(x) = -\frac{1}{4|x|}$ such that

$$\Lambda = \{(x, \xi) \in \mathbb{R}^2; \xi = d\varphi(x)\}. \tag{46}$$

We shall ignore in the rest of this discussion the technicalities related to the behaviour of φ at zero or infinity and shall do like if v where spectrally supported on a compact subset of $\mathbb{R}-\{0\}$, i.e. we shall assume (abusively) that $v = \mathrm{Op}_h^{\mathrm{W}}(\chi(\xi))v$ for some χ in $C_0^\infty(\mathbb{R}-\{0\})$. We take next γ in $C_0^\infty(\mathbb{R})$, equal to one close to zero, and define, inspired by Ifrim and Tataru [2015a],

$$v_\Lambda = \mathrm{Op}_h^{\mathrm{W}}\Big(\gamma\Big(\frac{x+p'(\xi)}{\sqrt{h}}\Big)\Big)v,\ v_{\Lambda^c} = v - v_\Lambda \tag{47}$$

where the choice of the width $\sqrt{h}$ in the cut-off function is the optimal one. Then our aim is to get for v_Λ an ordinary differential equation.

Proposition 6. *Let v be a solution of (44). Assume that for t in some interval $[1,T]$ the a priori estimates (36) hold true. Then, if we define $\omega(x) = xd\varphi(x) + p(d\varphi(x))$, v_Λ solves*

$$(D_t - \omega(x))v_\Lambda = h\big[\alpha_3 {v_\Lambda}^3 + \alpha_1|v_\Lambda|^2 v_\Lambda + \alpha_{-1}|v_\Lambda|^2\bar{v}_\Lambda + \alpha_{-3}\bar{v}_\Lambda^3\big] + O_{L^\infty}(\epsilon h^{1+\delta}), \tag{48}$$

where δ is a small positive number.

Idea of proof: The proof of the proposition relies on the following facts. First, the contribution v_{Λ^c} defined in (47) will have better estimates than v: this follows from the fact that by definition

$$v_{\Lambda^c} = \mathrm{Op}_h^{\mathrm{W}}\Big(\gamma_1\Big(\frac{x+p'(\xi)}{\sqrt{h}}\Big)\frac{x+p'(\xi)}{\sqrt{h}}\Big)v \simeq \mathrm{Op}_h^{\mathrm{W}}\Big(\gamma_1\Big(\frac{x+p'(\xi)}{\sqrt{h}}\Big)\Big)(\sqrt{h}\mathfrak{L}v) \tag{49}$$

where $\gamma_1(z) = \frac{(1-\gamma(z))}{z}$ and $\mathfrak{L} = \frac{1}{h}\mathrm{Op}_h^{\mathrm{W}}(x+p'(\xi))$. It turns out that $\mathfrak{L}$ may be expressed from the Klainerman vector field Z and the equation, so that an a priori bound of the form (A) in (36) implies that

$$\|\mathfrak{L}v\|_{L^2} = O(h^{-K\epsilon^2}) \tag{50}$$

and thus $\|v_{\Lambda^c}\|_{L^2} = O(h^{\frac12 - K\epsilon^2})$. As the cut-off γ_1 in (49) localizes essentially in a strip of frequencies of size $h^{\frac12}$, a semiclassical Sobolev inequality provides estimates of the form $\|v_{\Lambda^c}\|_{L^\infty} = O(\epsilon h^{\frac14 - K\epsilon^2})$. This allows to replace in the right hand side of (44) v by v_Λ, up to contributions to the remainder. One may also perform such a replacement in the left hand side using some commutation arguments, ending up with an equation of the form

$$\big(D_t - \mathrm{Op}_h^{\mathrm{W}}(x\xi + p(\xi)\big)v_\Lambda = h\big[\alpha_3 {v_\Lambda}^3 + \alpha_1|v_\Lambda|^2 v_\Lambda + \alpha_{-1}|v_\Lambda|^2\bar{v}_\Lambda + \alpha_{-3}\bar{v}_\Lambda^3\big] + O(\epsilon h^{\frac54 - \delta}). \tag{51}$$

Finally, as $\frac{d}{d\xi}(x\xi + p(\xi)) = x + p'(\xi)$ vanishes on Λ, and as this set may be represented using (46), one may write through a Taylor expansion at $\xi = d\varphi(x)$,

$$x\xi + p(\xi) = xd\varphi(x) + p(d\varphi(x)) + O((\xi - d\varphi(x))^2). \tag{52}$$

Since we have restricted our considerations to the case of ξ staying in a compact subset of $\mathbb{R}^*$ (which is equivalent to x staying in a compact subset of $\mathbb{R}^*$ when (x, ξ) is close to Λ), one may rewrite this as

$$x\xi + p(\xi) = \omega(x) + O((x + p'(\xi))^2).$$

Recalling the definition (47) of v_Λ, we deduce from that

$$\begin{aligned}\mathrm{Op}_h^{\mathrm{W}}(x\xi + p(\xi))v_\Lambda &= \omega(x)v_\Lambda \\ &\quad + \text{ term in } \mathrm{Op}_h^{\mathrm{W}}((x + p'(\xi))^2)\mathrm{Op}_h^{\mathrm{W}}\Big(\gamma\Big(\frac{x + p'(\xi)}{\sqrt{h}}\Big)\Big)v.\end{aligned} \tag{53}$$

The last term above may be written, up to remainders, as

$$\sqrt{h}\mathrm{Op}_h^{\mathrm{W}}\Big(\frac{x + p'(\xi)}{\sqrt{h}}\gamma\Big(\frac{x + p'(\xi)}{\sqrt{h}}\Big)\Big)\underbrace{\mathrm{Op}_h^{\mathrm{W}}(x + p'(\xi))v}_{=h\mathfrak{L}v}. \tag{54}$$

Combining with (50), and the fact that the localization of v_Λ allows one to estimate L^∞ norms from $h^{-\frac{1}{4}}$ times Sobolev ones, we obtain that the L^∞ norm of (54) is $O(\epsilon h^{\frac{5}{4}-\delta})$ for some small $\delta > 0$, so that the last term in (53) may be incorporated in the remainder (48). Plugging (53) in (51), one gets (48). ☐

End of proof of Theorem 2: As explained at the end of Section 5, to conclude the proof of the theorem, one has to show that, for a solution u of (33), the a priori bounds (36) imply (37). We have already seen that that (A') holds and we are left with showing the L^∞ estimate (B'). As we have seen that v_{Λ^c} enjoys good a priori bounds, one has to obtain a uniform bound for the solution v_Λ of the ODE (48). Performing a normal form, one may reduce (48) to an equation

$$(D_t - \omega(x))f = h\alpha_1|f|^2 f + O_{L^\infty}(\epsilon h^{1+\delta}),\ h = \frac{1}{t}, \tag{55}$$

where f is a new unknown related to v_Λ in such a way that a uniform control of f is equivalent to a uniform control of v_Λ. As we assumed that α_1 in (33) is *real*, if there were no remainder in $\epsilon t^{-1-\delta}$ in (55), one would get immediately that $\partial_t|f|^2 = 0$, whence the uniform control of f. Since the $O(\epsilon t^{-1-\delta})$ remainder in (55) is integrable, one may show that such a uniform control still holds for the solution of (55). If one makes this reasoning keeping track of the dependence of the constants on A, B, one may prove that (B') follows from (A), (B). Moreover, the analysis of the ODE provides as well the asymptotics of the solution when t goes to $+\infty$.

This outline of proof concerns the simplified model (32). In the case of the water waves system, the general strategy of the last part of the proof is similar. The fact that the coefficient α_1 in (33) is real (that plays a crucial role above) is a "null condition" in the sense of Christodoulou-Klainerman, that holds true because of the structure of the water waves system. □

7 Further results

Our goal in this section is to discuss further results of long time existence, concerning equation (9) under different assumptions.

We consider first the case of initial data that decay in space. We have discussed up to now, for such data, equation (9) when $g > 0, \kappa = 0$ and the depth of the fluid is infinite. We give here references to other global existence results, under other assumptions.

Water waves in infinite depth. We consider gravity waves ($g > 0, \kappa = 0$) in infinte depth, as in Theorem 2. In that statement, and in the results of Ionescu and Pusateri [2015b] and Ifrim and Tataru [2016], the assumptions of smallness of the initial data involve norms that control the energy. It turns out that one may weaken these conditions: Wang [2015a] proved a global existence result for the gravity water waves equation in one space dimension, for infinite depth fluids, when the initial data $(\eta, |D_x|^{\frac{1}{2}}\psi)$ belongs to some homogeneous Sobolev space that contains functions with infinite energy.

Capillary and capillary-gravity waves in infinite depth. Consider first equation (9) with $g = 0$ and $\kappa > 0$, still for a two dimensional fluid of infinite depth, (i.e. x in (9) varies in $\mathbb{R}$). For small, smooth and decaying initial data, global existence has been proved independently by Ifrim and Tataru [2017] and by Ionescu and Pusateri [2015a]. As far as we know, no (almost) global existence result is known for solutions of the full capillary-gravity problem ((9) with $g > 0, \kappa > 0$) in infinite depth, when the space dimension is equal to one. On the other hand, the similar problem in *two space dimensions* (i.e. for three dimensional fluids) has been solved by Deng, Ionescu, Pausader and Pusateri. They proved global existence for small smooth decaying data in Deng, Ionescu, Pausader, and Pusateri [2015].

Water waves in finite depth. Much less results are known concerning global existence of solutions when one works with a fluid of *finite depth*. The only results we are aware of concern the case of two space dimensions (three dimensional fluids) with $g > 0, \kappa = 0$ or $g = 0, \kappa > 0$ in (9). Wang studied Wang [2017b, 2015b, 2017a] the existence of global solutions for small, smooth, decaying initial data.

To finish this section, let us also discuss related results of long time existence, when one considers small but not necessarily decaying initial data. In this case, one cannot expect to use dispersion in order to get a longer existence time than the one that holds in general for a quadratic non linear hyperbolic equation with small data of size ϵ, namely $T_\epsilon \sim \epsilon^{-1}$. Nevertheless, we have seen in Section 4 that a normal forms procedure may allow one to reduce essentially the equation to a cubic one, and so allows to expect an existence time T_ϵ bounded from below by $c\epsilon^{-2}$.

For fluids of infinite depth, such a property has been proved in the case of gravity waves ($g > 0, \kappa = 0$) by Totz and Wu [2012] in one space dimension and by Totz [2015] in two space dimensions. In the case of capillary waves ($g = 0, \kappa > 0$) in one space dimension, a similar result has been obtained by Ifrim and Tataru [2017] and by Ionescu and Pusateri [2015a]. When one considers a constant non zero vorticity, a lower bound in $c\epsilon^{-2}$ for the time of existence of solutions has been shown by Ifrim and Tataru [2015b], still in one space dimension.

Regarding finite depth fluids, Harrop-Griffiths, Ifrim, and Tataru [2017] have proved a $c\epsilon^{-2}$ lower bound for the existence time, in the gravity waves case ($g > 0, \kappa = 0$) in one space dimension.

The above results apply in particular when one considers initial data that are periodic in space, i.e. defined on the circle. In such a case, better results may be obtained under stronger assumptions. First, it is possible to construct special classes of global solutions. Actually, Plotnikov and Toland [2001] (resp. Iooss, Plotnikov, and Toland [2005]) constructed, for the gravity waves system in finite (resp. infinite) depth, standing waves solutions. For the full gravity-capilarity system in infinite depth, Alazard and Baldi [2015] did the same. Later, Berti and Montalto [2017] built up time quasi-periodic solutions of system (9) in infinite depth, and more recently Baldi, Berti, Haus, and Montalto [2017] treated the same problem in finite depth.

The preceding results do not concern the Cauchy problem, as one constructs *special* solutions. But combining some of the ideas of Alazard and Baldi [2015] and normal forms methods, Berti and Delort [2017] proved that system (9), with even periodic initial data of size ϵ, has solutions defined up to time $c_N\epsilon^{-N}$ for any N, when the parameters (g, κ) avoid an exceptional subset of zero measure.

References

T. Alazard, N. Burq, and C. Zuily (2011a). "On the water-wave equations with surface tension". *Duke Math. J.* 158.3, pp. 413–499. MR: 2805065 (cit. on p. 2262).

– (2014). "On the Cauchy problem for gravity water waves". *Invent. Math.* 198.1, pp. 71–163. MR: 3260858 (cit. on p. 2262).

– (2016). "Cauchy theory for the gravity water waves system with non-localized initial data". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 33.2, pp. 337–395. MR: 3465379 (cit. on p. 2262).

Thomas Alazard and Pietro Baldi (2015). "Gravity capillary standing water waves". *Arch. Ration. Mech. Anal.* 217.3, pp. 741–830. MR: 3356988 (cit. on p. 2276).

Thomas Alazard, Nicolas Burq, and Claude Zuily (2011b). "Strichartz estimates for water waves". *Ann. Sci. Éc. Norm. Supér. (4)* 44.5, pp. 855–903. MR: 2931520 (cit. on p. 2262).

Thomas Alazard and Jean-Marc Delort (2015a). "Global solutions and asymptotic behavior for two dimensional gravity water waves". *Ann. Sci. Éc. Norm. Supér. (4)* 48.5, pp. 1149–1238. MR: 3429478 (cit. on pp. 2263–2265, 2270, 2271).

– (2015b). "Sobolev estimates for two dimensional gravity water waves". *Astérisque* 374, pp. viii+241. MR: 3460636 (cit. on pp. 2263, 2265).

Thomas Alazard and Guy Métivier (2009). "Paralinearization of the Dirichlet to Neumann operator, and regularity of three-dimensional water waves". *Comm. Partial Differential Equations* 34.10-12, pp. 1632–1704. MR: 2581986 (cit. on p. 2262).

S. Alinhac (1986). "Paracomposition et opérateurs paradifférentiels". *Comm. Partial Differential Equations* 11.1, pp. 87–121. MR: 814548 (cit. on p. 2262).

– (1988). "Ondes de raréfaction pour des systèmes quasi-linéaires hyperboliques multidimensionnels". In: *Journées "Équations aux Dérivées Partielles" (Saint Jean de Monts, 1988)*. École Polytech., Palaiseau, Exp. No. VIII, 7. MR: 975497 (cit. on p. 2262).

David M. Ambrose (2003). "Well-posedness of vortex sheets with surface tension". *SIAM J. Math. Anal.* 35.1, pp. 211–244. MR: 2001473 (cit. on p. 2262).

David M. Ambrose and Nader Masmoudi (2005). "The zero surface tension limit of two-dimensional water waves". *Comm. Pure Appl. Math.* 58.10, pp. 1287–1315. MR: 2162781 (cit. on p. 2262).

Pietro Baldi, Massimiliano Berti, Emanuele Haus, and Riccardo Montalto (2017). "Time quasi-periodic gravity water waves in finite depth". arXiv: 1708.01517 (cit. on p. 2276).

Massimiliano Berti and Jean-Marc Delort (2017). "Almost global existence of solutions for capillary-gravity water waves equations with periodic spatial boundary conditions". *hal-01457217* (cit. on p. 2276).

Massimiliano Berti and Riccardo Montalto (2017). "KAM for gravity capillary water waves". *Memoires of AMS* (cit. on p. 2276).

Jean-Michel Bony (1981). "Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires". *Ann. Sci. École Norm. Sup. (4)* 14.2, pp. 209–246. MR: 631751 (cit. on p. 2262).

Demetrios Christodoulou and Hans Lindblad (2000). "On the motion of the free surface of a liquid". *Comm. Pure Appl. Math.* 53.12, pp. 1536–1602. MR: 1780703 (cit. on p. 2262).

Daniel Coutand and Steve Shkoller (2007). "Well-posedness of the free-surface incompressible Euler equations with or without surface tension". *J. Amer. Math. Soc.* 20.3, pp. 829–930. MR: 2291920 (cit. on p. 2262).

W. Craig and C. Sulem (1993). "Numerical simulation of gravity waves". *J. Comput. Phys.* 108.1, pp. 73–83. MR: 1239970 (cit. on p. 2260).

Walter Craig (1985). "An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits". *Comm. Partial Differential Equations* 10.8, pp. 787–1003. MR: 795808 (cit. on p. 2261).

J.-M. Delort (2009). "Long-time Sobolev stability for small solutions of quasi-linear Klein-Gordon equations on the circle". *Trans. Amer. Math. Soc.* 361.8, pp. 4299–4365. MR: 2500890 (cit. on p. 2269).

– (2012). "A quasi-linear Birkhoff normal forms method. Application to the quasi-linear Klein-Gordon equation on $\mathbb{S}^1$". *Astérisque* 341, pp. vi+113. MR: 2952065 (cit. on p. 2269).

– (2015). "Quasi-linear perturbations of Hamiltonian Klein–Gordon equations on spheres". *Mem. Amer. Math. Soc.* 234.1103, pp. vi+80. MR: 3288855 (cit. on p. 2269).

Yu Deng, Alexandru D. Ionescu, Benoît Pausader, and Fabio Pusateri (2015). "Global solutions of the gravity-capilarity water waves system in 3 dimensions". arXiv: 1601.05685 (cit. on p. 2275).

P. Germain, N. Masmoudi, and J. Shatah (2012a). "Global solutions for 2D quadratic Schrödinger equations". *J. Math. Pures Appl. (9)* 97.5, pp. 505–543. MR: 2914945 (cit. on p. 2266).

– (2012b). "Global solutions for the gravity water waves equation in dimension 3". *Ann. of Math. (2)* 175.2, pp. 691–754. MR: 2993751 (cit. on p. 2263).

Benjamin Harrop-Griffiths, Mihaela Ifrim, and Daniel Tataru (2017). "Finite depth gravity water waves in holomorphic coordinates". *Ann. PDE* 3.1, Art. 4, 102. MR: 3625189 (cit. on p. 2276).

John K. Hunter, Mihaela Ifrim, and Daniel Tataru (2016). "Two dimensional water waves in holomorphic coordinates". *Comm. Math. Phys.* 346.2, pp. 483–552. MR: 3535894 (cit. on pp. 2263, 2269).

John K. Hunter, Mihaela Ifrim, Daniel Tataru, and Tak Kwong Wong (2015). "Long time solutions for a Burgers-Hilbert equation via a modified energy method". *Proc. Amer. Math. Soc.* 143.8, pp. 3407–3412. MR: 3348783 (cit. on p. 2269).

Mihaela Ifrim and Daniel Tataru (2015a). "Global bounds for the cubic nonlinear Schrödinger equation (NLS) in one space dimension". *Nonlinearity* 28.8, pp. 2661–2675. MR: 3382579 (cit. on pp. 2270, 2271, 2273).

– (2015b). "Two dimensional gravity water waves with constant vorticity: I. Cubic lifespan". arXiv: 1510.07732 (cit. on p. 2276).

– (2016). "Two dimensional water waves in holomorphic coordinates II: Global solutions". *Bull. Soc. Math. France* 144.2, pp. 369–394. MR: 3499085 (cit. on pp. 2263–2265, 2269–2271, 2275).

– (2017). "The Lifespan of Small Data Solutions in Two Dimensional Capillary Water Waves". *Arch. Ration. Mech. Anal.* 225.3, pp. 1279–1346. MR: 3667289 (cit. on pp. 2275, 2276).

Alexandru D. Ionescu and Fabio Pusateri (2015a). "Global regularity for $2D$ water waves with surface tension". arXiv: 1408.4428 (cit. on pp. 2275, 2276).

– (2015b). "Global solutions for the gravity water waves system in 2d". *Invent. Math.* 199.3, pp. 653–804. MR: 3314514 (cit. on pp. 2263–2265, 2269–2271, 2275).

G. Iooss, P. I. Plotnikov, and J. F. Toland (2005). "Standing waves on an infinitely deep perfect fluid under gravity". *Arch. Ration. Mech. Anal.* 177.3, pp. 367–478. MR: 2187619 (cit. on p. 2276).

David Lannes (2005). "Well-posedness of the water-waves equations". *J. Amer. Math. Soc.* 18.3, pp. 605–654. MR: 2138139 (cit. on p. 2262).

– (2013a). "Space time resonances [after Germain, Masmoudi, Shatah]". *Astérisque* 352. Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058, Exp. No. 1053, ix, 355–388. MR: 3087351 (cit. on p. 2266).

– (2013b). *The water waves problem*. Vol. 188. Mathematical Surveys and Monographs. Mathematical analysis and asymptotics. American Mathematical Society, Providence, RI, pp. xx+321. MR: 3060183 (cit. on p. 2260).

Hans Lindblad (2005). "Well-posedness for the motion of an incompressible liquid with free surface boundary". *Ann. of Math. (2)* 162.1, pp. 109–194. MR: 2178961 (cit. on p. 2262).

V. I. Nalimov (1974). "The Cauchy-Poisson problem". *Dinamika Splošn. Sredy* Vyp. 18 Dinamika Židkost. so Svobod. Granicami, pp. 104–210, 254. MR: 0609882 (cit. on p. 2261).

Tohru Ozawa, Kimitoshi Tsutaya, and Yoshio Tsutsumi (1997). "Remarks on the Klein-Gordon equation with quadratic nonlinearity in two space dimensions". In: *Nonlinear waves (Sapporo, 1995)*. Vol. 10. GAKUTO Internat. Ser. Math. Sci. Appl. Gakkōtosho, Tokyo, pp. 383–392. MR: 1602662 (cit. on p. 2267).

P. I. Plotnikov and J. F. Toland (2001). "Nash-Moser theory for standing water waves". *Arch. Ration. Mech. Anal.* 159.1, pp. 1–83. MR: 1854060 (cit. on p. 2276).

Jalal Shatah (1985). "Normal forms and quadratic nonlinear Klein-Gordon equations". *Comm. Pure Appl. Math.* 38.5, pp. 685–696. MR: 803256 (cit. on p. 2266).

Jalal Shatah and Chongchun Zeng (2008a). "A priori estimates for fluid interface problems". *Comm. Pure Appl. Math.* 61.6, pp. 848–876. MR: 2400608 (cit. on p. 2262).

– (2008b). "Geometry and a priori estimates for free boundary problems of the Euler equation". *Comm. Pure Appl. Math.* 61.5, pp. 698–744. MR: 2388661 (cit. on p. 2262).

– (2011). "Local well-posedness for fluid interface problems". *Arch. Ration. Mech. Anal.* 199.2, pp. 653–705. MR: 2763036 (cit. on p. 2262).

Nathan Totz (2015). "A justification of the modulation approximation to the 3D full water wave problem". *Comm. Math. Phys.* 335.1, pp. 369–443. MR: 3314508 (cit. on p. 2276).

Nathan Totz and Sijue Wu (2012). "A rigorous justification of the modulation approximation to the 2D full water wave problem". *Comm. Math. Phys.* 310.3, pp. 817–883. MR: 2891875 (cit. on pp. 2269, 2276).

Xuecheng Wang (2015a). "Global infinite energy solutions for the $2D$ gravity water waves system". arXiv: 1502.00687 (cit. on p. 2275).

– (2015b). "Global solutions for the $3D$ gravity water waves system above a flat bottom". arXiv: 1508.06227 (cit. on p. 2276).

– (2017a). "Global regularity for the $3D$ finite depth capillary water waves". arXiv: 1611.05472 (cit. on p. 2276).

– (2017b). "On the 3-dimensional water waves system above a flat bottom". *Anal. PDE* 10.4, pp. 893–928. MR: 3649370 (cit. on p. 2276).

Sijue Wu (1997). "Well-posedness in Sobolev spaces of the full water wave problem in 2-D". *Invent. Math.* 130.1, pp. 39–72. MR: 1471885 (cit. on p. 2261).

– (1999). "Well-posedness in Sobolev spaces of the full water wave problem in 3-D". *J. Amer. Math. Soc.* 12.2, pp. 445–495. MR: 1641609 (cit. on p. 2261).

– (2009). "Almost global wellposedness of the 2-D full water wave problem". *Invent. Math.* 177.1, pp. 45–135. MR: 2507638 (cit. on pp. 2263, 2269).

– (2011). "Global wellposedness of the 3-D full water wave problem". *Invent. Math.* 184.1, pp. 125–220. MR: 2782254 (cit. on p. 2263).

Hideaki Yosihara (1982). "Gravity waves on the free surface of an incompressible perfect fluid of finite depth". *Publ. Res. Inst. Math. Sci.* 18.1, pp. 49–96. MR: 660822 (cit. on p. 2261).

V. E. Zakharov (1968). "Stability of periodic waves of finite amplitude on the surface of a deep fluid". *Journal of Applied Mechanics and Technical Physics* 9.2, pp. 190–1943 (cit. on p. 2260).

Received 2017-12-05.

JEAN-MARC DELORT

Université Paris 13
Sorbonne Paris Cité, LAGA, CNRS (UMR 7539)
99, Avenue J.-B. Clément
F-93430 Villetaneuse
France
delort@math.univ-paris13.fr

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2279–2304)

SYMMETRY AND SYMMETRY BREAKING: RIGIDITY AND FLOWS IN ELLIPTIC PDES

Jean Dolbeault, Maria J. Esteban and Michael Loss

Abstract

The issue of symmetry and symmetry breaking is fundamental in all areas of science. Symmetry is often assimilated to order and beauty while symmetry breaking is the source of many interesting phenomena such as phase transitions, instabilities, segregation, self-organization, *etc.* In this contribution we review a series of sharp results of symmetry of nonnegative solutions of nonlinear elliptic differential equation associated with minimization problems on Euclidean spaces or manifolds. Nonnegative solutions of those equations are unique, a property that can also be interpreted as a rigidity result. The method relies on linear and nonlinear flows which reveal deep and robust properties of a large class of variational problems. Local results on linear instability leading to symmetry breaking and the bifurcation of non-symmetric branches of solutions are reinterpreted in a larger, global, variational picture in which our flows characterize directions of descent.

1 Introduction

Symmetries are fundamental properties of the laws of Physics. They impose constraints on modeling phenomena and, at a more basic level, they serve as criteria of classification. Inspired by his work in crystallography, Pierre Curie made an early attempt (in 1894) to investigate the consequences of symmetries. Since then, symmetry has been an important preoccupation for many scientists.

More intriguing than symmetry is the phenomenon of *symmetry breaking*, which asserts that the state of a system may have less symmetries than the underlying physical laws.

Partially supported by the projects *Kibord* and *EFI* (J.D.) of the French National Research Agency (ANR), and by the NSF grant DMS-1301555 (M.L.).

MSC2010: primary 35J20; secondary 49K30, 53C21.

Keywords: Symmetry, symmetry breaking, interpolation inequalities, Caffarelli–Kohn–Nirenberg inequalities, optimal constants, rigidity results, fast diffusion equation, carré du champ, bifurcation, instability.

Among various considerations on the causes of the symmetries and what these symmetries mean in physics, P. Curie wrote in Curie [1894] that

C'est la dissymétrie qui crée le phénomène.

In mathematical terms, "dissymétrie" shifts the attention to solutions which may have less symmetries than the problem they solve. *Symmetry breaking*, especially spontaneous symmetry breaking, has been an incredibly fruitful concept over the last century. It appears in mechanics (buckling instabilities), in particle physics, in the description of phase transitions or complex dynamics, *etc.* One of the basic mechanisms is the bifurcation phenomenon in nonlinear systems, which has to do with the stability analysis of symmetric states.

Symmetry has attracted the attention of mathematicians for diverse reasons which range from assertions like "symmetry is beautiful" to practical motivations: symmetry simplifies the search of solutions and makes their computation more tractable from a numerical point of view by reducing the number of degrees of freedom.

Entropy methods have a long history in various fields of Science and in particular of Mathematics. The notion of entropy that we shall consider here is inspired by results in the theory of nonlinear PDEs and especially nonlinear diffusion equations. It borrows tools from Kinetic Theory and from Information Theory. Other major sources of inspiration are the *carré du champ* method used in the study of Semi-groups and Markov processes as well as the *rigidity* (uniqueness) techniques in the Theory of Nonlinear Elliptic Equations. In addition to the application to symmetry issues, one of our contributions was to rephrase these two approaches in a common framework of parabolic equations and to emphasize the role of the nonlinear diffusions in the search for optimal ranges and optimal constants in related interpolation inequalities.

It is definitely out of reach to give even a partial account of all mathematical issues of symmetry and symmetry breaking in this paper, so we shall focus on PDEs with two main examples: the first one is the equation

$$-\operatorname{div}\left(|x|^{-\beta}\,\nabla w\right)=|x|^{-\gamma}\left(w^{2p-1}-w^{p}\right)\quad\text{in}\quad\mathbb{R}^d\setminus\{0\}\,,$$

which has an interesting feature: there is a competition between nonlinearities and weights. The solutions can be interpreted as critical points of an energy functional. Without weights, solutions are radially symmetric (up to translations). With weights and in some regime of the parameters β, γ and p, non-radial solutions are energetically more favorable. Since we are interested in energy minimizers, as a particular sub-problem, understanding who wins in the competition is a central question.

Alternatively, we shall consider the equation

$$-\Delta\varphi+\Lambda\,\varphi=\varphi^{p-1}\quad\text{on}\quad\mathfrak{M}\,,$$

where $\mathfrak{M}$ is a sphere, a compact manifold or a cylinder. In that case, the geometric properties of the manifold replace the weight and compete with the scale induced by the parameter Λ. If there is enough space, in a precise sense that can be measured, then solutions with less symmetry may have a lower energy.

These two equations, although very simple because the nonlinearities (and also the weights in the case of the first equation) obey power laws, are not purely academic. For one, the solutions (and the associated functional inequalities) are of direct interest for instance in some models of fluid mechanics. More important is the fact that power laws appear in many problems when scalings or blow-up methods are used to extract an asymptotic behavior. Hence, we expect that our model equations lie at the core of many nonlinear or weighted problems. Finally, models involving power laws have the advantage that they can be treated by using *nonlinear flows* and *entropy methods*. Indeed we are able to give sharp results of rigidity for the equation, and symmetry results for the optimal functions associated with related interpolation inequalities.

Because of the confluence of various branches of analysis such as non-linear diffusion and the calculus of variations, and the fundamental nature of the above equations, we believe that it is worth studying them in great detail, with sharp stability results and sharp constants in the functional inequalities. Note that this amounts to establishing the exact range of the parameters for which extremal functions are symmetric. Variational issues of the symmetry and symmetry breaking will be detailed below.

Let us fix some notations and conventions. Throughout this paper, we shall use the notation $2^*:=\frac{2\,d}{d-2}$ if $d\ge 3$, and $2^*:=\infty$ if $d=1$ or 2. We shall say that a function is an *extremal function* for an optimal functional inequality if equality holds in the inequality. To simplify notations, parameters will be omitted whenever they are not essential for the understanding of the strategy of proof. This paper is a review of various results which were published in several papers (references will appear in the text) and are collected together for the first time. The reader is invited to pay attention that some notations have been redefined compared to the original papers.

2 Interpolation inequalities and flows on compact manifolds

2.1 Interpolation inequalities on $\mathbb{S}^d$. Let us consider the inequality

$$(1)\qquad \|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}+\frac{d}{p-2}\,\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}\ge\frac{d}{p-2}\,\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)}\quad\forall\,u\in\mathrm{H}^1(\mathbb{S}^d,d\mu)$$

where $d\mu$ is the uniform probability measure induced by the Lebesgue measure on $\mathbb{S}^d \subset \mathbb{R}^{d+1}$. Here the exponent p is such that $1 \le p < 2$ or $2 < p < 2^*$, or $p = 2^*$ if $d \ge 3$. The case $p = 2^*$ corresponds to the usual Sobolev inequality on $\mathbb{S}^d$ or, using the stereographic projection, to the Sobolev inequality in $\mathbb{R}^d$. In the limit case as $p \to 2$, we recover the logarithmic Sobolev inequality

$$(2) \qquad \|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)} \ge \frac{d}{2} \int_{\mathbb{S}^d} |u|^2 \log\left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu \quad \forall\, u \in \mathrm{H}^1(\mathbb{S}^d, d\mu) \setminus \{0\}\,.$$

In (1) and (2), equality is achieved by any constant non-zero function. The value of the optimal constants, $d/(p-2)$ and $d/2$ is obtained by linearization: if φ is an eigenfunction associated with the first positive eigenvalue of the Laplace-Beltrami operator on $\mathbb{S}^d$, the infimum of

$$\frac{(p-2)\,\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\|u\|^2_{\mathrm{L}^p(\mathbb{S}^d)} - \|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}} \quad \text{and} \quad \frac{2\,\|\nabla u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}{\int_{\mathbb{S}^d} |u|^2 \log\left(\frac{|u|^2}{\|u\|^2_{\mathrm{L}^2(\mathbb{S}^d)}}\right) d\mu}\,,$$

respectively for $p \neq 2$ and for $p = 2$, is achieved by $u = 1 + \varepsilon\,\varphi$ in the limit as $\varepsilon \to 0$.

Inequality (1) has been established in Bidaut-Véron and Véron [1991] by rigidity methods, in Beckner [1993] by techniques of harmonic analysis, and using the *carré du champ* method in Bentaleb [1993], Bakry and Ledoux [1996], and Demange [2008], for any $p > 2$. The case $p = 2$ was studied in Mueller and Weissler [1982]. In Bakry and Émery [1984, 1985a,b], D. Bakry and M. Emery proved the inequalities under the restriction

$$2 < p \le 2^\# := \frac{2\,d^2 + 1}{(d-1)^2}\,.$$

Their method relies on a linear heat flow method which is presented below, as well as a nonlinear flow which allow us to get rid of this restriction.

2.2 Flows and *carré du champ* methods on $\mathbb{S}^d$.

We start by the linear heat flow method of Bakry and Émery [1985b]. For any function $\rho > 0$ we define a *generalized entropy functional* $\mathcal{E}_p$ and a *generalized Fisher information functional* I_p by

$$\mathcal{E}_p[\rho] := \frac{1}{p-2}\left[\left(\int_{\mathbb{S}^d} \rho\, d\mu\right)^{\frac{2}{p}} - \int_{\mathbb{S}^d} \rho^{\frac{2}{p}}\, d\mu\right],$$

$$\mathcal{E}_2[\rho] := \frac{1}{2}\int_{\mathbb{S}^d} \rho \log\left(\frac{\rho}{\|\rho\|_{\mathrm{L}^1(\mathbb{S}^d)}}\right) d\mu$$

if $p \neq 2$ or $p = 2$, respectively, and

$$I_p[\rho] := \int_{\mathbb{S}^d} |\nabla \rho^{\frac{1}{p}}|^2 \, d\mu \,.$$

With this notation, (1) and (2) amount to $I_p[\rho] \geq d\, \mathcal{E}_p[\rho]$ as can be checked using $\rho = |u|^p$. Let us consider the heat flow

$$\frac{\partial \rho}{\partial t} = \Delta \rho \tag{3}$$

where Δ denotes the Laplace-Beltrami operator on $\mathbb{S}^d$, and compute

$$\frac{d}{dt}\mathcal{E}_p[\rho] = -\,2\, I_p[\rho] \quad \text{and} \quad \frac{d}{dt} I_p[\rho] \leq -\,2\, d\; I_p[\rho]$$

where the differential inequality holds if $p \leq 2^\#$. Under this condition, we obtain that

$$\frac{d}{dt}\Big(I_p[\rho] - d\, \mathcal{E}_p[\rho]\Big) \leq 0\,.$$

On the other hand, $\rho(t,\cdot)$ converges as $t \to \infty$ to a constant, namely $\int_{\mathbb{S}^d} \rho\, d\mu$ since $d\mu$ is a probability measure and $\int_{\mathbb{S}^d} \rho\, d\mu$ is conserved by (3). As a consequence, $\lim_{t\to\infty} \big(I_p[\rho] - d\, \mathcal{E}_p[\rho$ 0, which proves that $I_p[\rho(t,\cdot)] - d\, \mathcal{E}_p[\rho(t,\cdot)]$ is nonnegative for any $t \geq 0$ and completes the proof. See Bakry and Émery [ibid.] for details. One may wonder whether the monotonicity property is also true for some $p > 2^\#$. The following result contains a negative answer to this question.

Proposition 1. *Dolbeault, Esteban, and Loss [2017] For any $p \in (2^\#, 2^*)$ or $p = 2^*$ if $d \geq 3$, there exists a function ρ_0 such that, if ρ is a solution of (3) with initial datum ρ_0, then*

$$\frac{d}{dt}\Big(I_p[\rho] - d\, \mathcal{E}_p[\rho]\Big)_{|t=0} > 0\,.$$

The function ρ_0 is explicitly constructed in Dolbeault, Esteban, and Loss [ibid.].

To overcome the limitation $p \leq 2^\#$, one can consider a nonlinear diffusion of fast diffusion or porous medium type

$$\frac{\partial \rho}{\partial t} = \Delta \rho^m\,. \tag{4}$$

With this flow, we no longer have $\frac{d}{dt}\mathcal{E}_p[\rho] = -\,I_p[\rho]$ but we can still prove that

$$\frac{d}{dt}\Big(I_p[\rho] - d\,\mathcal{E}_p[\rho]\Big) \le 0\,,$$

for any $p \in [1, 2^*]$. Proofs of the latter have been given in Demange [2008] and Dolbeault, Esteban, and Loss [2014]. We also refer to Dolbeault, Esteban, Kowalczyk, and Loss [2013] and Dolbeault, Esteban, and Kowalczyk [2014] for results which are more specific to the case of the sphere, and further references therein. Except for $p = 1$ and $p = 2^*$ with $d \ge 3$, there is some flexibility in the choice of m, which can be used to build deficit functionals and improved inequalities: see Demange [2008] and Dolbeault, Esteban, and Kowalczyk [2014]. Notice that ρ_0 in Proposition 1 is a function related with the nonlinear diffusion equation (4).

The case of $\mathbb{S}^d$ highlights the limitations of linear flows and shows the flexibility and strength of nonlinear flows. At least for $p < 2^*$, the optimal constant in (1) and (2) is established by proving that the minimum of $I_p[\rho] - d\,\mathcal{E}_p[\rho]$ is 0. Earlier results in Bidaut-Véron and Véron [1991], Bakry and Ledoux [1996], and Beckner [1993] can be reinterpreted as a purely elliptic method, which goes as follows. A positive minimizer actually exists by standard compactness arguments and any solution ρ satisfies an Euler-Lagrange equation. By testing the equation with $\Delta\rho^m$, we observe that the solution is a constant and, as a consequence, that $\rho \equiv 1$ because of the normalization. We will rely on a similar observation in the next two sections and refer to this method as the *elliptic method.*

The method applies not only to minimizers, but also to any positive solution of the Euler-Lagrange equations. What we prove is a uniqueness result. Since constant functions are solutions, this proves that there are no non-constant solutions. This is why it is called a *rigidity* result.

Compared to Bidaut-Véron and Véron [1991], Bakry and Ledoux [1996], and Beckner [1993], our approach provides a unified framework for $p > 2$ and $p < 2$ (which is not covered in the above mentioned results). However, the main advantage of the method is that it explains why a local result (the best constant is given by the linearization around the constant functions) is actually global: $I_p[\rho] - d\,\mathcal{E}_p[\rho]$ is strictly monotone decreasing under the action of the flow, unless the solution has reached the unique, trivial stationary state.

2.3 Inequalities on compact manifolds. The nonlinear diffusion flow method applies not only to spheres, but also to general compact manifolds. Without entering in the details, let us state a result of Dolbeault, Esteban, and Loss [2014]. Earlier important references are: Gidas and Spruck [1981], Bidaut-Véron and Véron [1991], Licois and Véron [1998], and Demange [2008], among many other contributions which are listed in Dolbeault, Esteban, and Loss [2014].

Let us assume that $(\mathfrak{M}, g)$ is a smooth compact connected Riemannian manifold of dimension $d \geq 1$, without boundary. We denote by dv_g the volume element, by Δ the Laplace-Beltrami operator on $\mathfrak{M}$, by Ric the Ricci tensor and assume for simplicity that $\mathrm{vol}_g(\mathfrak{M}) = 1$. Let λ_1 be the lowest positive eigenvalue of $-\Delta$ and

$$\lambda_\star := \inf_{u \in \mathrm{H}^2(\mathfrak{M})} \frac{\int_{\mathfrak{M}} \left[(1-\theta)\,(\Delta u)^2 + \frac{\theta\, d}{d-1}\,\mathrm{Ric}(\nabla u, \nabla u)\right] dv_g}{\int_{\mathfrak{M}} |\nabla u|^2\, dv_g}\,,$$

$$\theta = \frac{(d-1)^2\,(p-1)}{d\,(d+2)+p-1}\,.$$

Theorem 2. *With the above notations, if* $0 < \lambda < \lambda_\star$, *then for any* $p \in (1,2) \cup (2, 2^*)$, *the equation*

$$-\Delta v + \frac{\lambda}{p-2}\left(v - v^{p-1}\right) = 0$$

has a unique positive solution in $C^2(\mathfrak{M})$, *which is constant and equal to* 1.

It has been shown in Dolbeault, Esteban, and Loss [ibid.] that nonlinear diffusion flows provide a unified framework for *elliptic rigidity* and *carré du champ* methods. The computations heavily rely on the Bochner-Lichnerowicz-Weitzenböck formula

$$\tfrac{1}{2}\,\Delta\,(|\nabla f|^2) = \|\mathrm{Hess}\, f\|^2 + \nabla\cdot(\Delta f)\cdot\nabla f + \mathrm{Ric}(\nabla f, \nabla f)\,.$$

More general results can be established using the so-called $CD(\rho, N)$ condition (see Bakry, Gentil, and Ledoux [2014] and references therein), but they are formal in most of the cases covered only by nonlinear flows. In dimension $d = 2$, the Moser-Trudinger-Onofri inequality replaces in a certain sense Sobolev's inequality, and it is possible to extend the method described above to cover this case: see Dolbeault, Esteban, and Jankowiak [2017]. Bounded convex domains in $\mathbb{R}^d$ have also been considered in Dolbeault and Kowalczyk [2017] in relation with the Lin-Ni conjecture (homogeneous Neumann boundary conditions). Concerning unbounded domains, subcritical Gagliardo-Nirenberg have been established in the case of the line in Dolbeault, Esteban, Laptev, and Loss [2014] while Rényi entropy powers, which will be essential in Section 4, can be used in $\mathbb{R}^d$ to get sharp interpolation inequalities: see Savaré and Toscani [2014], Toscani [2014], and Dolbeault and Toscani [2016].

3 Rigidity on cylinders and sharp symmetry results in critical Caffarelli-Kohn-Nirenberg inequalities

In this section we use a nonlinear flow to prove rigidity results for nonlinear elliptic problems on non-compact manifolds: cylinders and weigthed Euclidean spaces. All results of this section, and their proofs, can be found in Dolbeault, Esteban, and Loss [2016b].

3.1 Three equivalent rigidity results. Let us consider the spherical cylinder $\mathcal{C} := \mathbb{R} \times \mathbb{S}^{d-1}$ and denote by $s \in \mathbb{R}$ and $\omega \in \mathbb{S}^{d-1}$ the coordinates. Let Δ_ω denote the Laplace-Beltrami operator on $\mathbb{S}^{d-1}$.

Theorem 3. *Let $d \geq 2$. For all $p \in (2, 2^*)$ and $0 < \Lambda \leq \Lambda_{\mathrm{FS}} := 4\,\frac{d-1}{p^2-4}$, any positive solution $\varphi \in \mathrm{H}^1(\mathcal{C})$ of*

$$-\,\partial_s^2\,\varphi - \Delta_\omega\,\varphi + \Lambda\,\varphi = \varphi^{p-1} \quad in \quad \mathcal{C} \tag{5}$$

is, up to a translation in the s-direction, equal to

$$\varphi_\Lambda(s) := \left(\tfrac{p}{2}\,\Lambda\right)^{\frac{1}{p-2}}\,\left(\cosh\left(\tfrac{p-2}{2}\,\sqrt{\Lambda}\,s\right)\right)^{-\frac{2}{p-2}} \quad \forall\, s \in \mathbb{R}\,.$$

For any $\Lambda > \Lambda_{\mathrm{FS}}$, there are also positive solutions which do not depend only on s.

A similar rigidity result holds for non-spherical cylinders $\mathbb{R}\times\mathfrak{M}$ where $\mathfrak{M}$ is a compact manifold, but in this case we cannot characterize the optimal set of parameters Λ with our method: see Dolbeault, Esteban, and Loss [ibid.].

Let

$$a_c := \frac{d-2}{2} \quad \text{and} \quad b_{\mathrm{FS}}(a) := \frac{d\,(a_c - a)}{2\,\sqrt{(a_c-a)^2 + d - 1}} + a - a_c\,.$$

By using the Emden-Fowler transformation

$$v(r,\omega) = r^{a-a_c}\,\varphi(s,\omega) \quad \text{with} \quad r = |x|\,, \quad s = -\log r \quad \text{and} \quad \omega = \frac{x}{r}\,, \tag{6}$$

Theorem 3 is equivalent to the following result.

Theorem 4. *Assume that $d \geq 2$, $a < a_c$ and $\min\{a, b_{\mathrm{FS}}(a)\} < b \leq a+1$. Then any nonnegative solution v of*

$$-\,\nabla\cdot\left(|x|^{-2\,a}\,\nabla v\right) = |x|^{-b\,p}\,|v|^{p-2}\,v \quad in \quad \mathbb{R}^d\setminus\{0\} \tag{7}$$

which satisfies $\int_{\mathbb{R}^d} \frac{|v|^p}{|x|^{b\,p}}, dx < \infty$, *is, up to a scaling, equal to*

$$v_\star(x) = \left(1 + |x|^{(p-2)\,(a_c - a)}\right)^{-\frac{2}{p-2}} \quad \forall\, x \in \mathbb{R}^d\,.$$

If $a < 0$ *and* $a < b < b_{\mathrm{FS}}(a)$, *there are also positive solutions which do not depend only on* $|x|$.

Let us define $\alpha_{\mathrm{FS}} := \sqrt{\frac{d-1}{n-1}}$ and pick n and α such that

$$n = \frac{d - b\,p}{\alpha} = \frac{d - 2\,a - 2}{\alpha} + 2 = \frac{2\,p}{p-2}\,,$$

so that we also have $p = 2\,n/(n-2)$. Next we consider the diffusion operator

$$\mathfrak{L}\,w := \alpha^2\left(w'' + \frac{n-1}{r}\,w'\right) - \frac{1}{r^2}\,\Delta_\omega\,w\,.$$

Then, with the change of variables

$$v(r,\omega) = w(r^\alpha,\omega) \quad \forall\,(r,\omega) \in \mathbb{R}^+ \times \mathbb{S}^{d-1}\,,$$

Theorem 4 is equivalent to

Theorem 5. *Assume that* $n > d \geq 2$ *and* $p = 2\,n/(n-2)$. *If* $0 < \alpha \leq \alpha_{\mathrm{FS}}$, *then any nonnegative solution* $w(x) = w(r,\omega)$ *with* $r \in \mathbb{R}_+$ *and* $\omega \in \mathbb{S}^{d-1}$ *of*

$$-\,\mathfrak{L}\,w = w^{p-1} \quad \textit{in} \quad \mathbb{R}^d \setminus \{0\} \tag{8}$$

which satisfies $\int_{\mathbb{R}^d} |x|^{n-d}\,|w|^p\,dx < \infty$, *is equal, up to a scaling, to*

$$w_\star(x) = \left(1 + |x|^2\right)^{-\frac{n-2}{2}} \quad \forall\, x \in \mathbb{R}^d\,.$$

If $\alpha > \alpha_{\mathrm{FS}}$, *there are also solutions which do not depend only on* $|x|$.

Let us complement these results with some remarks:
(i) If n is an integer, then (8) is the Euler-Lagrange equation associated with the standard Sobolev inequality

$$-\,\alpha^2\,\Delta w = w^{\frac{n+2}{n-2}} \quad \text{in} \quad \mathbb{R}^n\,,$$

where Δ denotes the Laplacian operator in $\mathbb{R}^n$, but in the class of functions which depend only on the first $d-1$ angular variables.

(ii) The conditions on the parameters in Theorems 3, 4 and 5 are equivalent:

$$0<\Lambda\le\Lambda_{\rm FS}\iff b_{\rm FS}^{-1}(b)\le a<a_c\iff 0<\alpha\le\alpha_{\rm FS}\,.$$

(iii) Solutions of (3), (7) and (8) are stable (in a sense defined below) among non-symmetric solutions, *i.e.*, solutions which explicitly depend on ω, if and only if the above condition on the parameters is satisfied. Such a condition has been introduced in Catrina and Wang [2001], but the sharp condition was established by V. Felli and M. Schneider in Felli and Schneider [2003], and this is why we use the notation $\Lambda_{\rm FS}$, $b_{\rm FS}$ and $\alpha_{\rm FS}$ (see Section 5.1). Notice that stability is a local property while our uniqueness (rigidity) results are global.

3.2 Optimal symmetry range in critical Caffarelli-Kohn-Nirenberg inequalities. The Caffarelli-Kohn-Nirenberg inequalities

$$\left(\int_{\mathbb R^d}\frac{|v|^p}{|x|^{b\,p}}\,dx\right)^{2/p}\le\mathsf C_{a,b}\int_{\mathbb R^d}\frac{|\nabla v|^2}{|x|^{2\,a}}\,dx\quad\forall\,v\in\mathfrak D_{a,b}\tag{9}$$

appear in Caffarelli, Kohn, and Nirenberg [1984], under the conditions that $a\le b\le a+1$ if $d\ge3$, $a<b\le a+1$ if $d=2$, $a+1/2<b\le a+1$ if $d=1$, and $a<a_c$ where the exponent

$$p=\frac{2\,d}{d-2+2\,(b-a)}$$

is determined by the invariance of the inequality under scalings. Here $\mathsf C_{a,b}$ denotes the optimal constant in (9) and the space $\mathfrak D_{a,b}$ is defined by

$$\mathfrak D_{a,b}:=\Big\{v\in\mathrm L^p\big(\mathbb R^d,|x|^{-b}\,dx\big)\;:\;|x|^{-a}\,|\nabla v|\in\mathrm L^2\big(\mathbb R^d,dx\big)\Big\}\,.$$

These inequalities were apparently introduced first by V.P. Il'in in Ilin [1961] but are more known as *Caffarelli-Kohn-Nirenberg inequalities*, according to Caffarelli, Kohn, and Nirenberg [1984]. Up to a scaling and a multiplication by a constant, any extremal function for the above inequality is a nonnegative solution of (7). It is therefore natural to ask whether $v_\star$ realizes the equality case in (9). Let

$$\mathsf C_{a,b}^\star:=\frac{\left(\int_{\mathbb R^d}\frac{|v_\star|^p}{|x|^{b\,p}}\,dx\right)^{2/p}}{\int_{\mathbb R^d}\frac{|\nabla v_\star|^2}{|x|^{2\,a}}\,dx}=\tfrac p2\,|\mathbb S^{d-1}|^{1-\frac2p}\,(a-a_c)^{1+\frac2p}\left(\frac{2\sqrt\pi\,\Gamma\big(\frac p{p-2}\big)}{(p-2)\,\Gamma\big(\frac{3\,p-2}{2\,(p-2)}\big)}\right)^{\frac{p-2}p}.$$

It was proved in Felli and Schneider [2003] that whenever $a<0$ and $b<b_{\rm FS}(a)$, the solutions of (7) are not radially symmetric: this is a *symmetry breaking* result, based on

the linear instability of $\mathcal{F}[v] := \mathsf{C}^\star_{a,b} \int_{\mathbb{R}^d} \frac{|\nabla v|^2}{|x|^{2a}}\,dx - \big(\int_{\mathbb{R}^d} \frac{|v|^p}{|x|^{b\,p}}\,dx\big)^{2/p}$ at $v = v_\star$. The main *symmetry* result of Dolbeault, Esteban, and Loss [2016b] is

Corollary 6. *Assume that* $d \geq 2$, $a < a_c$, *and* $b_{\rm FS}(a) \leq b \leq a+1$ *if* $a < 0$. *Then* $\mathsf{C}_{a,b} = \mathsf{C}^\star_{a,b}$ *and equality in* (9) *is achieved by a function* $v \in \mathfrak{D}_{a,b}$ *if and only if, up to a scaling and a multiplication by a constant,* $v = v_\star$.

In other words, whenever $\mathcal{F}[v]$ is linearly stable at $v = v_\star$, then $v_\star$ is a global extremal function for (9).

3.3 Sketch of the proof of Theorem 5.

The case $d = 2$ requires some specific estimates so we shall assume that $d \geq 3$ for simplicity. Let

$$u^{\frac{1}{2}-\frac{1}{n}} = w \quad \Longleftrightarrow \quad u = w^p \quad \text{with} \quad p = \frac{2\,n}{n-2}\,. \tag{10}$$

Up to a multiplicative constant, the right hand side in (9) is transformed into a generalized *Fisher information* functional

$$I[u] := \int_{\mathbb{R}^d} u\,|\mathsf{D}\mathsf{p}|^2\,d\mu \quad \text{where} \quad \mathsf{p} = \frac{m}{1-m}\,u^{m-1}\,. \tag{11}$$

Here $d\mu = |x|^{n-d}\,dx$, p is the *pressure function*, $\mathsf{D}\,\mathsf{p} := \big(\alpha\,\frac{\partial \mathsf{p}}{\partial r}, \frac{1}{r}\,\nabla_\omega \mathsf{p}\big)$, and $\mathsf{p}' = \frac{\partial \mathsf{p}}{\partial r}$ and $\nabla_\omega \mathsf{p}$ respectively denote the radial and the angular derivatives of p. The left hand side in (9) is now proportional to a *mass* integral, $\int_{\mathbb{R}^d} u\,d\mu$. In this section we consider the *critical case* and make the choice $m = 1 - 1/n$.

After these preliminaries, let us introduce the *fast diffusion* flow

$$\frac{\partial u}{\partial t} = \mathcal{L} u^m\,, \quad m = 1 - \frac{1}{n}\,, \tag{12}$$

where the operator $\mathcal{L}$, which has been considered in Theorem 5, is such that $\mathcal{L} w := -\mathsf{D}^*\,\mathsf{D}\,w$. The flow associated with (12) preserves the mass. At formal level, the key idea is to prove that $I[u(t,\cdot)]$ is decreasing w.r.t. t if u solves (12), and that the limit is $I[w_\star^p]$. A long computation indeed shows that, if u is a smooth solution of (12) with the appropriate behavior as $x \to 0$ and as $|x| \to +\infty$, then

$$\frac{d}{dt} I[u(t,\cdot)] \leq -2 \int_{\mathbb{R}^d} \mathsf{K}[\mathsf{p}(t,\cdot)]\,u(t,\cdot)^m\,d\mu$$

where, with $r = |x|$, we have

$$\mathsf{K}[\mathsf{p}] = \alpha^4 \left(1 - \frac{1}{n}\right) \left[\mathsf{p}'' - \frac{\mathsf{p}'}{r} - \frac{\Delta_\omega \mathsf{p}}{\alpha^2 (n-1) r^2}\right]^2 + 2\alpha^2 \frac{1}{r^2} \left|\nabla_\omega \mathsf{p}' - \frac{\nabla_\omega \mathsf{p}}{r}\right|^2 + (n-2)\big(\alpha_{\rm FS}^2 - \alpha^2\big) \frac{|\nabla_\omega \mathsf{p}|^2}{r^4} + \zeta_\star (n-d) \frac{|\nabla_\omega \mathsf{p}|^4}{r^4} \tag{13}$$

for some positive constant $\zeta_\star$. Hence, if $\alpha \le \alpha_{\rm FS}$, then $I[u(t,\cdot)]$ is nonincreasing along the flow of (12). However, regularity and decay estimates needed to justify such computations are not known yet and this parabolic approach is therefore formal. As in Section 2.2, we can instead rely on an elliptic method, which can be justified as follows.

If u_0 is a nonnegative critical point of I under mass constraint, then

$$0 = I'[u_0] \cdot \mathfrak{L} u_0^m = \frac{d I[u(t,\cdot)]}{dt}_{|t=0} \le -2 \int_{\mathbb{R}^d} \mathsf{K}[\mathsf{p}_0]\, \mathsf{u}_0^{1-n}\, d\mu$$

if u solves (12) with initial datum u_0. Here $I'[u_0]$ denotes the differential of I at u_0. With $\mathsf{p}_0 = \mathsf{p}(0,\cdot)$, this proves that $\nabla_\omega \mathsf{p}_0 = 0$: p_0 is radially symmetric. By solving $\mathsf{p}_0'' - \mathsf{p}_0'/r = 0$, we obtain that $\mathsf{p}_0(x) = a + b\,|x|^2$ for some constants $a, b \in \mathbb{R}^+$. The conclusion easily follows.

Proposition 7. *Let w be a nonnegative solution of (8) and $\mathsf{p} = (n-1)\, w^{-\frac{2}{n-2}}$. Under the assumptions of Theorem 5, if $\alpha \le \alpha_{\rm FS}$, then $\mathsf{K}[\mathsf{p}] = 0$.*

In practice, we prove that any solution of (5) on $\mathcal{C}$ has good decay properties as $s \to \pm\infty$, by delicate elliptic estimates, which rely on the fact that $p = 2n/(n-2) < 2d/(d-2)$ is a subcritical exponent on the d-dimensional manifold $\mathcal{C}$. This is enough to justify all integrations by parts and prove as a consequence that a nonnegative solution of (8) satisfies $\mathsf{K}[\mathsf{p}] = 0$: the conclusion follows as above. Notice that this amounts to test (8) by $\mathfrak{L} w^{2(n-1)/(n-2)}$.

4 Rigidity and sharp symmetry results in subcritical Caffarelli-Kohn-Nirenberg inequalities

In this section we consider a class of subcritical Caffarelli-Khon-Nirenberg inequalities and extend the results obtained for the critical case. Most results of this section have been published in Dolbeault, Esteban, Loss, and Muratori [2017], a joint paper of the authors with M. Muratori.

4.1 Subcritical Caffarelli-Kohn-Nirenberg inequalities. With the notation

$$\|w\|_{\mathrm{L}^{q,\gamma}(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |w|^q\,|x|^{-\gamma}\,dx\right)^{1/q}\,, \quad \|w\|_{\mathrm{L}^q(\mathbb{R}^d)} := \|w\|_{\mathrm{L}^{q,0}(\mathbb{R}^d)}\,,$$

we define $\mathrm{L}^{q,\gamma}(\mathbb{R}^d)$ as the space $\{w \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \{0\}) \,:\, \|w\|_{\mathrm{L}^{q,\gamma}(\mathbb{R}^d)} < \infty\}$. We shall work in the space $\mathrm{H}^p_{\beta,\gamma}(\mathbb{R}^d)$ of functions $w \in \mathrm{L}^{p+1,\gamma}(\mathbb{R}^d)$ such that $\nabla w \in \mathrm{L}^{2,\beta}(\mathbb{R}^d)$, which can also be defined as the completion of $\mathfrak{D}(\mathbb{R}^d \setminus \{0\})$ with respect to the norm

$$\|w\|^2 := (p_\star - p)\,\|w\|^2_{\mathrm{L}^{p+1,\gamma}(\mathbb{R}^d)} + \|\nabla w\|^2_{\mathrm{L}^{2,\beta}(\mathbb{R}^d)}\,.$$

Let us consider the family of subcritical *Caffarelli-Kohn-Nirenberg interpolation inequalities* that can be found in Caffarelli, Kohn, and Nirenberg [1984] and which is given by

$$\|w\|_{\mathrm{L}^{2p,\gamma}(\mathbb{R}^d)} \le \mathcal{C}_{\beta,\gamma,p}\,\|\nabla w\|^{\vartheta}_{\mathrm{L}^{2,\beta}(\mathbb{R}^d)}\,\|w\|^{1-\vartheta}_{\mathrm{L}^{p+1,\gamma}(\mathbb{R}^d)} \quad \forall\, w \in \mathrm{H}^p_{\beta,\gamma}(\mathbb{R}^d)\,. \tag{14}$$

Here the parameters β, γ and p are subject to the restrictions

$$d \ge 2\,, \quad \gamma - 2 < \beta < \frac{d-2}{d}\,\gamma\,, \quad \gamma \in (-\infty, d)\,, \quad p \in (1, p_\star] \tag{15}$$

with

$$p_\star := \frac{d-\gamma}{d-\beta-2} \quad \text{and} \quad \vartheta = \frac{(d-\gamma)\,(p-1)}{p\,\big(d+\beta+2-2\,\gamma-p\,(d-\beta-2)\big)}\,.$$

The critical case $p = p_\star$ determines $\vartheta = 1$ and has been dealt with in Section 3, so we shall focus on the subcritical case $p < p_\star$. Here by *critical* we simply mean that $\|w\|_{\mathrm{L}^{2p,\gamma}(\mathbb{R}^d)}$ scales like $\|\nabla w\|_{\mathrm{L}^{2,\beta}(\mathbb{R}^d)}$ and $\mathcal{C}_{\beta,\gamma,p}$ denotes the optimal constant in (14). The limit case $\beta = \gamma - 2$ and $p = 1$, which is an endpoint for (15), corresponds to Hardy-type inequalities: optimality is achieved among radial functions but there is no extremal function: see Dolbeault, Esteban, Loss, and Tarantello [2009]. The other endpoint is $\beta = (d-2)\,\gamma/d$, in which case $p_\star = d/(d-2)$: according to Catrina and Wang [2001] (also see Section 5.1), either $\gamma \ge 0$, symmetry holds and there exists a symmetric extremal function, or $\gamma < 0$, and then symmetry is broken but there is no extremal function. in all other cases, the existence of an extremal function for (14) follows from standard methods: see Catrina and Wang [2001], Dolbeault and Esteban [2012b], and Dolbeault, Muratori, and Nazaret [2017] for related results.

When $\beta = \gamma = 0$, (14) is a Gagliardo-Nirenberg interpolation inequality which is well known to be related to the fast diffusion equation $\frac{\partial u}{\partial t} = \Delta u^m$ in $\mathbb{R}^d$, not only for

$m = 1 - 1/d$ but also for any $m \in [1 - 1/d, 1)$. Here we generalize this observation to the weighted spaces.

Symmetry in (14) means that the equality case is achieved by Aubin-Talenti type functions

$$w_\star(x) = \left(1 + |x|^{2+\beta-\gamma}\right)^{-1/(p-1)} \quad \forall\, x \in \mathbb{R}^d\,.$$

On the contrary, there is *symmetry breaking* if this is not the case, because the equality case is then achieved by a non-radial extremal function. It has been proved in Bonforte, Dolbeault, Muratori, and Nazaret [2017] that *symmetry breaking* holds in (14) if

$$\gamma < 0 \quad \text{and} \quad \beta_{\mathrm{FS}}(\gamma) < \beta < \frac{d-2}{d}\,\gamma \tag{16}$$

where

$$\beta_{\mathrm{FS}}(\gamma) := d - 2 - \sqrt{(\gamma - d)^2 - 4\,(d-1)}\,.$$

Under Condition (15), *symmetry* holds in the complement of the set defined by (16).

Theorem 8. *Assume that* (15) *holds and that*

$$\beta \le \beta_{\mathrm{FS}}(\gamma) \quad \textit{if} \quad \gamma < 0\,. \tag{17}$$

Then the extremal functions for (14) *are radially symmetric and, up to a scaling and a multiplication by a constant, equal to* $w_\star$*.*

This means that (16) is the sharp condition for *symmetry breaking*.

4.2 A rigidity result. Up to a scaling and a multiplication by a constant, the Euler-Lagrange equation

$$-\,\mathrm{div}\left(|x|^{-\beta}\,\nabla w\right) = |x|^{-\gamma}\left(w^{2p-1} - w^p\right) \quad \text{in} \quad \mathbb{R}^d \setminus \{0\} \tag{18}$$

is satisfied by any extremal function for (14). In the range of parameters given by (15) and (17), our method establishes the symmetry of all positive solutions.

Theorem 9. *Assume that* (15) *and* (17) *hold. Then all positive solutions to* (18) *in* $\mathrm{H}^p_{\beta,\gamma}(\mathbb{R}^d)$ *are radially symmetric and, up to a scaling, equal to* $w_\star$*.*

This is again a *rigidity* result. Nonnegative solutions to (18) are actually positive by the standard Strong Maximum principle. Theorem 8 is therefore a consequence of Theorem 9.

4.3 Sketch of the proof of Theorem 9. Let us give an outline of the strategy of Dolbeault, Esteban, Loss, and Muratori [2017]. As in the critical case, Inequality (14) for a function w can be transformed by the change of variables

$$w(x) = v(r^\alpha, \omega)\,,$$

where $r = |x| \neq 0$ and $\omega = x/r$, in the new inequality

$$\left(\int_{\mathbb{R}^d} |v|^{2p}\, d\mu\right)^{\frac{1}{2p}} \le \mathcal{K}_{\alpha,n,p} \left(\int_{\mathbb{R}^d} |\mathsf{D}v|^2\, d\mu\right)^{\frac{\vartheta}{2}} \left(\int_{\mathbb{R}^d} |v|^{p+1}\, d\mu\right)^{\frac{1-\vartheta}{p+1}} \tag{19}$$

with $\mathcal{K}_{\alpha,n,p} = \alpha^{-\zeta}\, \mathcal{C}_{\beta,\gamma,p}$, $\zeta = \frac{\vartheta}{2} + \frac{1-\vartheta}{p+1} - \frac{1}{2\,p}$ and $d\mu = |x|^{n-d}\, dx$. The condition for the change of variables is

$$n = \frac{d-\beta-2}{\alpha} + 2 = \frac{d-\gamma}{\alpha}\,,$$

which reflects the fact that the weights are all the same in (19). It is solved by

$$\alpha = 1 + \frac{\beta-\gamma}{2} \quad \text{and} \quad n = 2\,\frac{d-\gamma}{\beta+2-\gamma}\,.$$

Inequality (19) is a Caffarelli-Kohn-Nirenberg inequality with weight $|x|^{n-d}$ in all terms, and $\mathsf{D}v := \left(\alpha\,\frac{\partial v}{\partial s}, \frac{1}{s}\,\nabla_\omega v\right)$. Notice that $p_\star = \frac{n}{n-2}$, so that $2\,p_\star$ is the critical Sobolev exponent associated with the *fractional dimension* n considered in (10).

With a generalized *Fisher information* I and the *pressure function* p defined by (11), we consider the *subcritical range* $m_1 := 1 - 1/n < m < 1$. If u is smooth solution of (12) with sufficient decay properties, we obtain that I evolves according to

$$\frac{d}{dt} I[u(t,\cdot)] = -2 \int_{\mathbb{R}^d} \mathcal{R}[\mathsf{p}(t,\cdot)]\, u(t,\cdot)^m\, d\mu \;\text{ with }\; \mathcal{R}[\mathsf{p}] := \mathsf{K}[\mathsf{p}] + (m-m_1)\,(\mathcal{L}\mathsf{p})^2,$$

where K is given by (13). We recover the result of the critical case of Section 3.3 by taking the limit as $m \to m_1$.

Inspired by tools of Information Theory and Savaré and Toscani [2014], Toscani [2014], and Dolbeault and Toscani [2016], we introduce the generalized *Rényi entropy power* functional

$$\mathcal{F}[u] := \left(\int_{\mathbb{R}^d} u^m\, d\mu\right)^\sigma \quad \text{with} \quad \sigma = \frac{2}{n}\,\frac{1}{1-m} - 1 > 1$$

and observe that $\mathcal{F}''$ has the sign of $-\mathcal{H}[u(t,\cdot)]$ where

$$\mathcal{H}[u] := (m - m_1) \int_{\mathbb{R}^d} \left| \mathfrak{L}\mathsf{p} - \frac{\int_{\mathbb{R}^d} u\, |\mathsf{Dp}|^2\, u^m\, d\mu}{\int_{\mathbb{R}^d} u^m\, d\mu} \right|^2 d\mu + \int_{\mathbb{R}^d} \mathfrak{R}[\mathsf{p}]\, u^m\, d\mu\,.$$

Here $\mathcal{F}'$ denotes the derivative with respect to t of $\mathcal{F}[u(t,\cdot)]$. The computation requires many integrations by parts. The fact that boundary terms do not contribute can be justified if u is a nonnegative critical point, *i.e.*, a minimizer of $\mathcal{F}'$ under mass constraint. Indeed, the minimization of

$$\left(\int_{\mathbb{R}^d} v^{p+1}\, d\mu \right)^{\sigma-1} \int_{\mathbb{R}^d} |\mathsf{D}v|^2\, d\mu \quad \text{with} \quad v = u^{m-1/2}$$

under the constraint that $\int_{\mathbb{R}^d} u\, d\mu = \int_{\mathbb{R}^d} v^{2p}\, d\mu$ takes a given positive value is equivalent to the *Caffarelli-Kohn-Nirenberg interpolation inequalities* (14).

To make the argument rigorous, we can argue as in Section 3.3 by taking u as initial datum and performing the computation of $\mathcal{F}''$ at $t = 0$ only. In other words, we are simply testing the Euler-Lagrange equation satisfied by u with $\mathfrak{L}u^m$. By elliptic regularity (the estimates are as delicate as in the critical case and we refer to Dolbeault, Esteban, Loss, and Muratori [2017] for details), we have enough estimates to prove that $\mathcal{H}[u] = 0$ and deduce that $\mathsf{p}(x) = \mathsf{a} + \mathsf{b}\, |x|^2$ for some real constants a and b.

4.4 Considerations on the optimality of the method. The *symmetry breaking* condition in (9) and (14) has been established by proving the linear instability of radial critical points, in Felli and Schneider [2003] and Bonforte, Dolbeault, Muratori, and Nazaret [2017] respectively. This amounts to a spectral gap condition in a Hardy-Poincaré inequality: see Bonforte, Dolbeault, Muratori, and Nazaret [ibid.] for details. It is remarkable that the symmetry holds whenever radial critical points are linearly stable and this deserves an explanation. The solution of (12) is attracted by self-similar Barenblatt functions as $t \to +\infty$. Since these Barenblatt functions are precisely the radial critical points of our variational problem, the asymptotic rate of convergence is determined by the previous spectral gap, in self-similar variables. It can be checked that the condition that appears in the *carré du champ* method, which amounts to prove that a quadratic form has a sign, is the same in the asymptotic regime as $t \to +\infty$ as the quadratic form which is used to check symmetry breaking. Hence either symmetry breaking occurs, or the *carré du champ* method shows that the *Rényi entropy power* functional is monotone non-increasing, at least in the asymptotic regime: see Dolbeault, Esteban, and Loss [2016a] for details. To conclude in the critical case, it is enough to observe that all terms in the expression of $\mathsf{K}[\mathsf{p}]$ in (13) are quadratic, except the last one, which has a sign and is negligible compared to the others in the asymptotic regime: the sign condition for $\mathsf{K}[\mathsf{p}]$ away from the asymptotic regime is the same as when $t \to +\infty$. This explains why our method for proving *symmetry* gives the optimal range in the critical case. In the subcritical regime, a similar observation can also be done.

5 Bifurcations and symmetry breaking

The results of this section are taken mostly from Dolbeault, Esteban, Tarantello, and Tertikas [2011], Dolbeault and Esteban [2012a], and Dolbeault and Esteban [2014].

5.1 Rigidity and bifurcations. Let us come back to the critical Caffarelli-Kohn-Nirenberg inequality and consider the Emden-Fowler transformation (6). As noted in Catrina and Wang [2001], Inequality (9) is transformed into the Gagliardo-Nirenberg-Sobolev inequality

$$\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathcal{C})} + \Lambda\,\|\varphi\|^2_{\mathrm{L}^2(\mathcal{C})} \geq \mu(\Lambda)\,\|\varphi\|^2_{\mathrm{L}^p(\mathcal{C})} \quad \forall\,\varphi \in \mathrm{H}^1(\mathcal{C})$$

where $\mu(\Lambda) = \mathsf{C}_{a,b}^{-1}\,\left|\mathbb{S}^{d-1}\right|^{1-2/p}$. Here $\mathcal{C} := \mathbb{R}\times\mathbb{S}^{d-1}$ is a cylinder and, as in Section 2, we adopt the convention that the measure on the sphere is the uniform probability measure. The extremal functions are, up to multiplication by a constant, and dilation, solutions of (5).

If we restrict the study to symmetric functions, that is, $v(r) = r^{a-a_c}\,\varphi(-\log r)$ with $r = |x|$, then the inequality degenerates into the simple Gagliardo-Nirenberg-Sobolev inequality

$$\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathbb{R})} + \Lambda\,\|\varphi\|^2_{\mathrm{L}^2(\mathbb{R})} \geq \mu_\star(\Lambda)\,\|\varphi\|^2_{\mathrm{L}^p(\mathbb{R})} \quad \forall\,\varphi \in \mathrm{H}^1(\mathbb{R})\,.$$

Here we denote by

$$\mu_\star(\Lambda) = \mu_\star(1)\,\Lambda^{\frac{p+2}{2\,p}}$$

the optimal constant and notice that $\varphi_\star(s) = \left(\frac{1}{2}\,p\,\Lambda\,\cosh\left(\frac{p-2}{2}\,\sqrt{\Lambda}\,s\right)^{-2}\right)^{1/(p-2)}$ is an optimal function, which is the unique solution of $-\varphi'' + \Lambda\,\varphi = |\varphi|^{p-2}\,\varphi$ on $\mathbb{R}$, up to translations. With this notation, we have $\mu_\star(\Lambda) = \|\varphi_\star\|^{p-2}_{\mathrm{L}^p(\mathbb{R})}$. If we linearize

$$\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathcal{C})} + \Lambda\,\|\varphi\|^2_{\mathrm{L}^2(\mathcal{C})} - \mu_\star(\Lambda)\,\|\varphi\|^2_{\mathrm{L}^p(\mathcal{C})}$$

around $\varphi = \varphi_\star$, V. Felli and M. Schneider found in Felli and Schneider [2003] that the lowest eigenvalue of the quadratic form, that is, the lowest positive eigenvalue of the Pöschl-Teller operator $-\frac{d^2}{ds^2} + \Lambda + d - 1 - (p-1)\,\varphi_\star^{p-2}$, is given by $\lambda_1(\Lambda) = -\frac{1}{4}\,(p^2 -$

4) $(\Lambda - \Lambda_{\mathrm{FS}})$, so that $\lambda_1(\Lambda_{\mathrm{FS}}) < 0$ if and only if

$$\Lambda > \Lambda_{\mathrm{FS}} := 4\,\frac{d-1}{p^2-4}\,.$$

See Lifchitz and Landau [1966, p. 74] for details. This condition is the *symmetry breaking* condition of Theorem 3. The branch of non-radial solutions bifurcating from $\Lambda = \Lambda_{\mathrm{FS}}$ has been computed numerically in Dolbeault and Esteban [2012a] and an example is shown in Figure 1. By construction, we know that $\Lambda \mapsto \mu(\Lambda)$ is increasing, concave, and we read from Theorem 3 that the non-symmetric branch bifurcates from $\Lambda = \Lambda_{\mathrm{FS}}$, and is such that $\mu(\Lambda) < \mu_\star(\Lambda)$ if $\Lambda > \Lambda_{\mathrm{FS}}$. This simple scenario explains the symmetry and symmetry breaking properties in (9), but is not generic as we shall see next in the case of more complicated interpolation inequalities.

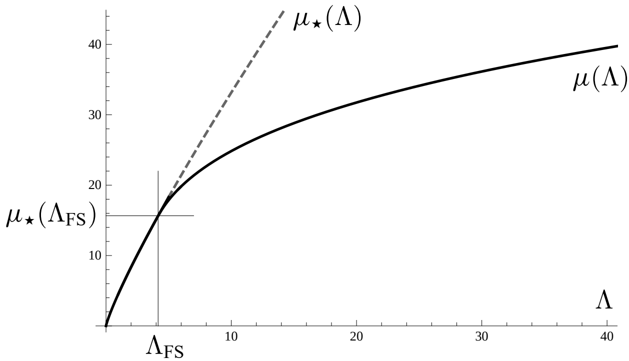


Figure 1: Branches for $p = 2.8$, $d = 5$, $\theta = 1$.

5.2 Bifurcations, reparametrization and turning points. Let us consider the interpolation inequality

$$\left(\int_{\mathbb{R}^d} \frac{|u|^p}{|x|^{bp}}\,dx\right)^{\frac{2}{p}} \le \mathsf{C}_{a,b,\theta} \left(\int_{\mathbb{R}^d} \frac{|\nabla u|^2}{|x|^{2a}}\,dx\right)^{\theta} \left(\int_{\mathbb{R}^d} \frac{|u|^2}{|x|^{2\,(a+1)}}\,dx\right)^{1-\theta} \tag{20}$$

with $d \ge 1$, $p \in (2, 2^*)$ or $p = 2^*$ if $d \ge 3$, and $\theta \in (\vartheta(p), 1]$ with $\vartheta(p) := d\,\frac{p-2}{2\,p}$. The scaling invariance imposes $p = 2\,d/\big(d-2+2\,(b-a)\big)$. As proved in Caffarelli, Kohn, and Nirenberg [1984], the above inequalities hold with a finite constant $\mathsf{C}_{a,b,\theta}$ if

$a < a_c = (d-2)/2$, and $b \in (a+1/2, a+1]$ when $d = 1$, $b \in (a, a+1]$ when $d = 2$ and $b \in [a, a+1]$ when $d \geq 3$. Moreover, there exist extremal functions for the inequalities (20) for any $p \in (2, 2^*)$ and $\theta \in (\vartheta(p), 1)$ or $\theta = \vartheta(p)$ and $d \geq 2$, with $a_c - a > 0$ not too large. On the contrary equality is never achieved for $p = 2$, or $a < 0$, $p = 2^*$ and $d \geq 3$, or $d = 1$ and $\theta = \vartheta(p, 1)$. The existence of extremal functions has been studied in Dolbeault and Esteban [2012b]. We may notice that

$$0 \leq \vartheta(p) \leq \theta < 1 \quad \Longleftrightarrow \quad 2 \leq p \leq p^*(d, \theta) := \frac{2\,d}{d - 2\,\theta} < 2^*\,.$$

With the same conventions as in the previous subsection, the Emden-Fowler change of variables (6) transforms (20) into the Gagliardo-Nirenberg-Sobolev inequality

$$\left(\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathcal{C})} + \Lambda\,\|\varphi\|^2_{\mathrm{L}^2(\mathcal{C})}\right)^\theta\,\|\varphi\|^{2(1-\theta)}_{\mathrm{L}^2(\mathcal{C})} \geq \mu(\theta, \Lambda)\,\|\varphi\|^2_{\mathrm{L}^p(\mathcal{C})} \quad \forall\,\varphi \in \mathrm{H}^1(\mathcal{C}) \tag{21}$$

on $\mathcal{C} := \mathbb{R}\times\mathbb{S}^{d-1}$, with $\Lambda = (a - a_c)^2$ and $\mu(\theta, \Lambda) = \mathrm{C}^{-1}_{a,b,\theta}\,\left|\mathbb{S}^{d-1}\right|^{1-2/p}$. Of course, the case $\theta = 1$ corresponds to the critical case and, consistently, we write $\mu(1, \Lambda) = \mu(\Lambda)$.

For $\theta < 1$, the Euler-Lagrange equation of an extremal function on $\mathcal{C}$ is

$$-\,\Delta\varphi + \frac{1}{\theta}\left((1-\theta)\,\frac{\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathcal{C})}}{\|\varphi\|^2_{\mathrm{L}^2(\mathcal{C})}} + \Lambda\right)\varphi - \frac{\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathcal{C})} + \Lambda\,\|\varphi\|^2_{\mathrm{L}^2(\mathcal{C})}}{\theta\,\|\varphi\|^p_{\mathrm{L}^p(\mathcal{C})}}\,\varphi^{p-1} = 0\,. \tag{22}$$

Up to the reparametrization

$$\Lambda \mapsto \lambda = \frac{1}{\theta}\Big[(1-\theta)\,t[\varphi] + \Lambda\Big] \quad \text{where} \quad t[\varphi] := \frac{\|\nabla\varphi\|^2_{\mathrm{L}^2(\mathcal{C})}}{\|\varphi\|^2_{\mathrm{L}^2(\mathcal{C})}}$$

and a multiplication by a constant, an extremal function φ for (21) solves (5). In other words, we can use the set of solutions in the critical case $\theta = 1$ to parametrize the solutions corresponding to $\theta < 1$.

Let us start with the symmetric functions. With an evident notation, we define $\mu_\star(\theta, \Lambda)$ as the optimal constant in the inequality corresponding to (21) restricted to symmetric functions, *i.e.*, functions depending only on $s \in \mathbb{R}$. If we denote by $\varphi_{\star,\lambda}$ the function

$$\varphi_{\star,\lambda}(s) = \left(\frac{1}{2}\,p\,\lambda\,\cosh\left(\tfrac{p-2}{2}\,\sqrt{\lambda}\,s\right)^{-2}\right)^{\frac{1}{p-2}}$$

for any $\lambda > 0$, then $t[\varphi_{\star,\lambda}]$ is explicit and we can parametrize the set $\left\{\big(\Lambda, \mu_\star(\theta, \Lambda)\big)\,:\,\Lambda > 0\right\}$ by $\left\{\big(\theta\,\lambda - (1-\theta)\,t[\varphi_{\star,\lambda}], \mu_\star(\lambda)\big)\,:\,\lambda > 0\right\}$. It turns out that the equation

$\Lambda = \theta\,\lambda - (1-\theta)\,t[\varphi_{\star,\lambda}]$ can be inverted, which allows us to obtain $\lambda = \Lambda_*^\theta(\Lambda)$ and get an explicit expression for

$$\mu_\star(\theta,\Lambda) = \mu_\star\big(\Lambda_*^\theta(\Lambda)\big) = \mu_\star\big(\Lambda_*^\theta(1)\big)\,\Lambda^{\theta-\frac{p-2}{2p}}\,.$$

According to del Pino, Dolbeault, Filippas, and Tertikas [2010], a Taylor expansion around $\varphi_{\star,\Lambda_{\mathrm{FS}}}$ shows that for any $\Lambda > \Lambda_{\mathrm{FS}}^\theta$, where

$$\Lambda_{\mathrm{FS}}^\theta := \theta\,\mu_{\mathrm{FS}} - (1-\theta)\,t[u_{\mathrm{FS}}]\,,$$

the function $\varphi_{\star,\lambda}$ with $\lambda = \Lambda_*^\theta(\Lambda)$ is linearly unstable, so that $\mu(\theta,\Lambda) < \mu_\star(\theta,\Lambda)$.

The case of non-symmetric functions is more subtle because we do not know the exact multiplicity of the solutions of (5) in the symmetry breaking range. There is a branch of non-symmetric solutions of (22) which bifurcates from the branch of symmetric solutions at $\Lambda = \Lambda_{\mathrm{FS}}^\theta$. This branch has been computed numerically in Dolbeault and Esteban [2012a] and a formal asymptotic expansion was performed in a neighborhood of the bifurcation point in Dolbeault and Esteban [2014]. Because of the reparametrization of the solutions of (22) by the solutions of (5), we can use the branch $\lambda \mapsto \varphi_\lambda$ of non-symmetric extremal functions for $\lambda > \Lambda_{\mathrm{FS}}$ to get an upper bound of $\mu(\theta,\Lambda)$:

$$\mu(\theta,\Lambda) \le \mu(\lambda) \quad \text{for any } \lambda > \Lambda_{\mathrm{FS}} \text{ such that} \quad \Lambda = \theta\,\lambda - (1-\theta)\,t[\varphi_\lambda]\,.$$

Actually, we deduce from the branch $\lambda \mapsto \varphi_\lambda$ of non-symmetric extremal functions an entire branch of non-symmetric solutions of (22) which is parametrized by λ and deduce a parametric curve $\mathfrak{B} := \big\{\big(\Lambda(\lambda) := \theta\,\lambda - (1-\theta)\,t[\varphi_\lambda], \mu(\lambda)\big)\ :\ \lambda > \Lambda_{\mathrm{FS}}\big\}$ which can be used to bound $\mu(\theta,\Lambda)$ from above. If $(\Lambda,\mu) \in \mathfrak{B}$, we have no proof that φ_λ is optimal if $\mu(\lambda) < \mu_\star(\theta,\Lambda)$, but at least we know that

$$\mu(\lambda) = \Big(\|\nabla\varphi_\lambda\|^2_{\mathrm{L}^2(\mathcal{C})} + \Lambda(\lambda)\,\|\varphi_\lambda\|^2_{\mathrm{L}^2(\mathcal{C})}\Big)^\theta\ \|\varphi_\lambda\|^{2(1-\theta)}_{\mathrm{L}^2(\mathcal{C})}\ \|\varphi_\lambda\|^{-2}_{\mathrm{L}^p(\mathcal{C})}\,.$$

Some numerical results are shown in Figure 2.

The formal asymptotic expansion of Dolbeault and Esteban [ibid.] suggests that there are only two possible generic scenarii:

(i) Either the curve $\mathfrak{B}$ bifurcates to the right, that is, $\mathfrak{B}$ is included in the region $\Lambda \ge \Lambda_{\mathrm{FS}}^\theta$, and $\Lambda \mapsto \mu(\theta,\Lambda)$ is qualitatively expected to be as in Figure 1. We know that this is what happens for $\theta = 1$ and expect a similar behavior for any θ close enough to 1. In this case, the region of symmetry breaking is characterized by the linear instability of the symmetric optimal functions.

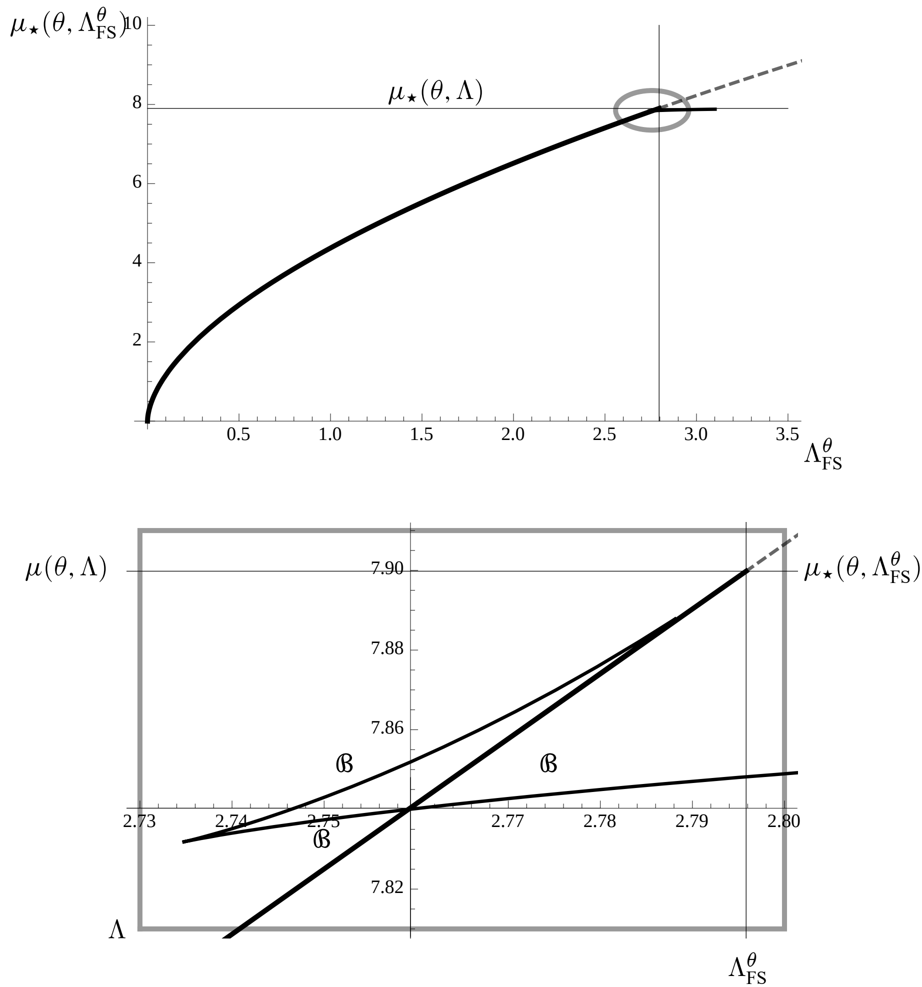


Figure 2: Branches for $p = 2.8$, $d = 5$, $\theta = 0.718$. Left: the bifurcation point $(\Lambda^\theta_{\rm FS}, \mu_\star(\theta,\Lambda^\theta_{\rm FS})$ is at the intersection of the horizontal and vertical lines. The area enclosed in the small ellipse is enlarged in the right plot: the branch has a *turning point* and $\mu(\theta,\Lambda^\theta_{\rm FS}) < \mu_\star(\theta,\Lambda^\theta_{\rm FS})$.

(ii) Or the curve $\mathfrak{B}$ bifurcates to the left. For $\lambda - \Lambda_{\rm FS} > 0$, small enough, the curve $\lambda \mapsto \big(\Lambda(\lambda), \mu(\lambda)\big)$ satisfies $\Lambda(\lambda) < \Lambda^\theta_{\rm FS}$ and $\mu(\lambda) > \mu_\star(\theta,\Lambda(\lambda))$. In that case, the region of symmetry breaking does not seem to be characterized by the linear instability of the symmetric optimal functions and we numerically observe a turning point as in Figure 2 (right).

In Dolbeault, Esteban, Filippas, and Tertikas [2015], *a priori* estimates for branches with $\theta < 1$ were deduced from the known symmetry results (later improved in Dolbeault, Esteban, and Loss [2016b]). This further constrains $\mathfrak{B}$ and the symmetry breaking region and determines a lower bound for the value of Λ corresponding to a turning point of the branch. There are many open questions concerning $\mathfrak{B}$ and the set of extremal functions when $\theta < 1$, but at least we can prove that the symmetry breaking range does not always coincide with the region of linear instability of symmetric optimal functions.

5.3 Symmetry breaking and energy considerations. The exponent $\vartheta(p)$ is the exponent which appears in the Gagliardo-Nirenberg inequality

$$\|\nabla u\|_{\mathrm{L}^2(\mathbb{R}^d)}^{2\vartheta(p)}\,\|u\|_{\mathrm{L}^2(\mathbb{R}^d)}^{2(1-\vartheta(p))} \ge \mathsf{C}_{\mathrm{GN}}(p)\,\|u\|_{\mathrm{L}^p(\mathbb{R}^d)}^2 \quad \forall\, u \in \mathrm{H}^1(\mathbb{R}^d)\,. \tag{23}$$

By considering an extremal function for this inequality and translations, for any $p \in (2, 2^*)$, one can check that

$$\mu(\vartheta(p), \Lambda) \le \mathsf{C}_{\mathrm{GN}}(p) \quad \forall\, \Lambda > 0\,.$$

Lemma 10. *Let $d \ge 2$. For any $p \in (2, 2^*)$, if $\mathsf{C}_{\mathrm{GN}}(p) < \mu_\star(\vartheta(p), \Lambda_{\mathrm{FS}}^{\vartheta(p)})$, there exists $\Lambda_s \in (0, \Lambda_{\mathrm{FS}}^{\vartheta(p)})$ such that $\mu(\vartheta(p), \Lambda) = \mu_\star(\vartheta(p), \Lambda)$ if and only if $\Lambda \in (0, \Lambda_s]$.*

The fact that the symmetry range is an interval of the form $(0, \Lambda_s]$ can be deduced from a scaling argument: see Dolbeault, Esteban, Loss, and Tarantello [2009] and Dolbeault, Esteban, Tarantello, and Tertikas [2011] for details. The result is otherwise straightforward but difficult to use because the value of $C_{\mathrm{GN}}(p)$ is not known explicitly. From a numerical point of view, it gives a simple criterion, which has been implemented in Dolbeault, Esteban, Tarantello, and Tertikas [2011]. Moreover, in Dolbeault and Esteban [2014], it has been observed numerically that the condition $\mathsf{C}_{\mathrm{GN}}(p) < \mu_\star(\vartheta(p), \Lambda_{\mathrm{FS}}^{\vartheta(p)})$ is equivalent to a *bifurcation to the left* as in Figure 2.

For θ and $p-2$ small enough, the assumption of Lemma 10 holds. Let us consider the Gaussian test function $\mathsf{g}(x) := (2\,\pi)^{-d/4}\,\exp(-|x|^2/4)$ in (20) and consider

$$h(p) := \frac{\|\nabla \mathsf{g}\|_{\mathrm{L}^2(\mathbb{R}^d)}^{2\theta}\,\|\mathsf{g}\|_{\mathrm{L}^2(\mathbb{R}^d)}^{2(1-\theta)}}{\|\mathsf{g}\|_{\mathrm{L}^p(\mathbb{R}^d)}^2}\,\frac{1}{\mu_\star(\theta, \Lambda_{\mathrm{FS}}^\theta)} \quad \text{with} \quad \theta = \vartheta(p)\,.$$

A computation shows that $\lim_{p\to 2_+} h(p) = 1$ and $\lim_{p\to 2_+} \frac{dh}{dp}(p) < 0$. For $p - 2 > 0$, small enough, we obtain that

$$\mathsf{C}_{\mathrm{GN}}(p) \le h(p) < \mu_\star(\theta, \Lambda_{\mathrm{FS}}^\theta)\,.$$

A perturbation argument has been used in Dolbeault, Esteban, Tarantello, and Tertikas [2011] to establish the following result.

Theorem 11. *Let $d \ge 2$. There exists $\eta > 0$ such that for any $p \in (2, 2+\eta)$,*

$$\mu(\theta,\Lambda) < \mu_\star(\theta,\Lambda) \quad \textit{if} \quad \Lambda_{\mathrm{FS}}^\theta - \eta < \Lambda < \Lambda_{\mathrm{FS}}^\theta \quad \textit{and} \quad \vartheta(p) < \theta < \vartheta(p) + \eta\,.$$

5.4 An open question. The criterion considered in Lemma 10 is based on energy considerations and provides only a sufficient condition for symmetry breaking. It is difficult to check it in practice, except in asymptotic regimes of the parameters. The formal expansions of the branch near the bifurcation points are based on a purely local analysis, and suggest another criterion: either the branch bifurcates to the right and the symmetry breaking range is characterized by the linear instability of the symmetric optimal functions, or the branch bifurcates to the left, and this is not anymore the case. Is such an observation, which has been made numerically only for some specific values of p, true in general? This seems to be true when θ is close enough to $\vartheta(p)$ and at least in this regime we can conjecture that *the symmetry breaking range is not characterized by the linear instability of the symmetric optimal functions if and only if the branch bifurcates to the left.*

An additional question, which corresponds to a limiting case, goes as follows. If $\theta = \vartheta(p)$, is the range of symmetry determined exactly by the value of the optimal constant in (23), when it is below $\mu_\star(\theta, \Lambda_{\mathrm{FS}}^\theta)$? Numerically, this is supported by the fact that, in this case, the curve $\mathfrak{B}$ is monotone increasing as a function of Λ.

In the study of the symmetry issue in (9) and (14), the key tool is the nonlinear flow, which extends a local result (linear stability) to a global result (rigidity). A similar tool would be needed to answer the conjecture. In the case $\theta = \vartheta(p)$, it would be crucial to obtain a variational characterization of the non-symmetric solutions in the curve of non-symmetric functions $\mathfrak{B}$ and a uniqueness result for any given Λ.

References

Dominique Bakry and Michel Émery (1984). “Hypercontractivité de semi-groupes de diffusion”. *C. R. Acad. Sci. Paris Sér. I Math.* 299.15, pp. 775–778. MR: 772092 (cit. on p. 2282).

Dominique Bakry and Michel Émery (1985a). "Diffusions hypercontractives". In: *Séminaire de probabilités, XIX, 1983/84*. Vol. 1123. Lecture Notes in Math. Springer, Berlin, pp. 177–206. MR: 889476 (cit. on p. 2282).

– (1985b). "Inégalités de Sobolev pour un semi-groupe symétrique". *C. R. Acad. Sci. Paris Sér. I Math.* 301.8, pp. 411–413. MR: 808640 (cit. on pp. 2282, 2283).

Dominique Bakry, Ivan Gentil, and Michel Ledoux (2014). *Analysis and geometry of Markov diffusion operators*. Vol. 348. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, pp. xx+552. MR: 3155209 (cit. on p. 2285).

Dominique Bakry and M. Ledoux (1996). "Sobolev inequalities and Myers's diameter theorem for an abstract Markov generator". *Duke Math. J.* 85.1, pp. 253–270. MR: 1412446 (cit. on pp. 2282, 2284).

William Beckner (1993). "Sharp Sobolev inequalities on the sphere and the Moser-Trudinger inequality". *Ann. of Math. (2)* 138.1, pp. 213–242. MR: 1230930 (cit. on pp. 2282, 2284).

Abdellatif Bentaleb (1993). "Inégalité de Sobolev pour l'opérateur ultrasphérique". *C. R. Acad. Sci. Paris Sér. I Math.* 317.2, pp. 187–190. MR: 1231419 (cit. on p. 2282).

Marie-Françoise Bidaut-Véron and Laurent Véron (1991). "Nonlinear elliptic equations on compact Riemannian manifolds and asymptotics of Emden equations". *Invent. Math.* 106.3, pp. 489–539. MR: 1134481 (cit. on pp. 2282, 2284).

Matteo Bonforte, Jean Dolbeault, Matteo Muratori, and Bruno Nazaret (2017). "Weighted fast diffusion equations (Part I): Sharp asymptotic rates without symmetry and symmetry breaking in Caffarelli-Kohn-Nirenberg inequalities". *Kinet. Relat. Models* 10.1, pp. 33–59. MR: 3579563 (cit. on pp. 2292, 2294).

L. Caffarelli, R. Kohn, and L. Nirenberg (1984). "First order interpolation inequalities with weights". *Compositio Math.* 53.3, pp. 259–275. MR: 768824 (cit. on pp. 2288, 2291, 2296).

Florin Catrina and Zhi-Qiang Wang (2001). "On the Caffarelli-Kohn-Nirenberg inequalities: sharp constants, existence (and nonexistence), and symmetry of extremal functions". *Comm. Pure Appl. Math.* 54.2, pp. 229–258. MR: 1794994 (cit. on pp. 2288, 2291, 2295).

Pierre Curie (1894). "Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique". *Journal de physique théorique et appliquée* 3.1, pp. 393–415 (cit. on p. 2280).

Jérôme Demange (2008). "Improved Gagliardo-Nirenberg-Sobolev inequalities on manifolds with positive curvature". *J. Funct. Anal.* 254.3, pp. 593–611. MR: 2381156 (cit. on pp. 2282, 2284).

J. Dolbeault and M. J. Esteban (2012a). "A scenario for symmetry breaking in Caffarelli-Kohn-Nirenberg inequalities". *J. Numer. Math.* 20.3-4, pp. 233–249. MR: 3043639 (cit. on pp. 2295, 2296, 2298).

Jean Dolbeault and Maria J Esteban (2014). "Branches of non-symmetric critical points and symmetry breaking in nonlinear elliptic partial differential equations". *Nonlinearity* 27.3, p. 435 (cit. on pp. 2295, 2298, 2300).

Jean Dolbeault and Maria J. Esteban (2012b). "Extremal functions for Caffarelli-Kohn-Nirenberg and logarithmic Hardy inequalities". *Proc. Roy. Soc. Edinburgh Sect. A* 142.4, pp. 745–767. MR: 2966111 (cit. on pp. 2291, 2297).

Jean Dolbeault, Maria J. Esteban, Stathis Filippas, and Achilles Tertikas (2015). "Rigidity results with applications to best constants and symmetry of Caffarelli-Kohn-Nirenberg and logarithmic Hardy inequalities". *Calc. Var. Partial Differential Equations* 54.3, pp. 2465–2481. MR: 3412380 (cit. on p. 2300).

Jean Dolbeault, Maria J. Esteban, and Gaspard Jankowiak (2017). "Onofri inequalities and rigidity results". *Discrete Contin. Dyn. Syst.* 37.6, pp. 3059–3078. MR: 3622074 (cit. on p. 2285).

Jean Dolbeault, Maria J. Esteban, and Michael Kowalczyk Michałand Loss (2014). "Improved interpolation inequalities on the sphere". *Discrete Contin. Dyn. Syst. Ser. S* 7.4, pp. 695–724. MR: 3177759 (cit. on p. 2284).

Jean Dolbeault, Maria J. Esteban, Michal Kowalczyk, and Michael Loss (2013). "Sharp interpolation inequalities on the sphere: new methods and consequences". *Chin. Ann. Math. Ser. B* 34.1, pp. 99–112. MR: 3011461 (cit. on p. 2284).

Jean Dolbeault, Maria J. Esteban, Ari Laptev, and Michael Loss (2014). "One-dimensional Gagliardo-Nirenberg-Sobolev inequalities: remarks on duality and flows". *J. Lond. Math. Soc. (2)* 90.2, pp. 525–550. MR: 3263963 (cit. on p. 2285).

Jean Dolbeault, Maria J. Esteban, and Michael Loss (2014). "Nonlinear flows and rigidity results on compact manifolds". *J. Funct. Anal.* 267.5, pp. 1338–1363. MR: 3229793 (cit. on pp. 2284, 2285).

– (2016a). "Interpolation inequalities, nonlinear flows, boundary terms, optimality and linearization". *J. Elliptic Parabol. Equ.* 2.1-2, pp. 267–295. MR: 3645948 (cit. on p. 2294).

– (2016b). "Rigidity versus symmetry breaking via nonlinear flows on cylinders and Euclidean spaces". *Invent. Math.* 206.2, pp. 397–440. MR: 3570296 (cit. on pp. 2286, 2289, 2300).

– (2017). "Interpolation inequalities on the sphere: linear vs. nonlinear flows". *Ann. Fac. Sci. Toulouse Math. (6)* 26.2, pp. 351–379. MR: 3640894 (cit. on p. 2283).

Jean Dolbeault, Maria J. Esteban, Michael Loss, and Matteo Muratori (2017). "Symmetry for extremal functions in subcritical Caffarelli-Kohn-Nirenberg inequalities". *C. R.*

Math. Acad. Sci. Paris 355.2, pp. 133–154. MR: 3612700 (cit. on pp. 2290, 2293, 2294).

Jean Dolbeault, Maria J. Esteban, Michael Loss, and Gabriella Tarantello (2009). "On the symmetry of extremals for the Caffarelli–Kohn–Nirenberg inequalities". *Adv. Nonlinear Stud.* 9.4, pp. 713–726. MR: 2560127 (cit. on pp. 2291, 2300).

Jean Dolbeault, Maria J. Esteban, Gabriella Tarantello, and Achilles Tertikas (2011). "Radial symmetry and symmetry breaking for some interpolation inequalities". *Calc. Var. Partial Differential Equations* 42.3-4, pp. 461–485. MR: 2846263 (cit. on pp. 2295, 2300, 2301).

Jean Dolbeault and Michal Kowalczyk (2017). "Uniqueness and rigidity in nonlinear elliptic equations, interpolation inequalities and spectral estimates". To appear in Annales de la Faculté de Sciences de Toulouse, Mathématiques, arXiv: 1412.2915 (cit. on p. 2285).

Jean Dolbeault, Matteo Muratori, and Bruno Nazaret (2017). "Weighted interpolation inequalities: a perturbation approach". *Math. Ann.* 369.3-4, pp. 1237–1270. MR: 3713540 (cit. on p. 2291).

Jean Dolbeault and Giuseppe Toscani (2016). "Nonlinear diffusions: extremal properties of Barenblatt profiles, best matching and delays". *Nonlinear Anal.* 138, pp. 31–43. MR: 3485137 (cit. on pp. 2285, 2293).

Veronica Felli and Matthias Schneider (2003). "Perturbation results of critical elliptic equations of Caffarelli-Kohn-Nirenberg type". *J. Differential Equations* 191.1, pp. 121–142. MR: 1973285 (cit. on pp. 2288, 2294, 2295).

B. Gidas and J. Spruck (1981). "Global and local behavior of positive solutions of nonlinear elliptic equations". *Comm. Pure Appl. Math.* 34.4, pp. 525–598. MR: 615628 (cit. on p. 2284).

V. P. Ilin (1961). "Some integral inequalities and their applications in the theory of differentiable functions of several variables". *Mat. Sb. (N.S.)* 54 (96), pp. 331–380. MR: 0136692 (cit. on p. 2288).

Jean Rene Licois and Laurent Véron (1998). "A class of nonlinear conservative elliptic equations in cylinders". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 26.2, pp. 249–283. MR: 1631581 (cit. on p. 2284).

E. Lifchitz and L. Landau (1966). *Physique théorique. Tome II: Théorie du champ*. Deuxième édition revue. Traduit du russe par Edouard Gloukhian. Éditions Mir, Moscow, p. 451. MR: 0218091 (cit. on p. 2296).

Carl E. Mueller and Fred B. Weissler (1982). "Hypercontractivity for the heat semigroup for ultraspherical polynomials and on the n-sphere". *J. Funct. Anal.* 48.2, pp. 252–283. MR: 674060 (cit. on p. 2282).

Manuel del Pino, Jean Dolbeault, Stathis Filippas, and Achilles Tertikas (2010). "A logarithmic Hardy inequality". *J. Funct. Anal.* 259.8, pp. 2045–2072. MR: 2671121 (cit. on p. 2298).

Giuseppe Savaré and Giuseppe Toscani (2014). "The concavity of Rényi entropy power". *IEEE Trans. Inform. Theory* 60.5, pp. 2687–2693. MR: 3200617 (cit. on pp. 2285, 2293).

Giuseppe Toscani (2014). "Rényi entropies and nonlinear diffusion equations". *Acta Appl. Math.* 132, pp. 595–604. MR: 3255069 (cit. on pp. 2285, 2293).

Received 2017-11-29.

Jean Dolbeault
CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL Research University, Place de Lattre de Tassigny, F-75016 Paris, France
dolbeaul@ceremade.dauphine.fr

Maria J. Esteban
CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL Research University, Place de Lattre de Tassigny, F-75016 Paris, France
esteban@ceremade.dauphine.fr

Michael Loss
School of Mathematics, Skiles Building, Georgia Institute of Technology, Atlanta GA 30332-0160, USA
loss@math.gatech.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2305–2328)

ON LARGE TIME BEHAVIOR OF GROWTH BY BIRTH AND SPREAD

Yoshikazu Giga (儀我美一)

Abstract

This is essentially a survey paper on a large time behavior of solutions of some simple birth and spread models to describe growth of crystal surfaces. The models discussed here include level-set flow equations of eikonal or eikonal-curvature flow equations with source terms. Large time asymptotic speed called growth rate is studied. As an application, a simple proof is given for asymptotic profile of crystal grown by anisotropic eikonal-curvature flow.

1 Introduction

Equations describing front propagation or surface evolution are very important in various fields of science and technology. Let Γ_t be a hypersurface in $\mathbf{R}^N$ depending on time t, which describes, for example wave front or crystal surface. For simplicity, Γ_t is assumed to be closed so that it is the boundary of some bounded open set D_t. Let V be the normal velocity of Γ_t in the direction of $\mathbf{n}$, a unit normal vector field of Γ_t outward from D_t. The evolution given by a constant speed is often called Huygens' principle. Its explicit form is

$$V = \sigma \quad \text{on} \quad \Gamma_t, \tag{1-1}$$

where σ is a constant. This is a famous eikonal equation. To describe evolution of crystal surface, one has to consider anisotropy called kinetic anisotropy. It can be written

$$V = M(\mathbf{n})\sigma \quad \text{on} \quad \Gamma_t, \tag{1-2}$$

This work was partly supported by Japan Society for the Promotion of Science (JSPS) through grants No. 26220702 (Kiban S), No. 17H01091 (Kiban A) and No. 16H03948 (Kiban B)..
MSC2010: primary 35B40; secondary 35F21, 35K61, 35K65, 49L25, 53C44.

where $M(\mathbf{n})$ is a given positive function defined on a unit sphere. The function M is called a mobility. We refer this equation anisotropic eikonal equation. These equations are equations for one parameter family $\{\Gamma_t\}$.

In modern materials sciences, one also has to consider the curvature effect. The evolution is given by

$$V = aH + \sigma \quad \text{on} \quad \Gamma_t \tag{1-3}$$

with $a \geq 0$, where H is the $(N-1)$ times mean curvature in the direction of $\mathbf{n}$, i.e., $H = -\operatorname{div}_\Gamma \mathbf{n}$, where $\operatorname{div}_\Gamma$ denotes the surface divergence. If $\sigma = 0$ and $a = 1$, this equation is known as the mean curvature flow equation, which stems from materials science, and has been widely studied in mathematical community. Thus the equation (1-3) is often called the eikonal-curvature flow equation if $a > 0$ and $\sigma \neq 0$.

In materials science, one has to consider another anisotropy not only kinetic anisotropy. It is given as an anisotropic mean curvature or weighted mean curvature. Let γ be a given nonnegative function in $\mathbf{R}^N$ which is positively homogeneous of degree one, i.e., $\gamma(\lambda p) = \lambda\gamma(p)$ for all $\lambda > 0$, $p \in \mathbf{R}^N$. The anisotropic mean curvature H_γ is defined at least formally by

$$H_\gamma = -\operatorname{div}_\Gamma \xi(\mathbf{n}), \quad \xi(p) = \nabla_p \gamma = (\partial\gamma/\partial p_1, \ldots, \partial\gamma/\partial p_N).$$

It is known as the first variation of the interfacial energy $\int_\Gamma \gamma(\mathbf{n}) d\mathcal{H}^{N-1}$ with respect to a variation of hypersurface Γ, where $\mathcal{H}^{N-1}$ denotes the $N-1$ dimensional Hausdorff measure. If $\gamma(p) = |p|$ so that $\xi(\mathbf{n}) = \mathbf{n}$, H_γ is nothing but standard H. A typical anisotropic version of (1-3) is

$$V = M(\mathbf{n})(aH_\gamma + \sigma) \quad \text{on} \quad \Gamma_t. \tag{1-4}$$

It is very fundamental to ask whether or not the initial value problem for these equations is uniquely solvable. More precisely, the problem is that for a given initial data Γ_0 find $\{\Gamma_t\}_{t>0}$ solving (1-4). If one considers the problem globally-in-time, the singularity may develop for some smooth initial data even for (1-1). Thus one needs some weak notion of the solution. A level-set formulation is by now standard to solve such a problem globally-in-time. A level-set equation for (1-4) is the equation for u in $\mathbf{R}^N \times (0,\infty)$ such that *each* level-set of u moves by (1-4). To fix the orientation, we take $\mathbf{n} = -\nabla u/|\nabla u|$, where ∇ denotes the spatial gradient, i.e., $\nabla = (\partial_{x_1}, \ldots, \partial_{x_N})$, $\partial_{x_j} = \partial/\partial x_j$. For example, the level-set equation for (1-1) and (1-3) are

$$u_t - \sigma|\nabla u| = 0, \tag{1-5}$$

$$u_t - |\nabla u|\left(a \ \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) + \sigma\right) = 0, \tag{1-6}$$

respectively, where $u_t = \partial u/\partial t$. Unfortunately, level-set equations are highly degenerate in parabolic sense because there is no diffusion in the direction of ∇u. Fortunately, there is notion of viscosity solutions, (see e.g. Crandall, Ishii, and Lions [1992] and Y. Giga [2006]), based on order-preserving structure to handle continuous but non C^1 solutions. It turns out that such a notion is adjustable to this setting. Here are typical results. We consider the initial value problem of the level-set equation for (1-4), namely,

(1-7)
$$u_t - M\left(\nabla u/|\nabla u|\right)\left(a\ \operatorname{div}\left(-\xi\left(-\nabla u/|\nabla u|\right)\right)+\sigma\right)|\nabla u| = 0 \quad \text{in} \quad \mathbf{R}^N \times (0,\infty),$$
$$u|_{t=0} = u_0.$$

We shall use a short-hand notation $\{u > \ell\}$, $\{u < \ell\}$, $\{u = \ell\}$ to represent sets $\{(x,t) \mid u(x,t) > \ell\}$, $\{(x,t) \mid u(x,t) < \ell\}$ and $\{(x,t) \mid u(x,t) = \ell\}$, respectively.

Theorem 1.1. *Assume that $a \geq 0$ and $\sigma \in \mathbf{R}$ and that $M \geq 0$ is continuous. Assume that γ is convex and $\gamma(p) > 0$ for $p \neq 0$. Assume that $u_0 \in C(\mathbf{R}^N)$ equals constant α outside a ball. For γ assume either*

(a) (smoothness) $\gamma \in C^2\left(\mathbf{R}^N \backslash \{0\}\right)$

or

(b) (crystalline) γ is piecewise linear.

Then the following statements hold.
(Global solvability) The initial value problem for (1-7) *admits a unique continuous viscosity solution globally-in-time which equals α outside some ball in each finite time interval* $(0,T)$.
(Uniqueness of level-sets) The set $\{u < \ell\}$ (resp. $\{u > \ell\}$) depends only on $\{u_0 < \ell\}$ (resp. $\{u_0 > \ell\}$) and independent of the choice of u_0. The set $\{u = \ell\}$ is called the level-set flow solution of (1-4) *with initial data $\Gamma_0 = \{u_0 = \ell\}$.*

The assumption at space infinity does not restrict application if one considers a closed hypersurface. This statement for (a) was first proved by Chen, Y. Giga, and Goto [1991] and simultaneously for the level-set mean curvature flow equation by Evans and Spruck [1991] (corresponding the case $a = 1$, $\sigma = 0$ in (1-3)); see e.g. Y. Giga [2006] for details of the theory as well as related references. The case (b) of crystalline is not a simple generalization because the equation is nonlocal like total variation flow. A crystalline curvature flow was first introduced by Angenent and Gurtin [1989] and independently by Taylor [1991]; see also Gurtin [1993]. For $N = 2$, the statement in Theorem 1.1 was proved by M.-H. Giga and Y. Giga [2001], where more general γ is treated. For higher dimension $N \geq 3$, it is quite recent that this statement was proved by Y. Giga and Požár [2016], Y. Giga and Požár [2018] based on the work of M.-H. Giga, Y. Giga, and Požár [2013], M.-H. Giga, Y. Giga, and Požár [2014] in the sprit of M.-H. Giga and Y. Giga [1998]. The crucial steps are comparison principle and approximation arguments to construct a solution. Independently, A. Chambolle et al. Chambolle, Morini, and Ponsiglione [2017], Chambolle, Morini, Novaga, and Ponsiglione [2017] proved such a result for "convex" mobility but general convex γ including (b) adjusting formulation based on distance functions introduced first by Soner [1993].

If one looks the level-set equations, each level-set propagates by a given propagation law or surface evolution equations. This is also considered as spreading effects. For example, if one considers (1-5), each level-set spreads horizontally with velocity σ. Consider a crystal surface so that u is now the height of crystal. Assume that initially it is flat so that $u_0 = 0$. Then it does not grow just by spreading effect. One needs birth of crystal so that crystal grows. There are two typical mechanism of growth of crystal surface, see Burton, Cabrera, and Frank [1951]. One is the two-dimensional nucleation. The crystal surface grows by external supply of crystal molecules for a flat surface. It grows by catching such molecules. It is easy to catch molecules at the place where the crystal shape is not flat, i.e., $\nabla u \neq 0$ because of existence of microscopic steps. However, in the place where the surface is flat, there are no way to catch molecules unless there are step sources. At a very initial stage of the two-dimensional nucleation the step source catches crystal molecules so that a small disk-like island is formed at the step source on a flat face. Then this island grows by spreading and there occurs another birth of small disk-like island. It results a "wedding cake" consisting of several disks. This is a way of birth of new crystal surface in the two-dimensional nucleation.

The other mechanism of crystal growth is the spiral growth which is more popular. As pointed out in Burton, Cabrera, and Frank [ibid.], a pair of spirals opposite orientation whose centers are very close essentially forms a small island just like two-dimensional nucleation; see Ohtsuka, Tsai, and Y. Giga [2015], Smereka [2000], and Ohtsuka, Tsai, and Y. Giga [2018] for recent developments.

There are several models describing birth and spread macroscopically (see Ohara and Reid [1973]). If one fixes location of step source, it is of the form

$$u_t + F\left(\nabla u, \nabla^2 u\right) = r(x), \tag{1-8}$$

where $u_t + F$ is the left-hand side of the level-set equation (1-7) and $r(x) \geq 0$ is positive where step source exists. The simplest model is

$$u_t - \sigma|\nabla u| = cI(x), \quad I(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases} \tag{1-9}$$

with $c > 0$. This model is actually proposed by Schulze and Kohn [1999] to describe some high-temperature superconductor by approximating spiral growth on a crystal surface.

Our goal in this paper is to study the large-time behavior of the solution. We are especially interested in proving the asymptotic speed or the growth rate

$$\lim_{t\to\infty} u(x,t)/t$$

and its property. This is a very general question for partial differential equations of evolution type. There are by now several general theory for first order problems and some for second order problems (see review article by Mitake and Tran [2017]) but our problem is not covered by known theories like the weak KAM theory so far; see SubSection 2.3 and SubSection 5.1. The next result is a straightforward generalization of the result of Y. Giga, Mitake, Ohtsuka, and Tran [n.d.].

Theorem 1.2 (Existence of asymptotic speed)**.** *Assume that $r \geq 0$ is Lipschitz (continuous) and compactly supported. Assume the same assumption in Theorem 1.1 on (1-7) with $a \geq 0$, $\sigma > 0$ and $M > 0$. Assume further that $\gamma \in C^2(\mathbf{R}^N \backslash \{0\})$. Let u be the viscosity solution of (1-8) having the same left-hand side as (1-7) with $u_0 = 0$. Then the asymptotic speed $R = \lim_{t\to\infty} u(x,t)/t$, which is nonnegative, exists and the convergence is locally uniform.*

Let 1_E denote the characteristic function of E, i.e.,

$$1_E(x) = \begin{cases} 1, & x \in E \\ 0, & \text{otherwise.} \end{cases}$$

If $r_\varepsilon(x)$ is close to $c1_E$ in the sense $r_\varepsilon = c\eta_\varepsilon * 1_E$, where η_ε is the Friedrichs mollifier i.e., $\eta \in C_c^\infty(\mathbf{R}^N)$, $0 \leq \eta \leq 1$, $\int \eta \, \mathrm{d}x = 1$ $\eta(x) \equiv 0$ for $|x| \geq 3/4$ and $\eta_\varepsilon(x) = \varepsilon^{-N}\eta(x/\varepsilon)$, one might expect the asymptotic speed R_ε for r_ε converges to c as $\varepsilon \downarrow 0$. This is true for the first order model like $u_t - \sigma|\nabla u| = r(x)$. Unfortunately, this is not true in general for the second-order models. The next result easily follows by the comparison principle from similar results in Y. Giga, Mitake, and Tran [2016], where the case $r = c1_E$ is considered. For this non-Lipschitz r, we do not know even the existence of asymptotic speed.

Theorem 1.3. *Consider (1-8) with $r = r_\varepsilon$ in the plane. Assume the left-hand side is the same as (1-6) with $a = \sigma = 1$ and that $\gamma(p) = |p|$, $M(p) = |p|$. Assume that E is a closed square whose edge length is $2d$ with $d \in (1/\sqrt{2}, 1)$ so that E is not contained in nor not contains a unit disk. Then $0 < \liminf_{\varepsilon\to 0} R_\varepsilon \leq \limsup_{\varepsilon\to 0} R_\varepsilon < c$.*

This is because of curvature effect of spreading. There are few literature on asymptotic speed of second-order problems, for example, the work by Xin and Yu [2013], Xin and Yu [2014] studied the turbulent flow speed for what is called G-equations.

Our next concern is the asymptotic shape. For this purpose, we introduce a notion of the support function W_M of the polar of $1/M$, i.e.,

$$W_M(x) = \sup\{x \cdot p \mid |p| \leq 1/(M(p/|p|))\},\ x \in \mathbf{R}^N.$$

Its one sub level-set is often called the Wulff shape

$$\begin{aligned}\mathfrak{W}_M &= \left\{x \in \mathbf{R}^N \mid W_M(x) \leq 1\right\} \\ &= \bigcap_{|m|=1} \left\{x \in \mathbf{R}^N \mid x \cdot m \leq M(m)\right\}.\end{aligned}$$

Theorem 1.4 (Asymptotic profile). *Let u be as in Theorem 1.2. Then*

$$\lim_{\lambda\to\infty} u(\lambda x, \lambda t)/\lambda = R\,(t - W_M(x)/\sigma)_+$$

locally uniformly for $(x,t) \in \mathbf{R}^N \times (0,\infty)$, where $b_+ = \max(b,0)$.

Remark 1.5. *The results in Theorem 1.2 and Theorem 1.4 can be easily extended for general bounded uniformly continuous initial data by simple comparison with constant initial data. Thus R is independent of u_0.*

As a byproduct of our analysis, we give a simple new proof of anisotropic profile of level-set flow of (1-4) when the shape is growing. Such a result is originally proved by Ishii, Pires, and Souganidis [1999] a long time ago. From the point of (1-8), it is asymptotic profile of the horizontal growth.

Theorem 1.6. *Let Γ_t be the level-set flow solution in Theorem 1.1. Assume that $\gamma \in C^2(\mathbf{R}^N\backslash\{0\})$. Assume that Γ_0 strictly encloses $(a(N-1)/\sigma)\,\mathfrak{W}_\gamma$ up to translation. Then $\Gamma_t/t \to \partial\mathfrak{W}_M$ as $t \to \infty$ in the sense of the Hausdorff distance.*

Note that our assumption for Γ_0 is weaker than that of Ishii, Pires, and Souganidis [ibid., Theorem 6.1] where they assured Γ_0 encloses a sufficiently large ball. Note that their proof based on characterization of $\lim_{t\to\infty} u(tx,t)$ works even when γ is crystalline, where u is in Theorem 1.1. For crystalline case evolution of a convex shape by (1-4) is analyzed in M.-H. Giga and Y. Giga [2013], where the role of anisotropy in M and γ is clarified. We expect that all results in Theorem 1.2 – Theorem 1.6 can be extended to crystalline γ if appropriate stability holds (See SubSection 5.2).

This paper is organized as follows. In Section 2 we discuss the first-order model while in Section 3 we discuss the second-order model and give a sketch of the proof of Theorem 1.2. In Section 4 we prove Theorem 1.3, Theorem 1.4 and Theorem 1.6. In Section 5 we discuss unscaled asymptotic profiles and open problems.

2 First order models

We consider

$$u_t - \sigma|\nabla u| = r(x) \tag{2-1}$$

or its anisotropic version

$$u_t - \sigma M(-\nabla u) = r(x), \quad \left(M(p) := M\left(p/|p|\right)|p| \text{ for } p \in \mathbf{R}^N\right) \tag{2-2}$$

for a bounded upper semicontinuous function r. Here M is assumed to be continuous and nonnegative and $\sigma > 0$. If r is continuous, the standard theory of viscosity solutions yields a unique global-in-time solution for any bounded uniformly continuous initial data. However, if r is not continuous, typically $r(x) = cI(x)$, the solution may not be unique. We need to consider a kind of maximal solution which is formulated as an envelope solution in Y. Giga and Hamamuki [2013]. If one applies this equation to describe height of crystal surface by u, then it seems to be natural to consider a maximal solution; see Schulze and Kohn [1999]. In fact, there exists a unique global-in-time envelope solution for any such r when initial data u_0 is bounded uniformly continuous as proved by Y. Giga and Hamamuki [2013, Theorem 3.20]. Note that M doesn't need to be convex. We shall discuss several explicit solutions.

2.1 Explicit solutions. We consider (2-1) with $r(x) = cI(x)$, $\sigma > 0$ with initial data $u_0 = 0$.

Proposition 2.1. *Assume that $\sigma > 0$.*

(i) If $c > 0$, then $u_R(x,t) = R\left(t - |x|/\sigma\right)_+$ for $0 \le R \le c$ is a viscosity solution of (1-9) with $u_R|_{t=0} = 0$. The solution u_c is the unique envelope solution with zero initial data.

(ii) If $c \le 0$, then $u \equiv 0$ is a viscosity solution of (1-9) with initial data $u|_{t=0} = 0$ (It is actually the unique viscosity solution.)

Proof. (i) It is rather trivial to see that u_R solves (1-9) except $x = 0$. At the origin assume that $u_R - \varphi$ takes its maximum at $\left(0, \hat{t}\right)$, $\hat{t} > 0$ for some C^1 function φ. Then $\varphi_t\left(0, \hat{t}\right) \le R$ so that

$$\varphi_t\left(0, \hat{t}\right) - \sigma\left|\nabla\varphi\left(0, \hat{t}\right)\right| - cI(0) \le R - c \le 0.$$

Thus, u_R is a subsolution. Note that at $t = |x|/\sigma$, there is no way to touch from above. The test from below at $t = |x|/\sigma$ yields that u_R is a supersolution. We thus conclude that u_R is a solution.

(ii) This is very easy to check, so the proof is safely left to the reader. □

Its anisotropic version is as follows. We consider (2-2) with $r(x) = cI(x)$.

Proposition 2.2. *Assume that $\sigma > 0$.*

(i) If $c > 0$, then $u_R(x,t) = R\left(t - W_M(x)/\sigma\right)_+$ for $0 \le R \le c$ is a viscosity solution of

$$u_t - \sigma M(-\nabla u) = cI(x) \tag{2-3}$$

with $u_R|_{t=0} = 0$. The solution u_c is the unique envelope solution starting from zero.

(ii) If $c \le 0$, then $u \equiv 0$ is a viscosity solution of (2-3) with $u|_{t=0} = 0$.

The proof of this Proposition 2.2 is of course more involved. However, if one notices that $M(\nabla W_M) = 1$, it is rather easy.

2.2 Asymptotic speed and profile. The next result is a special case of Hamamuki [2013, Theorem 2.3]. For $u : \mathbf{R}^N \times (0,\infty) \to \mathbf{R}$, let $u^\lambda(x,t)$ be a rescaled function defined by

$$u^\lambda(x,t) = u(\lambda x, \lambda t)/\lambda \quad \text{for} \quad \lambda > 0.$$

Theorem 2.3. *Assume that $r \ge 0$ is continuous and compactly supported. Let u be the viscosity solution of (2-2) with initial data $u|_{t=0} = 0$. Let $c = \max r(x)$. Then $u^\lambda \to c\left(t - W_M(x)/\sigma\right)_+$ locally uniform as $\lambda \to \infty$.*

Proof. The proof given in Hamamuki [ibid.] is studying relaxed upper and lower limit. We here give a simple proof. We may assume that zero is a maximum point of r by translation. We know the solution is Lipschitz independent of regularity of r if initial data is Lipschitz (see e.g. Y. Giga and Hamamuki [2013]) since the Hamiltonian is coercive in the sense that $M(p) \to \infty$ as $|p| \to \infty$. Thus $\left\{u^\lambda(x,t)\right\}$ is equi-Lipschitz in $\mathbf{R}^N \times (0,T)$. By the Ascoli-Arzelà theorem for each sequence, there is a convergent subsequence u^{λ_i} and limit v such that $v^{\lambda_i} \to v$ locally uniformly (by diagonal argument) as $\lambda \to \infty$. By the stability of viscosity solutions (see e.g. Y. Giga [2006]), v satisfies (2-3) with $c = \max r(x)$.

Fortunately, $r_\lambda(x) \ge cI(x)$ for $r_\lambda(x) = r(\lambda x)$ so v must be the envelope solution of (2-3) and it must be unique. Thus, the convergence $u^\lambda \to v$ becomes full convergence and $v(x) = c\left(t - W_M(x)/\sigma\right)_+$. □

This statement is not exactly contained in Theorem 1.4 where r is assumed to be Lipschitz. This asymptotic results yield asymptotic speed as a Corollary.

Corollary 2.4. *Under the same assumption of Theorem 2.3, the asymptotic speed*

$$\lim_{t\to\infty} u(x,t)/t = c$$

exists and it is equals to $\max r$.

This is easy to prove by taking $t = 1$, $\lambda = t$, $x = x/t$. Note that the asymptotic speed is nothing but the maximum of r. For the first-order problem, the situation like Theorem 1.3 does not occur.

2.3 Non-coercive case. For a coercive case, large-time behavior is well studied. It goes back to the work of G. Namah and Roquejoffre [1999] and A. Fathi [1998]. It gives even asymptotic expansion $u(x,t) \sim ct + w(x)$ in the sense for a given ball B

$$\sup_{x\in B} |u(x,t) - ct - w(x)| \to 0 \quad \text{as} \quad t \to \infty.$$

Here w is a viscosity solution of a cell problem.

$$c - M(-\nabla w) = r(x).$$

Solutions may not be unique because the set $\{\hat{x} \mid \max r = r(\hat{x})\}$ plays a role of Aubry set. See Section 5. We do not touch this problem. We say the equation (1-8) of the form

$$u_t + F(\nabla u) = r(x)$$

is coercive if

$$\lim_{|p|\to\infty} F(p) = -\infty.$$

We notice that if the problem is non-coercive, the large-time behavior is not well studied although there are several works by Yokoyama, Y. Giga, and Rybka [2008], Y. Giga, Liu, and Mitake [2012], Y. Giga, Liu, and Mitake [2014] related to crystal growth. For example, if

$$u_t - \frac{|\nabla u|}{|\nabla u| + 1} = r(x), \tag{2-4}$$

it is not yet clear what the asymptotic speed is. Moreover, for a constant $c > 0$, the uniqueness of a solution with $r = cI(x)$ is not guaranteed. In fact, according to Y. Giga and Hamamuki [2013, Example 5.15]

$$U_c(x,t) = \begin{cases} ct - \dfrac{c}{1-c}|x| & (|x| \le (1-c)^2 t) \\ \left(\left(\sqrt{t} - \sqrt{|x|}\right)_+\right)^2 & (|x| \ge (1-c)^2 t) \end{cases}$$

is a unique envelope solution of (2-4) with $r = cI(x)$ when $c \le 1$. However, if $c > 1$ even an envelope solution may not be unique.

3 Second order models

3.1 Models. We now consider the equation (1-8) for (1-7), namely

$$u_t - M(-\nabla u)\left(a \operatorname{div}\left(-\xi(-\nabla u)\right) + \sigma\right) = r(x). \tag{3-1}$$

The major difference from (2-2) is that the curvature effect is included in spreading process; see Figure 1 for the graph of u governed by (3-1). In particular, if the radius of island is too small, it does not spread. In the two-dimensional nucleation, it is more realistic to consider the case that the place of birth may depend on time, i.e., r may depend on t. However, in this note we only consider the case when r is independent of time because it is already complicated than what we expect. Moreover, if one uses this model to describe the spiral growth, r must be independent of time and this is better approximation than (2-2).

We first recall the well-posedness of the initial value problem for (3-1).

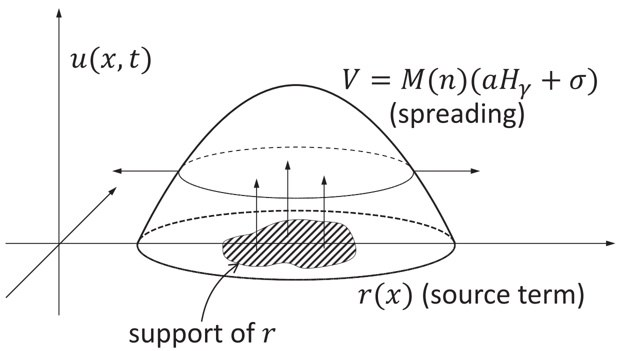


Figure 1: The graph of u at time t solving (3-1)

Theorem 3.1 (Solvability)**.** *Assume the same hypotheses of Theorem 1.1 concerning a, σ, M, γ and u_0. Assume that r is continuous and has compact support. Then the initial value problem for (3-1) with $u|_{t=0} = u_0$ admits a unique continuous viscosity solution u globally-in-time which equals α outside some ball in each finite time interval.*

Proof. The proof for the case when γ is C^2 outside the origin is by now standard and well known as in the book of Y. Giga [2006]. However, the case when γ is crystalline is quite new even if $r = 0$, see Y. Giga and Požár [2016], Y. Giga and Požár [2018]. We first prove the comparison principle. Suppose that u is viscosity subsolution and v is a viscosity supersolution as in Y. Giga and Požár [2016], Y. Giga and Požár [2018]. Then we have to conclude that $u \le v$ if initially $u \le v$. The definition for viscosity sub and supersolution with $r \not\equiv 0$ is not given there but it is obtained as a trivial modification. This comparison principle can be proved along the line of Y. Giga and Požár [2016], Y. Giga and Požár [2018] if one replaces doubling variable procedure with shift parameter ζ by

$$\Phi_\zeta = u(x,t) - v(y,s) - \frac{|x-y-\zeta|^2}{2\varepsilon} - S_{\varepsilon,\delta}(t,s),$$

where $S_{\varepsilon,\delta}(t,s)$ should be

$$S_{\varepsilon,\delta}(t,s) = \frac{|t-s|^2}{2\varepsilon} + \frac{\delta}{T-t} + \frac{\delta}{T-s};$$

in Y. Giga and Požár [2016], Y. Giga and Požár [2018] δ is taken to be equal to ε. We fix $\delta > 0$ small enough as in M.-H. Giga and Y. Giga [1998] unrelated to ε. We argue by contradiction as in Y. Giga and Požár [2016], Y. Giga and Požár [2018] and end up with

$$\frac{2\delta}{T^2} \le \frac{\delta}{(T-\hat{t})^2} + \frac{\delta}{(T-\hat{s})^2} \le r(\hat{x}) - r(\hat{y}),$$

where $(\hat{x}, \hat{t}, \hat{y}, \hat{s})$ is a maximum point of Φ_ζ for small ζ depending on ε. If $\varepsilon \to 0$ with $\zeta \to 0$, we observe $\hat{x} - \hat{y} \to 0$ so we get a contradiction. Existence can be proved by approximation as in Y. Giga and Požár [2016], Y. Giga and Požár [2018]. $\square$

3.2 Radial case and its generalization. We consider the special case when $M = \gamma$. In other words, kinetic anisotropy agrees with interfacial anisotropy. Moreover, assume that r depends only on $W_\gamma(x)$, i.e., $r(x) = h(W_\gamma(x))$, We postulate that the solution $u(x,t)$ only depends on $W_\gamma(x)$, i.e., u has the form

$$u(x,t) = U(W_\gamma(x), t).$$

Since we know that $\xi(\nabla W_\gamma(x)) = x/|x|$ so that $\operatorname{div}\xi(W_\gamma(x)) = N-1$ and $\gamma(\nabla W_\gamma) = 1$,

$$u_t - \gamma(-\nabla u)\left(a \operatorname{div}(-\xi(-\nabla u)) + \sigma\right) = h(W_\gamma(x)) \tag{3-2}$$

is reduced to

$$U_t - \frac{a(N-1)}{\rho} U_\rho + \sigma U_\rho = h(\rho) \tag{3-3}$$

if $U = U(\rho, t)$ is a nonincreasing function with respect to ρ, i.e., $U_\rho \le 0$. The equation for U is the same as in radial solution for isotropic case. The equation (3-3) is now linear first order but non-coercive Hamilton-Jacobi equation with singularity.

To see that the asymptotic speed may not be max h, we give a few examples. We rather consider discontinuous h of the form

$$h = c1_{[0,\rho_0]}. \tag{3-4}$$

Since the right-hand side is not continuous, we do not expect uniqueness of viscosity solutions. We rather consider the maximal solution. We set a critical number

$$\rho_* = a(N-1)/\sigma$$

and define

$$\psi_0(\rho) = c\left\{\rho + \rho_* \log|\sigma\rho - a(N-1)|\right\}/\sigma = c\left\{\rho + \rho_* \log|\sigma(\rho - \rho_*)|\right\}/\sigma,$$

which solves

$$\left(-\frac{a(N-1)}{\rho} + \sigma\right)\partial_\rho \psi = c \quad \text{for} \quad \rho \ne \rho_*.$$

Moreover, $\partial_\rho \psi_0(0) = 0$ as we expected.

Theorem 3.2. *Consider (3-2) with (3-4) and $u|_{t=0} = 0$. Assume that γ is C^2 outside the origin or that γ is crystalline with $N = 2$. Assume that $\sigma > 0$, $a > 0$.*

(i) If $\rho_0 < \rho_*$, then

$$u(x,t) = \begin{cases} \min\left\{\psi\left(W_\gamma(x)\right)_+, ct\right\} & \textit{for } x \textit{ with } W_\gamma(x) < \rho_* \\ 0 & \textit{for } x \textit{ with } W_\gamma(x) \ge \rho_* \end{cases}$$

with $\psi(\rho) = \psi_0(\rho) - \psi_0(\rho_0)$ is the maximal viscosity solution.

(ii) If $\rho_0 > \rho_*$, then

$$u(x,t) = \begin{cases} \min\left\{ct, \left(ct - \psi\left(W_\gamma(x)\right)\right)_+\right\} & \textit{for } x \textit{ with } W_\gamma(x) \ge \rho_* \\ ct & \textit{for } x \textit{ with } W_\gamma(x) < \rho_* \end{cases}$$

is the maximal viscosity solution.

(iii) If $\rho_0 = \rho_*$, then

$$u(x,t) = ct1_{W_\gamma}$$

is the maximal viscosity solution.

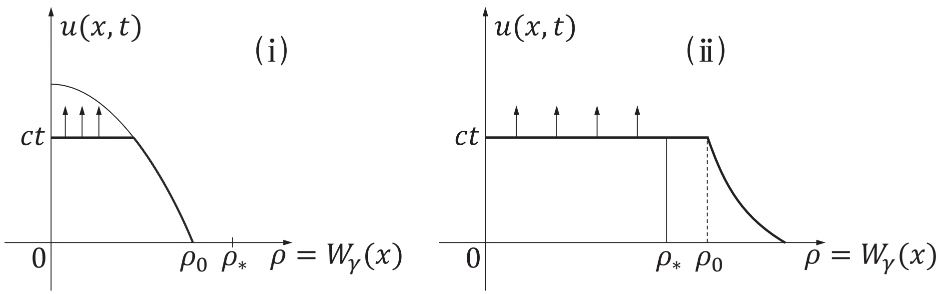


Figure 2: The graph of u at time t

From this we see the growth speed depends on geometry where u takes maximum. This is quite different from the first-order model. In fact in the case (i) $u(x,t)$ is bounded as $t \to \infty$ and $\lim_{t\to\infty} u(x,t)/t = 0$. In the case (ii) $c = \lim_{t\to\infty} u(x,t)/t$ for all $x \in \mathbf{R}^N$ while in the case (iii) $c = \lim_{t\to\infty} u(x,t)/t$ for $x \in \rho_* \mathfrak{W}_\gamma$ while outside $\rho_* \mathfrak{W}_\gamma$ we observe that $u(x,t) \equiv 0$. See Figure 2 for profiles of the graph of u at time t.

By the way, the function $U(\rho,t) = \min\{\psi(\rho), ct\}$ solves (3-3) with (3-4) for $\rho \in (0,\rho_*)$, $t \in \mathbf{R}$ if $\rho_0 < \rho_*$ while $U(\rho,t) = \min\{ct, ct - \psi(\rho)\}$ solves (3-3) with (3-4) for $\rho > \rho_*$, $t \in \mathbf{R}$ if $\rho_0 > \rho_*$ in viscosity sense. If one omits the plus part symbol in (i), (i i), then u is an entire viscosity solution, i.e., it solves (3-1) for all $t \in \mathbf{R}$, $x \in \mathbf{R}^N$; in the case of (i) one has to exclude the place where $\rho = \rho_*$.

The results in Theorem 3.2 is essentially proved in Y. Giga, Mitake, and Tran [2016], where the isotropic case, i.e., $\gamma(p) = |p|$ or $W_\gamma(x) = |x|$ is discussed. Theorem 3.2 is a trivial extension of results in Y. Giga, Mitake, and Tran [ibid., Sect. 4] to anisotropic case. For crystalline case, this result should be true. Although it is easy to see that the proposed solution is a viscosity solution, to show the maximality we need some stability of crystalline level-set equation which is so far not available for $N \geq 3$; see Section 5.2. The case $N = 2$ is proved in M.-H. Giga and Y. Giga [2001].

3.3 Lipschitz bounds. We shall derive Lipschitz bounds in time and space for (3-1) when the initial data $u_0 = 0$ and $r \geq 0$ is Lipschitz. These are straightforward extension of those in Y. Giga, Mitake, Ohtsuka, and Tran [n.d.], where the isotropic case is discussed.

Lemma 3.3 (Bound for time derivative). *Assume the same hypotheses of Theorem 3.1 concerning a, σ, M, r. Assume that $u_0 = 0$ and $r \geq 0$. Let u be the viscosity solution of (3-1) in Theorem 3.1. Then u is Lipschitz in t and*

$$0 \leq u_t(x,t) \leq c := \max_{\mathbf{R}^N} r(x)$$

for all $x \in \mathbf{R}^N$ *and almost all* $t \geq 0$.

Proof. Since $v(x,t) = ct$ is a viscosity supersolution (see e.g. Y. Giga [2006]) and $w \equiv 0$ is a viscosity supersolution of (3-1), by the comparison principle we easily see that

$$0 \leq u(x,t) \leq ct \quad \text{in} \quad \mathbf{R}^N \times [0,\infty)$$

For any given $s > 0$, both

$$u^s(x,t) := u(x,t+s) \quad \text{and} \quad u(x,t)$$

are viscosity solutions of (3-1). Since $u^s(x,0) \geq u(x,0) = u_0 = 0$, by the comparison principle we obtain

$$0 \leq (u^s - u)(x,t) \leq \sup_{\mathbf{R}^N}(u^s - u_0)\Big|_{t=0} \leq cs$$

which yields the desired estimate. □

Lemma 3.4 (Bound for spatial derivative)**.** *Assume the same hypotheses of Lemma 3.3. Assume furthermore that* $\gamma \in C^2\left(\mathbf{R}^N \backslash \{0\}\right)$ *and that* r *is Lipschitz. Then* u *is spatially Lipschitz and it gradient is essentially bounded. More precisely, its* L^∞*-norm has a bound*

$$\|\nabla u\|_{L^\infty(\mathbf{R}^N)}(t) \leq K$$

with K *independent of* $t \in (0,\infty)$.

This can be proved by what is called Bernstein's method. We first recall a simple matrix inequality.

Lemma 3.5. *Let* A *and* B *be real symmetric matrices. Assume that* A *is nonnegative definite, i.e.,* $A \geq 0$. *Then*

$$(\operatorname{tr} AB)^2 \leq \operatorname{tr}(ABB)\ \operatorname{tr} A.$$

This follows from the Schwarz inequality

$$\left(\operatorname{tr}({}^t ab)\right)^2 \leq \operatorname{tr}{}^t aa\ \operatorname{tr}{}^t bb$$

for general real square matrices a, b by setting $a = A^{1/2}$, $b = A^{1/2}B$, where ${}^t a$ denotes the transpose of a.

Formal proof of Lemma 3.4. We write (3-1) in the form of (1-8) with $F = F(p,X)$. Pretending that everything is smooth, we differentiate (1-8) in x_k to get

$$u_{kt} + \sum_{\ell=1}^{N} \frac{\partial F}{\partial p_\ell} u_{kl} + \sum_{i,j=1}^{N} \frac{\partial F}{\partial X_{ij}} u_{kij} = r_k,$$

where we use a shorthand notation $u_k = \partial_{x_k} u$, $u_{k\ell} = \partial_{x_k}\partial_{x_\ell} u$ and so on. We multiply u_k and add from 1 to N to get differential inequality for $w = \sum_{k=1}^N u_k^2/2$ of the form

$$w_t + \sum_{\ell=1}^N \frac{\partial F}{\partial p_\ell} w_\ell + \sum_{\ell,i,j=1}^N \frac{\partial F}{\partial X_{ij}} (w_{ij} - u_{j\ell} u_{i\ell}) = \nabla r \cdot \nabla u.$$

We set

$$a_{ij} = -\frac{\partial F}{\partial X_{ij}} = aM(-p)\frac{\partial^2 \gamma}{\partial p_i \partial p_j}(-p) \quad \text{with} \quad p = \nabla u.$$

By Lemma 3.5 we observe that

$$\sum_{i,j,\ell}^N a_{ij} u_{j\ell} u_{i\ell} \ge \left(\sum_{i,j} a_{ij} u_{ij}\right)^2 \Big/ \sum_i a_{ii} \ge \left(\sum_{i,j} a_{ij} u_{ij}\right)^2 \Big/ A$$

with some constant $A > 0$ independent of p since $\sum_i a_{ii} \le A$. We now obtain

$$w_t + \sum_\ell \frac{\partial F}{\partial p_\ell} w_\ell - \sum_{i,j} a_{ij} w_{ij} + \left(\sum_{i,j} a_{ij} u_{ij}\right)^2 \Big/ A \le \nabla r \cdot \nabla u.$$

Since the equation (1-8) is quasilinear, we observe that

$$\sum_{i,j} a_{ij} u_{ij} = u_t - \sigma M(-\nabla u) - r.$$

Since $|u_t| \le c$ and $\sigma M(p) \ge m_0 |p|$ with some constant $m_0 > 0$, we see that

$$\left(\sum_{i,j} a_{ij} u_{ij}\right)^2 \ge (m_0|\nabla u| - c)^2 \ge \frac{m_0}{2}|\nabla u|^2 - c^2 \quad \text{if} \quad m_0|\nabla u| > c.$$

Let β_0 be the largest zero of

$$\left(\frac{m_0}{2}\beta^2 - c^2\right) \Big/ A - \|\nabla r\|_{L^\infty(\mathbf{R}^N)} \beta = 0.$$

We thus conclude that

$$w_t + \sum_\ell \frac{\partial F}{\partial p_\ell} w_\ell - \sum_{i,j} a_{ij} w_{ij} \le 0 \quad \text{if} \quad |\nabla u| \ge \max(\beta_0, c/m_0) =: \beta_1.$$

By the comparison principle (assuming that the space infinity is well controlled), we observe that $w \le \beta_1^2/2$. We now obtain the desired bound $K = \beta_1$ at least formally.

To realize the idea, we fix time and approximate the equation so that the singularity near $\nabla u = 0$ zero is removed and that the problem is uniformly elliptic to get a smooth solution. We have skipped all this procedure and have left the details to Y. Giga, Mitake, Ohtsuka, and Tran [n.d.]. □

3.4 Existence of asymptotic speed. We are in position to prove the existence of asymptotic speed (Theorem 1.2). For this purpose, we check the motion of the top. We set

$$m(t) = \sup_{x \in \mathbf{R}^N} u(x,t).$$

Lemma 3.6. *Assume the same hypothesis of Theorem 1.2. Then $m(t)$ is subadditive and $R := \lim_{r\to\infty} m(t)/t$ exists and equals $\inf_{t>0} m(t)/t$ with $R \in [0,\infty)$.*

Proof. Since $v(x,t) = u^s(x,t) - m(s)$ is a subsolution of (3-1) with $u^s(x,t) = u(x,t+s)$ for $s > 0$ and since $v(x,0) \le u(x,s) - m(t) \le 0 = u(x,0)$, by the comparison principle we see that $v(x,t) \le u(x,t)$ in $\mathbf{R}^N \times (0,\infty)$. Take sup in both sides in x to get

$$m(t+s) - m(s) \le m(t)$$

which implies the subadditivity. The other assertion follows by Fekete's lemma (see e.g. Barles [2013, p. 95] for the proof) for a subadditive function.

In this argument, we do not use Lipschitz bound so Lemma 3.6 is still valid for continuous r. Also it applies to the case of crystalline. □

Proof of Theorem 1.2. Since $u \ge 0$, if $R = 0$ in Lemma 3.6, the convergence

$$R = \lim_{t\to\infty} u(x,t)/t$$

immediately follows with $R = 0$. We may assume that $R > 0$.

It suffices to prove that for a given ball B and $\varepsilon > 0$ there exists T such that

$$u(x,t)/t \ge R - \varepsilon \quad \text{for} \quad t > T,\ x \in B.$$

We may assume that B includes $\operatorname{supp} r$, the support of r. Assume that x_t is the maximizer of $u(x,t)$, i.e., $u(x_t,t) = m(t)$. Since the support of u is contained in some ball depending only on $t_0 > 0$ for $t \in (0,t_0)$, the existence of x_t is trivial. By a Lipschitz bound in Equation (3-4), we see that

$$\frac{u(x,t)}{t} \ge \frac{u(x,t) - u(x_t,t)}{t} + \frac{m(t)}{t} \ge -K\frac{|x - x_t|}{t} + \frac{m(t)}{t}.$$

If we admit that x_t is in the convex hull of $\operatorname{supp} r$ as stated in the next lemma, we take T large such that $m(t)/t \geq R - \varepsilon/2$ for $t > T$ and $2K\rho_1/T < \varepsilon/2$ to get $u(x,t)/t \geq R - \varepsilon$ for $t > T$, $x \in B_{\rho_1}$. Note that almost the same argument is found in the proof of Barles [ibid., Theorem 10.2]. □

Lemma 3.7. *Assume the same hypotheses of Theorem 1.2. Let S denote the convex hull of* $\operatorname{supp} r$. *Then* $\max_S u(\cdot,t) \geq \sup_{S^c} u(\cdot,t)$, *where S^c denotes the complement of S in* $\mathbf{R}^N$. *In particular, $x_t \in S$.*

This is nontrivial because x_t may not be a maximum point of r which is quite different from the first-order case. Such a difference essentially comes from the monotonicity of the geometric flow in the first-order case which determines the way of spreading. To see that x_t may not be a maximum point of r for the second-order case, it suffices to consider $r = \left(1_{B_{\rho_1}(0)} + \frac{1}{2}1_{B_{\rho_2}(q)}\right) * \eta_\varepsilon$ in $\mathbf{R}^2$ for isotropic case $a = \sigma = 1$, $\gamma(p) = |p| = M(p)$ with $\rho_1 < \rho_* - \varepsilon$, $1 = \rho_* < \rho_2 < 2 - \varepsilon$, $\varepsilon \in (0,1)$ and $|q| = 3$. From the observation for radial case the effect of $1_{B_{\rho_1}(0)}$ will eventually negligible for large time and the maximum is taken in $B_{\rho_2}(q)$. Moreover, $\sup_{S^c} u \leq \inf_S u$ may not hold.

Proof of Lemma 3.7. We set that $c(t) = \max_{\partial S} u(\cdot,t)$ and

$$w(x,t) = \left(c(t) - c'(t)\rho(x)/\sigma\right)_+, \ x \in S^c, \ t \geq 0$$

with

$$\rho(x) = d_{W_M}(x,S) := \inf\{W_M(x-y) \mid y \in S\}.$$

Formally, it is clear that w solves $u_t - \sigma M(-\nabla u) = 0$ in $S^c \times (0,\infty)$ with $w = c(t)$ on ∂S provided that the time derivative $c' \geq 0$. This can be proved rigorously as in Proposition 2.1 and Proposition 2.2. Since $u_t \geq 0$, we see that $c' \geq 0$. Since S is convex, so is ρ. Thus w is a viscosity supersolution of (1-7) in $S^c \times (0,\infty)$. By a comparison principle (see e.g. Y. Giga [2006]), we see that $u \leq w$ in $S^c \times (0,\infty)$. This yields the desired result. □

4 Asymptotic profile

4.1 Limit equations. We shall prove Theorem 1.4 in the second-order case. A stronger result for the first-order model is stated as Theorem 2.3.

Proof of Theorem 1.4. As in the first-order case, we may assume that the origin is contained in the interior of $\operatorname{supp} r$. As in the first-order case, $u^\lambda(x,t) = u(\lambda x, \lambda t)/\lambda$ has uniform Lipschitz bound (Lemmas 3.3, 3.4) in space-time, for each subsequence of $\lambda \to \infty$

there is a converges subsequence u^{λ} and a limit v such that $u^{\lambda_j} \to v$ uniformly. Moreover, by the stability of viscosity solution v must solve (2-3) outside the origin. Note that the second-order term disappears. At the origin by Theorem 1.2

$$u^{\lambda}(0, \lambda t)/\lambda \to Rt \quad \text{as} \quad \lambda \to \infty$$

locally uniformly in t. Thus $v(0,t) = Rt$. Since it is not difficult to show that

$$w(x,t) = R\left(t - W_M(x)\right)_+$$

is the unique Lipschitz solution of (2-3) outside the origin with the Dirichlet boundary condition $w(0,t) = Rt$, we conclude that $v = w$ and the convergence becomes full convergence. The proof is now complete. □

4.2 Case of intermediate speed. Note that our limit function satisfies (2-3) but it is not an envelope solution if $R < c$. Our Theorem 1.3 actually shows that there is an intermediate case.

Proof of Theorem 1.3. We know by Y. Giga, Mitake, and Tran [2016] that there is an intermediate case for the maximal solution of

$$u_t - |\nabla u|\left(\operatorname{div}\left(\nabla u/|\nabla u|\right) + 1\right) = c 1_{E_\ell} \tag{4-1}$$

if $E_\ell \subset \mathbf{R}^2$ is a square of edge length 2ℓ with $\ell \in \left(1/\sqrt{2}, 1\right)$. For given d in Theorem 2.3, we take $\varepsilon > 0$ small so that $1_{E_{\ell_1}} \geq r_\varepsilon \geq 1_{E_{\ell_2}}$ so that $\ell_2 < d < \ell_1$ and $\ell_i \in \left(1/\sqrt{2}, 1\right)$ $(i = 1, 2)$. Let u_i be the maximal solution of (4-1) with initial data $u_i|_{t=0} = 0$ and $\ell = \ell_i$ $(i = 1, 2)$. We know $\limsup_{t\to\infty} u_1/t \leq c_1 < c$, $\liminf_{t\to\infty} u_2/t \geq c_2 > 0$ by Y. Giga, Mitake, and Tran [ibid.]. By comparison, $u_1 \leq u \leq u_2$ thus $c_2 \leq R_\varepsilon \leq c_1$ for sufficiently small ε. □

4.3 Asymptotic profile of large level set. We shall give a simple proof for Theorem 1.6 based on Theorem 1.4. Let $\{E_t\}_{t\geq 0}$ be an increasing family of bounded closed sets which exhausts $\mathbf{R}^N$, i.e., for any compact set K there is t such that $K \subset E_t$. If $\{E_t\}$ is exhaustive,

$$q(x, E_0) = \inf\{t \geq 0 \mid x \in E_t\}.$$

is well defined for all $x \in \mathbf{R}^N$. It is continuous if E_t is continuous in t in Hausdorff distance sense and strictly monotone in the sense that $E_t \subset \inf E_s$ for $s > t \geq 0$.

For a given bounded closed set E_0, let S_t be a level-set flow of (1-4) starting from $S_0 = \partial E_0$. As is in Y. Giga [2006], we say that the open set D_t enclosed by S_t is called an open evolution while $E_t = D_t \cup S_t$ is called a closed evolution starting, respectively, D_0 and E_0.

Lemma 4.1. *Let $E_0 = \kappa \mathfrak{W}_\gamma$ with $\kappa > \rho_*(= a(N-1)/\sigma)$. Let E_t be the closed evolution of (1-4) starting from E_0. Then the function $w_\kappa(x,t) = (t - q(x))_+$ with $q = q(x, E_0)$ is a viscosity solution of (1-8) with $r = 1_{E_0}$ and $w|_{t=0} = 0$.*

Proof. It is easy to see that w is a viscosity solution of (1-8) once q is a well-defined continuous function. Since initially $E_0 \subset \inf E_s$ for $s > 0$, the strict monotonicity $E_{t_1} \subset \operatorname{int} E_{t_2}$ for $t_1 < t_2$ is clear by comparison. Upper semicontinuity of E_t is trivial and left lower semicontinuity follows from a general theory Y. Giga [ibid., Theorem 4.5.5]. The right lower semicontinuity follows from the monotonicity, so E_t is continuous in t in the Hausdorff distance sense.

To show that E_t is exhaustive, we compare with a special solution of $V = m_0\gamma(aH_\gamma + \sigma)$ such that a constant $m_0 > 0$ is taken so that $M(p) \geq m_0\gamma(p)$ for $p \in \mathbf{R}^N$. Since this equation has a self-similar growing solution (see Soner [1993], Gurtin [1993]) of the form $\lambda(t)W_\gamma$ with $\lambda(t) \to \infty$ as $t \to 0$ and since such a solution is a subsolution of of (1-4) in the level-set sense, by comparison $\{E_t\}$ is exhaustive. □

Proof of Theorem 1.6. We shall prove that

$$\lim_{t\to\infty} q(tx)/t = W_M(x)$$

locally uniformly in $x \in \mathbf{R}^N$. There is a Lipschitz function r such that $0 \leq r \leq 1$ and the set $\{r = 1\}$ equals $\kappa' W_\gamma$, $\kappa' > \rho_*$, $\kappa' < \kappa$ and $\operatorname{supp} r \subset \kappa W_\gamma$. By comparison, it is clear that

$$w_{\kappa'}(x,t) \leq u(x,t) \leq w_\kappa(x,t),$$

where u is the solution of (3-1) with zero initial data. The estimate $w_{\kappa'} \leq u$ implies that the asymptotic speed of u must be one. By Theorem 1.4 and $u \leq w_\kappa$, we see that

$$\lim_{t\to\infty} u(tx,t)/t = (1 - W_M(x)/\sigma)_+ \leq (1 - \limsup{}^*_{t\to\infty} q(tx)/t)_+\,,$$

where $\limsup^*$ is a relaxed limit, i.e., it is defined as

$$\limsup{}^*_{t\to\infty} f(t,x) = \lim_{t\to\infty} \sup\{f(s,y) \mid s \geq t, |y-x| \leq 1/t\}.$$

This in particular implies that

$$\text{(4-2)} \qquad \limsup{}^*_{t\to\infty} q(tx)/t \leq W_M(x)/\sigma \quad \text{for} \quad x \quad \text{satisfying} \quad W_M(x) \leq 1.$$

The other estimate is easy. It is easy to see that $(\sigma t + \beta)\partial\mathfrak{W}_M$ with $\beta > 0$ is a level-set supersolution of (1-4) (which is a solution of (1-2)). We take β large enough so that $\beta\mathfrak{W}_M$ includes $\kappa\mathfrak{W}_\gamma$. By comparison, $E_t \subset (\sigma t + \beta)\mathfrak{W}_M$ since $E_0 = \kappa\mathfrak{W}_\gamma$. This implies

$$\liminf{}_{*t\to\infty} q(tx)/t \geq W_M(x)/\sigma,$$

where $\liminf_{*t\to\infty} f = -\limsup^*_{t\to\infty}(-f)$. This implies $\lim_{t\to\infty} q(tx)/t = W_M(x)/\sigma$ locally uniformly since $W_M(x)$ is positively homogeneous of degree one.

The estimate (4-2) implies that for any $\varepsilon > 0$

(4-3) $$\sigma \mathcal{W}_{M,\varepsilon} \subset t^{-1} D_t$$

for sufficiently large t, where $\mathcal{W}_{M,\varepsilon} = \left\{x \in \mathcal{W}_M \mid \operatorname{dist}(x, \mathcal{W}_M^c) > \varepsilon\right\}$. Note that there is no fattening in this setting (1-8) so that $\operatorname{int} E_t = D_t$. The estimate $E_t \subset (\sigma t + \beta)\mathcal{W}_M$ implies

(4-4) $$t^{-1} E_t \subset \sigma \mathcal{W}_M^\varepsilon$$

for sufficiently large t, where $\mathcal{W}_M^\varepsilon = \{x \in \mathbf{R}^n \mid \operatorname{dist}(x, \mathcal{W}_M) < \varepsilon\}$.

For general Γ_0, we compare with $\kappa\mathcal{W}_\gamma$ and $\kappa'\mathcal{W}_\gamma$ so that $\kappa\mathcal{W}_\gamma \supset \Gamma_0$ or Γ_0 encloses $\kappa'\mathcal{W}_\gamma$ for a suitable choice of $\kappa, \kappa' > \rho_*$. The desired results like (4-3) and (4-4) for this initial data follows from comparison principle and behavior of solutions starting from $\kappa\mathcal{W}_\gamma$ or $\kappa'\mathcal{W}_\gamma$. □

Remark 4.2. *In Ishii, Pires, and Souganidis [1999, Theorem 6.1], a more general equation like*

$$V = v_1(\mathbf{n}, \mathbf{A}) + \sigma M(\mathbf{n}) \quad \sigma > 0, M > 0$$

is handled under the assumptions that v_1 is monotone nondecreasing in the second fundamental form $\mathbf{A}$ in the direction of $\mathbf{n}$ and positively homogeneous of degree one in $\mathbf{A}$, i.e., $v_1(\mathbf{n}\ \lambda\mathbf{A}) = \lambda v_1(\mathbf{n}, \mathbf{A})$, $\lambda > 0$ not necessarily linear. This case can be handled in our setting. The crucial step is to obtain a Lipschitz bound where we have used

$$-\sum_{i,j} \frac{\partial F}{\partial X_{ij}} X_{ij} = u_t - \sigma M(-\nabla u) - r.$$

Fortunately, this equality still holds if v_1 satisfies the Euler equation, i.e., $\sum_i \frac{\partial f}{\partial p_i} p_i = f$ for homogeneous functions.

5 Unscaled asymptotic profile

5.1 Large time convergence of a solution. We next try to find an unscaled asymptotic profile in the sense that we seek a function w such that for any ball B

(5-1) $$\sup_{x\in B} |u(x,t) - Rt - w(x)| \to 0,$$

as $t \to \infty$. Here, (w, R) satisfies a stationary problem $R + F(\nabla w, \nabla^2 w) = r(x)$ in $\mathbf{R}^n$. We emphasize here that, in general, solutions to this stationary problem are not unique even up to additive constants. See examples in Mitake and Tran [2017, Chapter 6] for instance. Therefore, the convergence (5-1) is not trivial in general. Such a problem is well studied in the first order model. It was started by Namah and Roquejoffre [1999] and Fathi [1998]. The problem is especially well studied for the Hamilton-Jacobi equations $u_t + H(x, \nabla u) = 0$ for convex Hamiltonian H in a periodic setting. For $\mathbf{R}^N$ setting, see the work of H. Ishii [2008]. These results are based on approach by dynamical systems. There is a PDE approach by Barles and Souganidis [2000] which covers some class of non convex Hamiltonian. However, for the second-order problems less is known especially parabolicity is degenerated. Recently, nonlinear adjoint method introduced by L. C. Evans [2010] is adjusted to apply such a kind of problems of large time behavior by Cagnetti, Gomes, Mitake, and Tran [2015]. This method allows some degenerate second order term but it does not apply to our second model because the degeneracy depends on a solution. The reader is referred to a recent nice survey by Mitake and Tran [2017] for more details and references.

For the solution u to the initial value problem of (3-1), what we know is that $u(x,t)-Rt$ converges locally uniformly as $t \to \infty$ to some function w by taking a subsequence because of Lipschitz bound. Moreover, w solves $R + F(\nabla w, \nabla^2 w) = r(x)$. However, such an equation is not well studied even under periodic setting. Therefore, the full convergence (5-1) is not yet known.

We finally point out that the asymptotic speed R is independent of the choice of initial data but the profile may depend on the initial data in a nonlinear way. Therefore, a key question here could be how w depends on u_0. For the first order case with a convex Hamiltonian, a representation formula for w is given by Davini and Siconolfi [2006], where the values of initial data on the Aubry set and the infimum stability of viscosity solutions essentially play a role. In second order case, this question is rather open even in case when the equation is linear in $\nabla^2 w$. Also, in first order case, if the Hamiltonian is non-convex, then it is hard to study the structure of the above stationary problem and it is rather open, as the weak KAM theory does not work well under such situation.

5.2 Some open problems.

We conclude this paper to give a couple of open problems.

Problem 1. *Show the full convergence* (5-1) *even if the equation is isotropic like*

$$u_t - (\operatorname{div}(\nabla u/|\nabla u|) + 1)\,|\nabla u| = r(x).$$

Study the uniqueness set for equation $R + F(\nabla w, \nabla^2 w) = r(x)$, the dependence of w in (5-1) *on the initial data u_0 in the case of second order equations.*

Problem 2. *Show the existence of R when r is discontinuous. Study how R depends on r both qualitatively and quantitively.*

In crystal growth problems, it is important to know how the growth rate depends on configuration of sources, i.e., geometric configuration of E when $r = c(x)1_E$ with some positive function depending on x; see Ohtsuka, Tsai, and Y. Giga [2015], Ohtsuka, Tsai, and Y. Giga [2018] for spiral growth. Several examples are studied in Y. Giga, Mitake, Ohtsuka, and Tran [n.d.].

Problem 3. *How regular is the solution u when $r \geq 0$ is regular and initial data is zero?*

These problems are very natural goals to derive unscaled asymptotic profile. The next problems are related to crystalline flow.

Problem 4 (Crystalline flow)**.** *Prove that if u_j is a viscosity solution of* (3-1) *with crystalline γ, so is its locally uniform limit u as $j \to \infty$.*

This is only proved for $N = 2$ in M.-H. Giga and Y. Giga [2001]. The problem for $N \geq 3$ is that definition in Y. Giga and Požár [2016], Y. Giga and Požár [2018] is not stable under such a limiting procedure. Once this is settled, an explicit solution given in Theorem 3.2 is also the maximal viscosity solution for crystalline case when $N \geq 3$.

Problem 5. *It seems that the spatially Lipschitz bound should be true for crystalline spreading law. Extend Lemma 3.4 to crystalline case.*

If so, this would yield the existence of the asymptotic speed for crystalline case.

We conclude this paper by pointing out that there are several potential applications of birth and spread models to other fields not limited in the field of crystal growth by considering various spreading laws. In this paper, we consider the spreading law $V = M(\mathbf{n})(aH_\gamma + \sigma)$ but it is interesting to consider more general spreading law as $V = g(\mathbf{n}, H_\gamma)$. For example, in Y. Giga, Mitake, Ohtsuka, and Tran [n.d.] formation of volcano profile is explained by taking inverse curvature like flow as the spreading law. It is worth to study above problems in these more general setting.

Acknowledgments. The author is grateful to Professor Hiroyoshi Mitake for valuable comments.

References

Sigurd Angenent and Morton E. Gurtin (1989). "Multiphase thermomechanics with interfacial structure. II. Evolution of an isothermal interface". *Arch. Rational Mech. Anal.* 108.4, pp. 323–391. MR: 1013461 (cit. on p. 2307).

G. Barles and Panagiotis E. Souganidis (2000). "On the large time behavior of solutions of Hamilton-Jacobi equations". *SIAM J. Math. Anal.* 31.4, pp. 925–939. MR: 1752423 (cit. on p. 2325).

Guy Barles (2013). "An introduction to the theory of viscosity solutions for first-order Hamilton-Jacobi equations and applications". In: *Hamilton-Jacobi equations: approximations, numerical analysis and applications*. Vol. 2074. Lecture Notes in Math. Springer, Heidelberg, pp. 49–109. MR: 3135340 (cit. on pp. 2320, 2321).

W. K. Burton, N. Cabrera, and F. C. Frank (1951). "The growth of crystals and the equilibrium structure of their surfaces". *Philos. Trans. Roy. Soc. London. Ser. A.* 243, pp. 299–358. MR: 0043005 (cit. on p. 2308).

Filippo Cagnetti, Diogo Gomes, Hiroyoshi Mitake, and Hung V. Tran (2015). "A new method for large time behavior of degenerate viscous Hamilton-Jacobi equations with convex Hamiltonians". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 32.1, pp. 183–200. MR: 3303946 (cit. on p. 2325).

Antonin Chambolle, Massimiliano Morini, Matteo Novaga, and Marcello Ponsiglione (Feb. 2017). "Existence and uniqueness for anisotropic and crystalline mean curvature flows". arXiv: 1702.03094 (cit. on p. 2308).

Antonin Chambolle, Massimiliano Morini, and Marcello Ponsiglione (2017). "Existence and uniqueness for a crystalline mean curvature flow". *Comm. Pure Appl. Math.* 70.6, pp. 1084–1114. MR: 3639320 (cit. on p. 2308).

Yun Gang Chen, Yoshikazu Giga, and Shun'ichi Goto (1991). "Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations". *J. Differential Geom.* 33.3, pp. 749–786. MR: 1100211 (cit. on p. 2307).

Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions (1992). "User's guide to viscosity solutions of second order partial differential equations". *Bull. Amer. Math. Soc. (N.S.)* 27.1, pp. 1–67. MR: 1118699 (cit. on p. 2307).

Andrea Davini and Antonio Siconolfi (2006). "A generalized dynamical approach to the large time behavior of solutions of Hamilton-Jacobi equations". *SIAM J. Math. Anal.* 38.2, pp. 478–502. MR: 2237158 (cit. on p. 2325).

L. C. Evans and J. Spruck (1991). "Motion of level sets by mean curvature. I". *J. Differential Geom.* 33.3, pp. 635–681. MR: 1100206 (cit. on p. 2307).

Lawrence C. Evans (2010). "Adjoint and compensated compactness methods for Hamilton-Jacobi PDE". *Arch. Ration. Mech. Anal.* 197.3, pp. 1053–1088. MR: 2679366 (cit. on p. 2325).

Albert Fathi (1998). "Sur la convergence du semi-groupe de Lax-Oleinik". *C. R. Acad. Sci. Paris Sér. I Math.* 327.3, pp. 267–270. MR: 1650261 (cit. on pp. 2313, 2325).

Mi-Ho Giga and Yoshikazu Giga (1998). "Evolving graphs by singular weighted curvature". *Arch. Rational Mech. Anal.* 141.2, pp. 117–198. MR: 1615520 (cit. on pp. 2308, 2315).

– (2001). "Generalized motion by nonlocal curvature in the plane". *Arch. Ration. Mech. Anal.* 159.4, pp. 295–333. MR: 1860050 (cit. on pp. 2307, 2317, 2326).

Mi-Ho Giga and Yoshikazu Giga (2013). "On the role of kinetic and interfacial anisotropy in the crystal growth theory". *Interfaces Free Bound.* 15.4, pp. 429–450. MR: 3148629 (cit. on p. 2310).

Mi-Ho Giga, Yoshikazu Giga, and Norbert Požár (2013). "Anisotropic total variation flow of non-divergence type on a higher dimensional torus". *Adv. Math. Sci. Appl.* 23.1, pp. 235–266. MR: 3155453 (cit. on p. 2308).

– (2014). "Periodic total variation flow of non-divergence type in $\mathbb{R}^n$". *J. Math. Pures Appl. (9)* 102.1, pp. 203–233. MR: 3212254 (cit. on p. 2308).

Yoshikazu Giga (2006). *Surface evolution equations*. Vol. 99. Monographs in Mathematics. A level set approach. Birkhäuser Verlag, Basel, pp. xii+264. MR: 2238463 (cit. on pp. 2307, 2312, 2314, 2318, 2321–2323).

Yoshikazu Giga and Nao Hamamuki (2013). "Hamilton-Jacobi equations with discontinuous source terms". *Comm. Partial Differential Equations* 38.2, pp. 199–243. MR: 3009078 (cit. on pp. 2311–2313).

Yoshikazu Giga, Qing Liu, and Hiroyoshi Mitake (2012). "Large-time asymptotics for one-dimensional Dirichlet problems for Hamilton-Jacobi equations with noncoercive Hamiltonians". *J. Differential Equations* 252.2, pp. 1263–1282. MR: 2853538 (cit. on p. 2313).

– (2014). "Singular Neumann problems and large-time behavior of solutions of noncoercive Hamilton-Jacobi equations". *Trans. Amer. Math. Soc.* 366.4, pp. 1905–1941. MR: 3152717 (cit. on p. 2313).

Yoshikazu Giga, H. Mitake, T. Ohtsuka, and H. V. Tran (n.d.). "Existence of asymptotic speed for birth and spread type nonlinear partial differential equations". In preparation (cit. on pp. 2309, 2317, 2320, 2326).

Yoshikazu Giga, Hiroyoshi Mitake, and Hung V. Tran (2016). "On asymptotic speed of solutions to level-set mean curvature flow equations with driving and source terms". *SIAM J. Math. Anal.* 48.5, pp. 3515–3546. MR: 3562367 (cit. on pp. 2309, 2317, 2322).

Yoshikazu Giga and N. Požár (2018). "Approximation of general facets by regular facets with respect to anisotropic total variation energies and its application to the crystalline mean curvature flow". *Comm. Pure Appl. Math.* (Cit. on pp. 2308, 2314, 2315, 2326).

Yoshikazu Giga and Norbert Požár (2016). "A level set crystalline mean curvature flow of surfaces". *Adv. Differential Equations* 21.7-8, pp. 631–698. MR: 3493931 (cit. on pp. 2308, 2314, 2315, 2326).

Morton E. Gurtin (1993). *Thermomechanics of evolving phase boundaries in the plane*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, pp. xi+148. MR: 1402243 (cit. on pp. 2307, 2323).

Nao Hamamuki (2013). "On large time behavior of Hamilton-Jacobi equations with discontinuous source terms". In: *Nonlinear analysis in interdisciplinary sciences—modellings,* *theory and simulations*. Vol. 36. GAKUTO Internat. Ser. Math. Sci. Appl. Gakkōtosho, Tokyo, pp. 83–112. MR: 3205346 (cit. on p. 2312).

Hitoshi Ishii (2008). "Asymptotic solutions for large time of Hamilton-Jacobi equations in Euclidean n space". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 25.2, pp. 231–266. MR: 2396521 (cit. on p. 2325).

Hitoshi Ishii, Gabriel E. Pires, and Panagiotis E. Souganidis (1999). "Threshold dynamics type approximation schemes for propagating fronts". *J. Math. Soc. Japan* 51.2, pp. 267–308. MR: 1674750 (cit. on pp. 2310, 2324).

H. Mitake and H. V. Tran (2017). "Dynamical properties of Hamilton-Jacobi equations via the nonlinear adjoint method". In: *Dynamical and Geometric Aspects of Hamilton-Jacobi and Linearized Monge-Ampère Equations, VIASM 2016*. Ed. by H. Mitake and H. V. Tran. Vol. 2183. Lecture Notes in Math. Springer, pp. 125–228 (cit. on pp. 2309, 2325).

Gawtum Namah and Jean-Michel Roquejoffre (1999). "Remarks on the long time behaviour of the solutions of Hamilton-Jacobi equations". *Comm. Partial Differential Equations* 24.5-6, pp. 883–893. MR: 1680905 (cit. on pp. 2313, 2325).

M. Ohara and R. C. Reid (1973). *Modeling crystal growth rates from solution*. Englewood Cliffs, NJ: Prentice–Hall (cit. on p. 2308).

T. Ohtsuka, Y.-H. R. Tsai, and Yoshikazu Giga (2015). "A level set approach reflecting sheet structure with single auxiliary function for evolving spirals on crystal surfaces". *J. Sci. Comput.* 62.3, pp. 831–874. MR: 3304334 (cit. on pp. 2308, 2326).

– (2018). "Growth rate of crystal surfaces with several dislocation centers". *Crystal Growth* (cit. on pp. 2308, 2326).

T. P. Schulze and R. V. Kohn (1999). "A geometric model for coarsening during spiral-mode growth of thin films". *Phys. D* 132.4, pp. 520–542. MR: 1704821 (cit. on pp. 2309, 2311).

Peter Smereka (2000). "Spiral crystal growth". *Phys. D* 138.3-4, pp. 282–301. MR: 1744632 (cit. on p. 2308).

Halil Mete Soner (1993). "Motion of a set by the curvature of its boundary". *J. Differential Equations* 101.2, pp. 313–372. MR: 1204331 (cit. on pp. 2308, 2323).

Jean E. Taylor (1991). "Constructions and conjectures in crystalline nondifferential geometry". In: *Differential geometry*. Vol. 52. Pitman Monogr. Surveys Pure Appl. Math. Longman Sci. Tech., Harlow, pp. 321–336. MR: 1173051 (cit. on p. 2307).

Jack Xin and Yifeng Yu (2013). "Sharp asymptotic growth laws of turbulent flame speeds in cellular flows by inviscid Hamilton-Jacobi models". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 30.6, pp. 1049–1068. MR: 3132416 (cit. on p. 2309).

– (2014). "Asymptotic growth rates and strong bending of turbulent flame speeds of G-equation in steady two-dimensional incompressible periodic flows". *SIAM J. Math. Anal.* 46.4, pp. 2444–2467. MR: 3228465 (cit. on p. 2309).

Etsuro Yokoyama, Yoshikazu Giga, and Piotr Rybka (2008). "A microscopic time scale approximation to the behavior of the local slope on the faceted surface under a nonuniformity in supersaturation". *Phys. D* 237.22, pp. 2845–2855. MR: 2514066 (cit. on p. 2313).

Received 2018-01-17.

Yoshikazu Giga (儀我美一)
Graduate School of Mathematical Sciences
University of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153-8914
Japan
labgiga@ms.u-tokyo.ac.jp

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2329–2356)

A PANORAMA OF SINGULAR SPDES

Massimiliano Gubinelli

Abstract

I will review the setting and some of the recent results in the field of singular stochastic partial differential equations (SSPDEs). Since Hairer's invention of regularity structures this field has experienced a rapid development. SSPDEs are non-linear equations with random and irregular source terms which make them ill-posed in classical sense. Their study involves a tight interplay between stochastic analysis, analysis of PDEs (including paradifferential calculus) and algebra.

1 Introduction

This contribution aims to give an overview of the recent developments at the interface between stochastic analysis and PDE theory where a series of new tools have been put in place to analyse certain classes of stochastic PDEs (SPDEs) whose rigorous understanding was, until recently, very limited. Typically these equations are non-linear and the randomness quite ill behaved from the point of view of standard functional spaces. In the following I will use the generic term *singular stochastic PDEs* (SSPDEs) to denote these equations.

The interplay between the algebraic structure of the equations, the irregular behaviour of the randomness and the weak topologies needed to handle such behaviour provide a fertile ground where new point of views have been developed and old tools put into work in new ways Hairer [2014], Gubinelli, Imkeller, and Perkowski [2015], Otto and Weber [2016], Kupiainen [2016], Bailleul and Bernicot [2016a], Bruned, Hairer, and Zambotti [2016], Chandra and Hairer [2016], and Bruned, Chandra, Chevyrev, and Hairer [2017].

MSC2010: primary 60H15; secondary 35S50.
Keywords: Singular SPDEs, Rough paths, paraproducts, regularity structures, Kardar-Parisi-Zhang equation, stochastic quantisation.

2 Ways of describing a function

2.1 From ODEs to rough differential equations. The simpler setting we can discuss is that of an ordinary differential equation perturbed by a random function in a non-linear way. Consider the Cauchy problem for $y : \mathbb{R}_+ \to \mathbb{R}^d$,

$$\begin{cases} \dot{y}(t) = \varepsilon^{1/2} f(y(t))\eta(t), & t > 0, \\ y(0) = y_0 \in \mathbb{R}^d \end{cases} \tag{1}$$

where the dot denotes time derivative, $f : \mathbb{R}^d \to \mathfrak{L}(\mathbb{R}^n; \mathbb{R}^d)$ is a family of smooth vector fields in $\mathbb{R}^d$ ($\mathfrak{L}(\mathbb{R}^n; \mathbb{R}^d)$ are the linear maps form $\mathbb{R}^n$ to $\mathbb{R}^d$), $\eta : \mathbb{R}_+ \to \mathbb{R}^n$ is a smooth centered $\mathbb{R}^m$–valued Gaussian random function and $\varepsilon > 0$ a small parameter.

If we are interested in the $\varepsilon \to 0$ limit of this equation we would better rescale it to see some interesting dynamics going on. In term of the rescaled variable $y_\varepsilon(t) = y(t/\varepsilon)$ the equation has the form

$$\begin{cases} \dot{y}_\varepsilon(t) = f(y_\varepsilon(t))\eta_\varepsilon(t), & t > 0, \\ y_\varepsilon(0) = y_0 \in \mathbb{R}^d \end{cases} \tag{2}$$

where $\eta_\varepsilon(t) = \varepsilon^{-1/2}\eta(t/\varepsilon)$. If we assume that η is stationary, has fast decaying correlations (e.g. exponentially fast) and independent components, then we can prove that η_ε converges in law to a white noise ξ, that is the Gaussian random distribution with covariance given by

$$\mathbb{E}[\xi(t)\xi(s)] = \delta(t - s).$$

This convergence takes place as random elements of the Hölder–Besov space[1] $\mathcal{C}^\alpha = B^\alpha_{\infty,\infty}$ for any $\alpha < -1/2$. Alternatively, and more in line with classical probability theory, one could look at the integral function $x_\varepsilon(t) = \int_0^t \eta_\varepsilon(s)\mathrm{d}s$ and conclude that it converges in the Hölder topology $\mathcal{C}^{\alpha+1}$ to the Brownian motion.

This procedure is reminiscent of homogenisation E [2011] but while there there are essentially only two (or a finite number) of scales which play a fundamental role here all the scales remain coupled also after the passage to the limit. Indeed one would now like to argue that the solution y_ε of Equation (2) converges to the solution z of the ODE

$$\begin{cases} \dot{z}(t) = f(z(t))\xi(t), & t > 0, \\ z(0) = y_0 \in \mathbb{R}^d. \end{cases} \tag{3}$$

where ξ is the white noise on $\mathbb{R}$. However we quickly realise that this equation is not well posed. Indeed we cannot hope better regularity for z than $\mathcal{C}^{1+\alpha}$ (e.g. in the simple

[1] The choice of this space is not canonical for this convergence in law but will fit our intended applications, other choices do not lead to substantial improvements in the arguments below.

setting where f is constant) and in this situation the pointwise product of $f(z)$ (still a $\mathcal{C}^{1+\alpha}$ function) and the distribution ξ of regularity $\mathcal{C}^{\alpha}$ is not a well defined operation.

Remark 2.1. *That this is not only a technical difficulty can be understood easily by considering the following example. Take $f_\varepsilon(t) = \varepsilon^{1/2}\sin(t/\varepsilon)$ and $g_\varepsilon(t) = \varepsilon^{-1/2}\sin(t/\varepsilon)$. Then for any $\alpha < -1/2$, $f_\varepsilon \to 0$ in $\mathcal{C}^{\alpha+1}$ and $g_\varepsilon \to 0$ in $\mathcal{C}^{\alpha}$ but $h_\varepsilon(t) := f_\varepsilon(t)g_\varepsilon(t) = \sin^2(t/\varepsilon) = 1 - \cos(2t/\varepsilon)/2$ and $h_\varepsilon \to 1$ in $\mathcal{C}^{2\alpha+1}$. We see that the product cannot be extended continuously in $\mathcal{C}^{\alpha+1} \times \mathcal{C}^{\alpha}$ as we would need to have a robust meaning for Equation (3) in the framework of Hölder–Besov spaces.*

This difficulty has been realised quite early in stochastic analysis and is at the origin of the invention of stochastic calculus by Itō (and the independent work of Dœblin) and has shaped ever since the study of stochastic processes, see e.g. Watanabe and Ikeda [1981] and Revuz and Yor [2004]. Itō's approach give a meaning to (3) by prescribing a certain preferred approximation scheme (the forward Riemman sum) to the integral version of the r.h.s. of the equation. The resulting Itō integral comes with estimates which are at the core of stochastic integration theory. However the Itō integral in *not* the right description for the limiting Equation (3). Indeed Wong and Zakai proved that the limit is given by another interpretation of the product, that provided by the Stratonovich integral.

2.2 Reconstruction of a coherent germ. From a strictly analytic viewpoint, without resorting to probabilistic techniques, Equation (3) should stand for a description of the possible limit points of the sequence $(y_\varepsilon)_\varepsilon$. Compactness arguments should provide methods to prove limits exists along subsequence and under nice conditions we would hope to be able to prove that there is only such a limit point, settling the problem of the convergence of the whole sequence. This is the standard approach. Usually in PDE theory some refined methods are put in place in order to establish some sort of compactness (e.g. convexity, maximum principle, concentrated or compensated compactness, Young measures), others maybe are needed to show uniqueness (e.g. entropy solutions, energy estimates, viscosity solutions), see e.g. Evans [1990]. But here we are stuck at a more primitive level, we do not have an effective description to start with, and this inability comes (as is fast to realise) with the inability to obtain useful and general apriori bounds for the compactness step.

With the aim of identifying such a description we could think of constructing a good local approximations of the function z. Around a given time $s > 0$ we imagine that z has the behaviour obtained by freezing the vectorfield f at that given time: write the ODE in the weak form and assume that for any smooth, compactly supported $\varphi : \mathbb{R}_+ \to \mathbb{R}$ we have

$$\left|\int_{\mathbb{R}} \varphi_s^\lambda(t)(\dot z(t) - f(z(s))\xi(t))\mathrm{d}t\right| \leqslant \lambda^\gamma N(\varphi) \tag{4}$$

where $\varphi_s^\lambda(t) := \lambda^{-1}\varphi((t-s)/\lambda)$ is a rescaled and centred version of φ, $N(\varphi), \gamma > 0$ constants not depending on λ or s (we let $z(t) = y_0$ and $\xi(t) = 0$ if $t < 0$). If ξ is a smooth function then this bound implies Equation (3) since $\varphi^\lambda \to \delta$ as $\lambda \to 0$, so this description of z is as detailed as the ODE. On the other hand it has the fundamental advantage of having *decoupled* the product $f(z(t))\xi(t)$ into $f(z(s))\xi(t)$. Equation (4) makes sense even when ξ is a distribution and in particular in the case of the white noise we are looking after. That this description is quite powerful is witnessed by Hairer's *reconstruction theorem* Hairer [2014].

Theorem 2.2 (Hairer's reconstruction). *Let $\gamma > 0$, $\alpha \in \mathbb{R}$ and $G = (G_x)_{x\in\mathbb{R}^d}$ be a family of distributions in $\mathcal{S}'(\mathbb{R}^d)$ such that there exists a constant $L(G)$ and a constant $N(\varphi)$ for which*

$$\sup_x |G_x(\varphi_x^\lambda)| \leqslant \lambda^{-\alpha} N(\varphi) L(G), \qquad \lambda \in (0,1) \tag{5}$$

$$|G_y(\varphi_x^\lambda) - G_x(\varphi_x^\lambda)| \leqslant \lambda^\gamma N(\varphi) L(G) P(|x-y|/\lambda), \qquad \lambda \in (0,1) \tag{6}$$

where P is a continuous function with at most polynomial growth. Then there exists a universal constant C_γ and a unique distribution $g = \mathcal{R}(G) \in \mathcal{S}'(\mathbb{R}^d)$, the reconstruction of G, such that

$$|g(\varphi_x^\lambda) - G_x(\varphi_x^\lambda)| \leqslant C_\gamma \lambda^\gamma N(\varphi) L(G), \qquad \lambda \in (0,1).$$

For the sake of the exposition I simplified a bit the setting and gave a slightly different formulation of this results which can be appreciated independently of other details of Hairer's theory of regularity structures Hairer [ibid.]. We give here an idea of proof, without pretension to make it fully rigorous. We call the family G a *germ* and the quantity in the l.h.s. of Equation (6) the *coherence* of the germ G. The theorem states a relation between coherent germs and distributions.

Proof. (sketch) *Uniqueness.* Assume $g, \tilde g$ are two possible reconstructions of G, then

$$|g(\varphi_x^\lambda) - \tilde g(\varphi_x^\lambda)| \lesssim \lambda^\gamma.$$

For any given test function $\psi \in \mathcal{S}(\mathbb{R}^d)$ we let $T_\lambda\psi(y) = \int \psi(x)\varphi_x^\lambda(y)\mathrm{d}x$. Then $T_\lambda\psi \to \psi$ in $\mathcal{S}(\mathbb{R}^d)$ and $g = \tilde g$ since

$$|(g-\tilde g)(\psi)| = \lim_{\lambda\to 0} |(g-\tilde g)(T_\lambda\psi)| \lesssim \liminf_{\lambda\to 0} \int |g(\varphi_x^\lambda) - \tilde g(\varphi_x^\lambda)||\psi(x)|\mathrm{d}x = 0.$$

Existence. We follow an idea of Otto and Weber Otto and Weber [2016]. Introduce the heat semigroup to perform a multiscale decomposition. Let $T_i = P_{2^{-i}}$ where $(P_t)_t$ is the heat kernel, then $T_{i+1}T_{i+1} = T_i$ for $i \geqslant 0$. Let

$$\mathcal{R}_N G(x) := \int_{y,z} T_N(x-y)T_N(y-z)G_y(z),$$

where $\int_{y,z}$ denotes the integral in (y,z) over $\mathbb{R}^d \times \mathbb{R}^d$. Note that if $G_y(x) = f(x)$ for some Schwartz distribution $f \in \mathcal{S}'(\mathbb{R}^d)$ then $\mathcal{R}_N G(x) = \int_{y,z} T_N(x-y)T_N(y-z)f(z) = f(T_{N+1}(x-\cdot)) \to f(x)$ in $\mathcal{S}'(\mathbb{R}^d)$. In general, in order to control the limit $\mathcal{R} = \lim_{N\to\infty} \mathcal{R}_N$ for more general germs we look at

$$(\mathcal{R}_{n+1}G - \mathcal{R}_n G)(x) := \mathcal{A}_n G(x) + \mathcal{B}_n G(x),$$

where

$$\begin{aligned} \mathcal{A}_n G(x) &:= \textstyle\int_{y,z}(T_{n+1}-T_n)(x-y)(T_n+T_{n+1})(y-z)G_y(z), \\ \mathcal{B}_n G(x) &:= \textstyle\int_{y,z,r} T_{n+1}(x-r)T_{n+1}(r-y)T_{n+1}(y-z)(G_y(z)-G_r(z)). \end{aligned}$$

Using (6) the terms $\mathcal{B}_n G(x)$ can be estimated by $|\mathcal{B}_n G(x)| \lesssim 2^{-n\gamma} N(\varphi)L(G)$, and they can be resummed over n since $\gamma > 0$. The terms $\mathcal{A}_n G$ are localized at scale 2^{-n} thanks to the factor $(T_{n+1}-T_n)$ and they behave as "orthogonal" contributions: once tested agains a test function ψ they can be estimated as $|\mathcal{A}_n G(\psi)| \lesssim \|(T_{n+1}-T_n)\psi\|_{L^1} 2^{-\alpha n} N(\varphi)L(G)$, thanks to the Equation (5). From we deduce that

$$\sum_n |\mathcal{A}_n G(\psi)| \lesssim N(\varphi)L(G)\sum_n \|(T_{n+1}-T_n)\psi\|_{L^1}2^{-\alpha n} \lesssim N(\varphi)L(G)\|\psi\|_{B^{\alpha}_{1,1}},$$

where $B^{\alpha}_{1,1}$ is the Besov space with norm $\|\psi\|_{B^{\alpha}_{1,1}} = \sum_{n\geqslant 1} 2^{-\alpha n}\|\Delta_n \psi\|_{L^1}$. From these observations is easy to deduce that $\mathcal{R}_N G \to \mathcal{R}G$ as a distribution. In order to identify $\mathcal{R}G$ we observe that, for fixed $L > 0$,

$$\mathcal{R}_L G(x) - \mathcal{R}_L G_h(x) = \int_{y,z} T_L(x-y)T_L(y-z)(G_y(z)-G_h(z)),$$

and if $|x-h| \simeq 2^{-L}$ we have $|\mathcal{R}_L G(x) - \mathcal{R}_L G_h(x)| \lesssim 2^{-L\gamma}$, while

$$\begin{aligned} \mathcal{R}G(x) - G_h(x) &= \mathcal{R}(G-G_h)(x) \\ &= \mathcal{R}_L(G-G_h)(x) + \sum_{n>L}\mathcal{A}_n(G-G_h)(x) + \sum_{n>L}\mathcal{B}_n G(x) \end{aligned}$$

It is not difficult to estimate $|\mathcal{A}_n(G-G_h)(\psi_h^{2^{-L}})| \lesssim 2^{-n\gamma}$ and finally deduce that $|(\mathcal{R}G - G_h)(\psi_h^{2^{-L}})| \lesssim 2^{-L\gamma}$. □

Let us go back to our equation. Let $G_s(t) = f(z(s))\xi(t)$ be our germ and consider its coherence:

$$G_u(\varphi_s^\lambda)-G_s(\varphi_s^\lambda) = \int_t \varphi^\lambda(t-s)(f(z(u))-f(z(s)))\xi(t) = (f(z(u))-f(z(s)))\xi(\varphi_s^\lambda).$$

Assuming that $z \in \mathcal{C}^{\alpha+1}$ and that $\xi \in \mathcal{C}^\alpha$ we have, for some polynomially growing P,

$$|G_u(\varphi_s^\lambda) - G_s(\varphi_s^\lambda)| \lesssim \lambda^{2\alpha+1} \|z\|_{\mathcal{C}^{\alpha+1}} \|\xi\|_{\mathcal{C}^\alpha} P((u-s)/\lambda).$$

We see that if $\gamma = 2\alpha - 1 > 0$ we can meet the conditions of Theorem 2.2. In this case the ODE can be replaced by the formulation (4) and the resulting theory coincides with the theory of differential equations build upon the Young integral Young [1936] and P. K. Friz and Hairer [2014].

This is not yet enough for us. White noise restrict the allowed values for α in the range $\alpha < 1/2$ and in this case $2\alpha - 1 < 0$. In this case the description is not precise enough to uniquely determine the distribution $\dot{z}(t)$ using only the assumption $z \in \mathcal{C}^\alpha$. Going back to the ODE and thinking about a Taylor expansion for the r.h.s. we come up with a refined description of the solution given by the new germ:

$$G_s(t) = f(z(s))\xi(t) + f'(z(s))f(z(s)) \int_s^t \xi(u)\mathrm{d}u,$$

where we denoted f' the gradient of f. Its coherence is given by (we let $f_2(z) = f'(z)f(z)$)

$$\begin{aligned} G_u(\varphi_s^\lambda) - G_s(\varphi_s^\lambda) \;&=\; \left[f(z(u)) - f(z(s)) - f_2(z(s)) \textstyle\int_s^u \xi(r)\mathrm{d}r\right]\xi(\varphi_s^\lambda) \\ &\quad + [f_2(z(u)) - f_2(z(s))]\left[\textstyle\int_t \varphi^\lambda(t-s)\left(\int_s^t \xi(r)\mathrm{d}r\right)\xi(t)\right]. \end{aligned}$$

In order to meet the conditions of the reconstruction theorem we can require

$$\left|\int_t \varphi^\lambda(t-s)\left(\int_s^t \xi(r)\mathrm{d}r\right)\xi(t)\right| \lesssim \lambda^{2\alpha+1}, \tag{7}$$

and

$$\left|f(z(u)) - f(z(s)) - f_2(z(s))\int_s^u \xi(r)\mathrm{d}r\right| \lesssim \lambda^{2\alpha+2}, \tag{8}$$

from which we see that $|G_u(\varphi_s^\lambda) - G_s(\varphi_s^\lambda)| \lesssim \lambda^{3\alpha+2}$. Provided $\alpha > 3/2$ we can reconstruct in a unique way a distribution g from this germ and verify the equation $\dot{z} = g$

(at least in the weak sense). Equation (7) is a condition on ξ, Equation (8) one on z. In particular, by Taylor expansion, this latter holds if the bound

$$\left| z(u) - z(s) - f(z(s)) \int_s^u \xi(r)\mathrm{d}r \right| \lesssim \lambda^{2\alpha+2}, \tag{9}$$

holds for z. This is a refinement of the Hölder assumption $z \in \mathcal{C}^{\alpha+1}$. Building on these basic observation is possible to develop a complete well–posedness theory showing that, provided $\alpha > -3/2$ there is a continuous map $\Phi : \Xi \mapsto z$ taking the germ

$$\Xi_s(t) = (\Xi_s^{(1)}(t), \Xi_s^{(2)}(t)) = \left(\xi(t), \left(\int_s^t \xi(r)\mathrm{d}r\right)\xi(t)\right)$$

satisfying $|\Xi_s^{(1)}(\varphi_s^\lambda)| \lesssim \lambda^\alpha$ and $|\Xi_s^{(2)}(\varphi_s^\lambda)| \lesssim \lambda^{2\alpha+1}$ to the unique Hölder function z satisfying the relation

$$\left| \int_{\mathbb{R}} \varphi_s^\lambda(t) \left(\dot{z}(t) - f(z(s))\xi(t) - f_2(z(s)) \int_s^t \xi(r)\mathrm{d}r \right) \mathrm{d}t \right| \lesssim \lambda^{3\alpha+2}.$$

The original difficulties are here not completely solved, indeed the germ $\Xi^{(2)}$ is not apriori well defined given that it contains a pointwise product between the distribution ξ and the function $\int_s^{\cdot} \xi(r)\mathrm{d}r$. However in this new perspective we have accomplished a major step: restricting the difficulty to a well defined quantity which can be analysed from the point of view of stochastic analysis without any reference to the ODE problem and its non-linearity.

The map Φ is called Itō–Lyons map P. K. Friz and Hairer [2014]. Its regularity and the fact that it provides an extension of the solution map for the classical ODE allows to control the limit of the Equation (2) as $\varepsilon \to 0$. Indeed *provided* we can show that the germ

$$\Xi_s^\varepsilon(t) = \left(\eta_\varepsilon(t), \left(\int_s^t \eta_\varepsilon(r)\mathrm{d}r\right)\eta(t)\right) \tag{10}$$

converges in the appropriate topology to Ξ then we can conclude that $y_\varepsilon = \Phi(\Xi^\varepsilon) \to \Phi(\Xi) = z$.

In this generalisation however there is a catch. The limiting problem is now defined in terms of a more complex object Ξ than the original white noise ξ. Any two sequences $(\eta_\varepsilon)_\varepsilon$ and $(\tilde{\eta}_\varepsilon)_\varepsilon$ approximating ξ and lifted into germs Ξ^ε and $\tilde{\Xi}^\varepsilon$ can converge to two limits Ξ and $\tilde{\Xi}$ for which $\Xi^{(1)} = \tilde{\Xi}^{(1)} = \xi$ but $\Xi^{(2)} \neq \tilde{\Xi}^{(2)}$. In this case the corresponding solutions y_ε and $\tilde{y}_\varepsilon$ to Equation (2) will in general converge to different limits $z = \Phi(\Xi)$, $\tilde{z} = \Phi(\tilde{\Xi})$.

2.3 From ODEs to PDEs. We have given an outlook of the use of Hairer's reconstruction theorem in the analysis of a controlled ODE. The Itō–Lyons map has been invented and initially studied by Lyons T. Lyons [1998], T. Lyons and Qian [2002], T. J. Lyons, Caruana, and Lévy [2007], and P. K. Friz and Hairer [2014] and is at the base of the *theory of rough paths* (RPT). Lyons' theory goes far beyond to the limit $\alpha > -3/2$ down to any $\alpha > -1$. This full range corresponds to the *subcritical* regime where scaling dictates that the noise is a perturbation of the first order differential operator ∂_t. The reformulation given here is essentially that introduced in the context of RPT by Davie Davie [2007] and here reshaped in the language of Theorem 2.2. Functions satisfying conditions like eq. (9) are called *controlled paths* in RPT and were introduced in Gubinelli [2004] in order to provide a nice analytical setting for a fixed point argument leading to the Itō–Lyons map and moreover to decouple the reconstruction of germs from the construction of the fixpoint. In the case of ODEs one can avoid the use of the reconstruction theorem by using the *sewing lemma* Gubinelli [2004], Feyel and De La Pradelle [2006], and Feyel, de La Pradelle, and Mokobodzki [2008]:

Lemma 2.3 (Sewing lemma)**.** *Let $\gamma > 0$. Let $G : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ a function such that*

$$|G(s,t) - G(s,u) - G(u,t)| \leqslant L_G |t-s|^{1+\gamma}, \qquad s \leqslant u \leqslant t.$$

for some $L_G > 0$. Then there exists a unique function $g : \mathbb{R}_+ \to \mathbb{R}$ such that

$$|g(t) - g(s) - G(s,t)| \leqslant C_\gamma L_G |t-s|^{1+\gamma}$$

for a universal constant C_γ.

Sewing and reconstruction are not equivalent. The sewing lemma combines in one operation the reconstruction operated by Theorem 2.2 and the integration needed to pass from $\dot{z}$ and z.

Hairer's regularity structure theory Hairer [2014] and P. K. Friz and Hairer [2014] builds over Theorem 2.2 a vast generalisation of Lyons' rough path theory and provides a solution theory for a large class of subcritical parabolic SPDEs. A recent series of three other papers Bruned, Hairer, and Zambotti [2016], Chandra and Hairer [2016], and Bruned, Chandra, Chevyrev, and Hairer [2017] complete the construction of this theory by "automatizing" the lifting of all the structures needed to deal with the various aspects of the solution theory for a generic singular SPDE: the construction of the appropriate model and regularity structure, the stochastic estimates and the identification of a suitable class of regular equations which possess limits (i.e. that can be *renormalized*).

2.4 Paraproducts and the paracontrolled Ansatz. In Gubinelli, Imkeller, and Perkowski [2015] an alternative approach has been introduced to handle the difficult product in (3) by decoupling it according to a multiscale decomposition. Write

$$f(z(t))\xi(t) = \sum_{n,m\geqslant -1} \int_{s,r} K_n(t-s)K_m(t-r)f(z(s))\xi(r),$$

where we let $(K_n)_{n\geqslant -1}$ to be kernels of Littlewood–Paley (LP) type which provide a resolution of a given distribution into "blocks" with specific frequency localization. See Bahouri, Chemin, and Danchin [2011] and Gubinelli and Perkowski [2015] for details on LP decomposition, Besov spaces and for the paraproduct estimates discussed below. Writing Δ_n for the operator of convolution with the kernel K_n we can decompose the product of two distributions g, h as above into three contributions according to the case where $n \leqslant m-1$, $|n-m| \leqslant 1$ and $n \geqslant m+1$:

$$gh = g \varolessthan h + g \circledcirc h + g \varogreaterthan h,$$

where we let

$$g \varolessthan h = h \varogreaterthan g := \sum_{n<m-1} (\Delta_n g)(\Delta_n h), \quad g \circledcirc h := \sum_{|n-m|\leqslant 1} (\Delta_n g)(\Delta_n h).$$

These operators are well behaved in several function spaces. For illustrative purpose we will use them mainly in the Hölder–Besov spaces $\mathcal{C}^\alpha = B^\alpha_{\infty,\infty}$ but other choices are possible. The LP decomposition can be chosen in such a way that these operators can be extended to bilinear bounded operators in Hölder–Besov spaces according to the following estimates:

$$\begin{aligned}
\|g \varolessthan h\|_{\mathcal{C}^\alpha} &\lesssim \|g\|_{L^\infty}\|h\|_{\mathcal{C}^\alpha}, && \alpha \in \mathbb{R},\\
\|g \varolessthan h\|_{\mathcal{C}^{\alpha+\beta}} &\lesssim \|g\|_{\mathcal{C}^\beta}\|h\|_{\mathcal{C}^\alpha}, && \alpha \in \mathbb{R}, \beta < 0,\\
\|g \circledcirc h\|_{\mathcal{C}^{\alpha+\beta}} &\lesssim \|g\|_{\mathcal{C}^\beta}\|h\|_{\mathcal{C}^\alpha}, && \alpha + \beta > 0.
\end{aligned}$$

We see that the *resonant product* $g \circledcirc h$ is defined only for functions whose sum of regularities is positive while the *paraproduct* $g \varolessthan h$ is always well defined. Another key observation is that the paraproduct does not improve the regularity of its r.h.s. while the resonant product (when it is well defined) improves the regularity of its factor of lower regularity.

Paraproducts and related operations were introduced by Bony and Meyer Bony [1981] and Meyer [1981] for the use in the regularity theory of fully–nonlinear hyperbolic equation. It is not the aim of the present exposition to cover the vast literature these ideas generated, which includes the calculus of paradifferential operators. The reader can refer to Bahouri, Chemin, and Danchin [2011], Metivier [2008], Taylor [2000], Tao [2006], and Alinhac and Gérard [1991] for some expositions on the results and applications of these tools to PDEs.

An basic result in the theory of paraproducts is Bony's *paralinearization* Bony [1981], Bahouri, Chemin, and Danchin [2011], and Gubinelli, Imkeller, and Perkowski [2015] which in our context reads

$$z \in \mathcal{C}^{2\alpha} \mapsto R_f(z) := f(z) - f'(z) \varolessthan z \in \mathcal{C}^{2\alpha}, \qquad \alpha > 0. \tag{11}$$

Moreover we have a commutator lemma proved in Gubinelli, Imkeller, and Perkowski [2015].

Lemma 2.4. *If $\beta + \gamma + \delta > 0$ there exists a bounded trilinear operator $Q : \mathcal{C}^{\beta} \times \mathcal{C}^{\gamma} \times \mathcal{C}^{\delta} \to \mathcal{C}^{\gamma+\delta}$ such that for smooth function g, h, l*

$$Q(g, h, l) = (g \varolessthan h) \circledcirc l - g(h \circledcirc l).$$

Going back to our ODE we can expand its r.h.s. as

$$\dot{z} = f(z)\xi = f(z) \varolessthan \xi + f(z) \varogreaterthan \xi + f(z) \circledcirc \xi.$$

If we assume that $\xi \in \mathcal{C}^{\alpha}$ and $z \in \mathcal{C}^{1+\alpha}$ then $f(z) \varolessthan \xi \in \mathcal{C}^{\alpha}$ and $f(z) \circledast \xi \in \mathcal{C}^{2\alpha+1}$ (at least when $2\alpha + 1 > 0$). The key idea is to perform a change of variables to encode the heuristic that the more irregular contribution to $\dot{z}$ comes from the paraproduct $f(z) \varolessthan \xi$. We formulate a *paracontrolled Ansatz* by introducing a new unknown $z^{\sharp} \in \mathcal{C}^{2\alpha+2}$ such that

$$z = f(z) \varolessthan X + z^{\sharp}, s \tag{12}$$

where X solves the equation $\dot{X} = \xi$. Doing so gives

$$\dot{z}^{\sharp} = \dot{z} - f(z) \varolessthan \dot{X} - (\partial_t f(z)) \varolessthan X = f(z) \circledcirc \xi + f(z) \varogreaterthan \xi - (\partial_t f(z)) \varolessthan X.$$

From the *paralinearization* (11) and the commutator Lemma 2.4 follows that

$$f(z) \circledcirc \xi = f'(z)(z \circledcirc \xi) + Q(f'(z), z, \xi) + R_f(z) \circledcirc \xi,$$

in the sense that the difference $f(z) \circledcirc \xi - f'(z)(z \circledcirc \xi)$ is well defined as soon as $3\alpha + 2 > 0$. Recalling (12) we can further simplify this expression into

$$\begin{aligned} f(z) \circledcirc \xi = f'(z) f(z)(X \circledcirc \xi) + f'(z) Q(f(z), X, \xi) + f'(z)(z^{\sharp} \circledcirc \xi) + \\ + Q(f'(z), z, \xi) + R_f(z) \circledcirc \xi \end{aligned}$$

Finally our original ODE is transformed into the following equation for $z^{\sharp}$:

$$\dot{z}^{\sharp} = f_2(z)(X \circledcirc \xi) + \Psi(z, z^{\sharp}, \xi) \tag{13}$$

where we collected into Ψ all the less interesting contributions which are well under control (as the reader can check) assuming $z, z^\sharp, \xi$ have regularities $\alpha + 1, 2\alpha + 2, \alpha$.

The problematic term $X \circledcirc \xi$ here plays the role of the term (7) in the rough path approach. The paracontrolled Ansatz (12), the role of the Equation (9). If we assume that $X \circledcirc \xi \in \mathcal{C}^{2\alpha+1}$ (as scaling considerations and Equation (7) suggests) then Equation (13) is a well defined differential equation (non-local, with some low order paradifferential terms) which can be solved for $z^\sharp \in \mathcal{C}^{2\alpha+2}$. Technically, in order for (13) to be an equation for $z^\sharp \in \mathcal{C}^{2\alpha+2}$ we need to solve for z in (12) or to consider as unknown the system $(z, z^\sharp)$. Both approaches are possible provided small modifications are introduced in the considerations above. For more details Gubinelli, Imkeller, and Perkowski [2015] and Gubinelli and Perkowski [2015].

As a consequence we can identify the Itō–Lyons map Ψ as the map going from the *enhanced noise* $\Xi = (\xi, X \circledcirc \xi) \in \mathcal{C}^\alpha \times \mathcal{C}^{2\alpha+1}$ to the solution $z \in \mathcal{C}^{\alpha+1}$ via $Z = (z, z^\sharp) \in \mathcal{C}^{\alpha+1} \times \mathcal{C}^{2\alpha+2}$. As before this solution maps agrees with the solution of the ODE whenever ξ is smooth and can be used to control the limit $y_\varepsilon \to z$ provided we can prove that the enhanced noise $J(\eta_\varepsilon) := (\eta_\varepsilon, (\partial_t^{-1}\eta_\varepsilon) \circledcirc \eta_\varepsilon)$ converges in $\mathcal{C}^\alpha \times \mathcal{C}^{2\alpha+1}$ as $\varepsilon \to 0$ (recall that the assumption $3\alpha + 2 > 0$ is in force here).

We record these basic relations into the following diagram:

$$
\begin{tikzcd}
\varepsilon \to 0 & \eta_\varepsilon \arrow[r, leftrightarrow, "J"] \arrow[d] & \Xi_\varepsilon \arrow[r, mapsto, "\Psi"] \arrow[d] & Y_\varepsilon = (y_\varepsilon, y_\varepsilon^\sharp) \arrow[r, mapsto] \arrow[d] & y_\varepsilon \arrow[d] \\
& \xi & \Xi \arrow[l, mapsto] \arrow[r, mapsto, "\Psi"] & Z = (z, z^\sharp) \arrow[r, mapsto] & z \\
& \Cap & \Cap & \Cap & \Cap \\
& \mathcal{C}^\alpha & \mathcal{C}^\alpha \times \mathcal{C}^{2\alpha+1} & \mathcal{C}^{\alpha+1} \times \mathcal{C}^{2\alpha+2} & \mathcal{C}^{\alpha+1}
\end{tikzcd}
$$

The paracontrolled Ansatz transforms a problem of singular SPDEs into well–posed PDE problem featuring some paradifferential operators. The major drawback is that certain equations are out of reach for this technique, essentially because we understand quite poorly a systematic paradifferential development of generic non-linearities beyond the first order. Higher–order paralinearization has been investigated long ago by Chemin Chemin [1988a,b] and some higher–order commutators introduced in the work of Bailleul and Bernicot Bailleul and Bernicot [2016b,a] but the technical advantage over regularity structures tends to be less evident.

It should be remarked that the core of all these approaches (regularity structures, rough paths theory or paracontrolled distributions) lies in three basic steps:

a) Transform the original equation into a well-posed analytical problem either via lifting into regularity structures (that is constructing and manipulating local germs as in

Sect. 2.2) or performing a change of variable by removing some leading order paradifferential term (like in Sect. 2.4);

b) Analyze the resulting systems in terms of a finite family of basic non-linear functionals Ξ of the given data (which could be stochastic or not), their are called, according to the approach used, *enhanced noise* (in the paracontrolled approach) or *model* (in Hairer's regularity structures) or *rough path* (in Lyons' RPT theory);

c) Construct the associated solution (Itō–Lyons) map Ψ and determine the relevant topologies on the enhanced noise Ξ with respect to which Ψ has nice continuity properties.

These three steps provide the analytical backbone around which other considerations can be developed. For example, in problems related to scaling limits like those described by Equation (2), one is led to study the probabilistic convergence of lifts $\Xi_\varepsilon = J(\eta_\varepsilon)$ of smooth random fields η_ε to limiting enhanced noises Ξ in the appropriate topology. This convergence will carry on to solutions of Equation (2) via the continuous solution map Ψ.

2.5 Ambiguities. Even in the situation where the approximation η_ε converges towards a smooth object θ but only in a very weak topology (like $\mathcal{C}^\alpha$ in the setting described above, with $\alpha \in (-3/2, -1/2)$), it is *not true* that the solutions y_ε converge to the solution of the ODE

$$\dot{z} = f(z)\theta.$$

Indeed if we assume that $\Xi = \lim_\varepsilon \Xi_\varepsilon$ exists we should have $\Xi^1 = \theta$ but in general $\Theta = J(\theta) \neq \Xi$. Going back to the definition of the solution map we find out that if we let $\sigma = \Theta^{(2)} - \Xi^{(2)}$ we have $y_\varepsilon \to z = \Psi(\Xi)$ where z satisfies

$$\dot{z} = f(z)\theta + f_2(z)\sigma.$$

A correction term appears in the formulation of the limiting problem, a relic of the limiting procedure. This phenomenon has been studied in stochastic analysis McShane [1972] and Sussmann [1991], in rough path theory for ODEs P. Friz and Oberhauser [2009] and P. Friz, Gassiat, and T. Lyons [2015] but also in relation to some SPDEs Hairer and Maas [2012]. Under certain conditions one can have $\theta = 0$ and $\sigma \neq 0$. In this case the final result is a form of stochastic homogenisation and, from the point of view of the techniques we discuss here, has been considered for certain SPDEs in Hairer, Pardoux, and Piatnitski [2013].

2.6 Other approaches. Other possible frameworks for the analysis of singular SPDEs have been developed recently. Bailleul and Bernicot Bailleul and Bernicot [2016a,b] introduced a semigroup approach to paraproducts with the aim of extending the paracontrolled calculus to manifolds via invariant constructions. They also investigated higher order versions of the paracontrolled calculus as we already remarked. Kupiainen Kupiainen [2016] and Kupiainen and Marcozzi [2017] introduced a renormalization group approach where the solution is described at every scale by an *effective* equation which do not possess any singularity. The main task of the analysis is to construct these effective description satisfying recursive equations. Finally Otto and Weber Otto and Weber [2016] use a semigroup approach to decouple the singular products and identify a suitable family of stochastic objects playing the role of *enhanced noise*. In their approach the necessary Schauder estimates are derived via an extension of the Krylov–Safanov kernel–free method and this allows them to treat certain classes of quasi-linear equations. Following their pioneering work Bailleul, Debussche and Hofmanová Bailleul, Debussche, and Hofmanová [2016] used the paradifferential Ansatz to solve quasi-linear equations. A key idea of Otto and Weber approach is the introduction of a parametric family of enhanced noises which take into account the modulation of the parabolic regularisation effects given by the quasi-linear nature of the equation. Transporting this idea in the paracontrolled framework Furlan and the author Furlan and Gubinelli [2016] introduced a non–linear paraproduct and related operator which allows to cover the results of Otto and Weber in the framework of paracontrolled distributions.

3 Weak universality

One motivation for the study of SSPDE is the phenomenon of *weak universality*. This term refers to the fact that the large scale behaviour of certain classes of random PDEs or other Markovian random fields with small non–linearities or small noise depends on very few details of the exact model under consideration and that it can be described by singular SPDEs. I will illustrate this phenomenon describing recent results of Hairer and Quastel Hairer and J. Quastel [2015] on the convergence of a large class of $1+1$ interface growth models to the Kardar–Parisi–Zhang (KPZ) equation.

3.1 The Hairer–Quastel universality result. Consider a continuous growth model Halpin-Healy and Zhang [1995] given by an height function $h : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ solving the equation

$$\partial_t h = \partial_x^2 h + \sigma F(\partial_x h) + \delta \eta, \tag{14}$$

where ∂_t, ∂_x denote time and space derivatives, σ, δ are parameters, η is a smooth space–time Gaussian process and F an even polynomial. The various contributions in the r.h.s accounts for various phenomena: smoothing of the surface ($\partial_x^2 h$) (e.g. due to thermal fluctuations), lateral growth mechanism ($F(\partial_x h)$) and microscopic fluctuations in the growth rate (η). There are two interesting regimes in this equation: according to whether the non-linearity dominates or the noise dominates the behaviour at scales of order 1:

a) *Intermediate disorder regime* ($\sigma = 1$ and $\delta \ll 1$) : the noise is small. In this case we let $\varepsilon = \delta^2$ and consider the rescaled field $\tilde{h}_\varepsilon(t, x) = h(t/\varepsilon^2, x/\varepsilon)$ which satisfies

$$\partial_t \tilde{h}_\varepsilon = \partial_x^2 \tilde{h}_\varepsilon + \varepsilon^{-2} F(\varepsilon \partial_x \tilde{h}_\varepsilon) + \eta_\varepsilon, \tag{15}$$

where $\eta_\varepsilon(t, x) = \varepsilon^{-3/2}\eta(t/\varepsilon^2, x/\varepsilon)$. Formal Taylor expansion of the non-linear term gives

$$\partial_t \tilde{h}_\varepsilon = \partial_x^2 \tilde{h}_\varepsilon + \varepsilon^{-2} F(0) + \frac{1}{2} F''(0)(\partial_x \tilde{h}_\varepsilon)^2 + \mathcal{O}(\varepsilon^2(\partial_x \tilde{h}_\varepsilon)^4) + \eta_\varepsilon; \tag{16}$$

b) *Weak asymmetry* ($\delta = 1$ and $\sigma \ll 1$) : the non–linearity is small. We let $\varepsilon = \sigma^2$ and consider $\tilde{h}_\varepsilon(t, x) = \varepsilon^{1/2} h(t/\varepsilon^2, x/\varepsilon)$ which satisfies

$$\partial_t \tilde{h}_\varepsilon = \partial_x^2 \tilde{h}_\varepsilon + \varepsilon^{-1} F(\varepsilon^{1/2} \partial_x \tilde{h}_\varepsilon) + \eta_\varepsilon, \tag{17}$$

where η_ε is define as in the intermediate disorder regime and Taylor expansion gives now

$$\partial_t \tilde{h}_\varepsilon = \partial_x^2 \tilde{h}_\varepsilon + \varepsilon^{-1} F(0) + \frac{1}{2} F''(0)(\partial_x \tilde{h}_\varepsilon)^2 + \mathcal{O}(\varepsilon(\partial_x \tilde{h}_\varepsilon)^4) + \eta_\varepsilon. \tag{18}$$

The parameter ε has been chosen as a measure of the microscopic spatial scale. The random field η_ε converges (under appropriate conditions on the covariance of η) to the space–time white noise $\xi = \xi(t, x)$ with covariance

$$\mathbb{E}[\xi(t, x)\xi(s, y)] = \delta(t - s)\delta(x - y).$$

In both regimes and as $\varepsilon \to 0$, one would like to argue formally that there are constant c_ε, λ such that the random field $h_\varepsilon(t, x) = \tilde{h}_\varepsilon(t, x) - c_\varepsilon t$ converges to the solution of the Kardar–Parisi–Zhang Kardar, Parisi, and Zhang [1986] equation

$$\partial_t h = \partial_x^2 h + \frac{\lambda}{2}(\partial_x h)^2 + \xi. \tag{19}$$

Unfortunately these heuristic considerations do not stand up to further scrutiny. First the Taylor approximations turn out to be partially justified in the intermediate disorder regime but not in the weak asymmetric one, second, and more importantly the KPZ equation is strongly ill posed since the presence of the space time white noise imposes a very weak regularity on h which makes the nonlinear term not well defined.

3.2 KPZ universality. These problems has been open for very long time since the original work of Kardar, Parisi and Zhang Kardar, Parisi, and Zhang [ibid.] in the '80 where they introduced the equation to describe the universality class of one dimensional growth models. Their hypothesis was that a large class of models featuring the basic mechanisms at work in Equation (14) must show characteristic *universal* large scale properties. This conjecture is mathematically quite open, even if there have been recent important progress to prove rigorously the existence of this *KPZ universality class*. For an introduction to the mathematical literature the reader can consult the contribution of Quastel to the 2014 ICM J. D. Quastel [2014] or the lecture notes of Corwin Corwin [2012].

The universal object behind this universality class, the *KPZ fixpoint* has been described recently by Matetski, Quastel and Remenik Matetski, J. Quastel, and Remenik [2016] via exact formulas for its finite dimensional marginals. The KPZ Equation (19) itself does *not* corresponds to this fixpoint. Kardar Parisi and Zhang introduced their equation as one of many possible models whose large scale properties were universal. In this respect the KPZ fixpoint is the large scale limit of the KPZ equation. Some rigorous results are available which partially confirm this conjecture Spohn [2011], Amir, Corwin, and J. Quastel [2011], Balázs, J. Quastel, and Seppäläinen [2011], and Borodin and Corwin [2014]. The large scale limit of the KPZ equation should correspond to a vanishing viscosity and vanishing noise limit, in the precise form

$$\partial_t H_\rho = \rho \partial_x^2 H_\rho + \frac{1}{2}(\partial_x H_\rho)^2 + \rho^{1/2}\xi.$$

where $\rho \to 0$ J. Quastel [2012]. In this regime the function H_ρ should converge to the random field $\mathcal{H}$ described in Matetski, J. Quastel, and Remenik [2016].

Weak universality of KPZ stands for the fact that the KPZ Equation (19) itself can be understood as a common limit to many models under the more restrictive conditions we discussed before, namely weak asymmetry or intermediate disorder. The first mathematical result in this direction is due to Bertini and Giacomin Bertini and Giacomin [1997] in 1997. They showed that the integrated density field h_ε of a weakly asymmetric version of the exclusion process on $\mathbb{Z}$ converges upon rescaling and appropriate recentering the "solution" of the KPZ equation. As we already observed the KPZ equation is a singular SPDE which is not classically well-posed. What Bertini and Giacomin really did was to prove the convergence of the field $\phi_\varepsilon = \exp(h_\varepsilon)$ to the unique positive solution ϕ of the stochastic heat equation (SHE)

$$\partial_t \phi = \partial_x^2 \phi + \phi\xi \tag{20}$$

where the product $\phi\xi$ is understood via Itō stochastic calculus. The SHE is a standard SPDE which can be solved via standard tools (see e.g. the classic lecture notes of Walsh Walsh [1986] for the solution theory in bounded domain). This exponential transformation to a linear PDE is called Hopf–Cole transformation and the convergence results of Bertini and Giacomin justifies the fact that the *correct* notion of solution to the KPZ Equation (19) should have the property that $\phi = \exp(h)$ satisfies the SHE, this is called the Hopf–Cole solution.

3.3 A notion of solution for KPZ. After the Bertini and Giacomin convergence result there were available a candidate solution (the Hopf–Cole solution) but not an equation yet! The situation remained unclear until Hairer Hairer [2013] used the ideas and the tools of rough path theory to formulate the KPZ equation as a well–posed SPDE. This first work was instrumental to the development of the far reaching theory of regularity structures and inspired the construction of alternative theories, like that of paracontrolled distributions.

I will sketch the solution theory for the KPZ equation in terms of paracontrolled distributions as described in Gubinelli and Perkowski [2017b]. We proceed like in the analysis of the ODE by transforming the problem in order to obtain a formulation which is amenable to standard techniques. In this respect we consider the model equation

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + \frac{\lambda}{2}(\partial_x h_\varepsilon)^2 + \xi_\varepsilon,$$

where ξ_ε is a smooth approximation of the white noise and for technical reasons we consider it on the periodic domain $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$. This equation has smooth local solution (fixing some nice initial condition), however as $\varepsilon \to 0$ we loose all the useful estimates since $\xi_\varepsilon \to \xi$ only in as a space–time distribution (with parabolic regularity $-3/2$).

We split the unknown h_ε into four components as $h_\varepsilon = X_\varepsilon + Y_\varepsilon + Z_\varepsilon + H_\varepsilon$ and let $X_\varepsilon, Y_\varepsilon, Z_\varepsilon$ be solutions of

$$\mathcal{L} X_\varepsilon = \xi_\varepsilon, \;\; \mathcal{L} Y_\varepsilon = \frac{\lambda}{2}(\partial_x X_\varepsilon)^2, \;\; \mathcal{L} Z_\varepsilon = \frac{\lambda}{2}(\partial_x Y_\varepsilon)^2 + \lambda \partial_x X_\varepsilon \partial_x Y_\varepsilon + \lambda \partial_x (X_\varepsilon + Y_\varepsilon)\partial_x Z_\varepsilon$$

where $\mathcal{L} = \partial_t - \Delta$. Then the function H_ε solves

$$\mathcal{L} H_\varepsilon = \frac{\lambda}{2}(\partial_x H_\varepsilon)^2 + \frac{\lambda}{2}(\partial_x Z_\varepsilon)^2 + \lambda \partial_x (X_\varepsilon + Y_\varepsilon + Z_\varepsilon)\partial_x H_\varepsilon. \tag{21}$$

This transformation isolates the most singular contributions in the equation into the functions $X_\varepsilon, Y_\varepsilon, Z_\varepsilon$ which depends in an explicit fashion on the underlying noise ξ. The regularisation properties of the heat semigroup allows to prove that X_ε is uniformly in $C_t \mathcal{C}^{1/2-\kappa}$. In the following κ denotes some arbitrarily small positive constant and $C_t \mathcal{C}^\alpha$ denotes the space of continuous functions of time with values in $\mathcal{C}^\alpha$. It can also be shown that $(\partial_x X_\varepsilon)^2 \to +\infty$ as $\varepsilon \to 0$ (almost surely) and that there exists a constant $C_\varepsilon \to +\infty$ such that $[\![(\partial_x X_\varepsilon)^2]\!] = (\partial_x X_\varepsilon)^2 - C_\varepsilon$ converges to a well defined random field in $C_t \mathcal{C}^{-1-\kappa}$. Here the notation $[\![(\partial_x X_\varepsilon)^2]\!]$ stands for the Wick product Janson [1997].

This is an hint that our original formulation was not quite correct. In order to hope for some well defined limit, we should *renormalize* the equation and consider instead

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + \frac{\lambda}{2}[(\partial_x h_\varepsilon)^2 - C_\varepsilon] + \xi_\varepsilon,$$

and accordingly redefine Y_ε as the solution to the equation

$$\mathcal{L} Y_\varepsilon = \frac{\lambda}{2}[(\partial_x X_\varepsilon)^2 - C_\varepsilon] = \frac{\lambda}{2}[\![(\partial_x X_\varepsilon)^2]\!].$$

After these changes one can show that Y_ε converges in $C_t\mathcal{C}^{1-\kappa}$. Similar problems arise with the non-linear terms in the definition of Z_ε. A priori other renormalizations are expected whenever we try to multiply terms whose sums of regularities is not strictly positive. In the following I will assume that these renormalization have been performed by a modification in the equation for Z_ε in such a way that Z_ε has a limit in $C_t\mathcal{C}^{3/2-\kappa}$ and that the Equation (21) maintains the same form. The reader interested in the details of the precise renormalization procedure needed here can refer to the original paper of Hairer Hairer [2013] or to Hairer [2014] and Gubinelli and Perkowski [2017b]. At this point it seems that Equation (21) could be used to get uniform estimates for H_ε in $C_t\mathcal{C}^{3/2-\kappa}$, however a crucial difficulty still remains, due to the product $\partial_x X_\varepsilon \partial_x H_\varepsilon$. The sum of regularities is barely negative: $-1/2-\kappa$ for the factor $\partial_x X_\varepsilon$ and $1/2-\kappa$ for $\partial_x H_\varepsilon$. Note that

$$\mathcal{L} H_\varepsilon = \lambda \partial_x X_\varepsilon \circledcirc\!\!\!\!\!\!\;\; \partial_x H_\varepsilon + \lambda \partial_x X_\varepsilon \circledcirc \partial_x H_\varepsilon + \cdots$$

where from now on the dots $(\cdots)$ means terms of higher regularity which do not pose problems. Taking into account this expansion we introduce the paracontrolled Ansatz

$$H_\varepsilon = \lambda Q_\varepsilon \otimes \partial_x H_\varepsilon + H_\varepsilon^\sharp,$$

where $\mathcal{L} Q_\varepsilon = \partial_x X_\varepsilon$. Using the approach described in Section 2.4 we can verify that $H_\varepsilon^\sharp$ solves a parabolic equation of the form

$$\mathcal{L} H_\varepsilon^\sharp = \lambda^2 (\partial_x H_\varepsilon)(\partial_x X_\varepsilon \circledcirc \partial_x Q_\varepsilon) + \cdots$$

for which well–posedness holds provided we give an *off line* definition of $\partial_x X_\varepsilon \circledcirc \partial_x Q_\varepsilon$ as usual by now. At the end of the analysis one obtain, locally in time, a continuous solution map

$$\Psi : \Xi_\varepsilon := (X_\varepsilon, Y_\varepsilon, Z_\varepsilon, \partial_x X_\varepsilon \circledcirc \partial_x Q_\varepsilon) \mapsto (H_\varepsilon, H_\varepsilon^\sharp)$$

which allows to pass to the limit for $(h_\varepsilon)_\varepsilon$ as $\varepsilon \to 0$ and obtain a random field $h \in C_t\mathcal{C}^{1/2-\kappa}$ provided we show that

$$\Xi_\varepsilon = (X_\varepsilon, Y_\varepsilon, Z_\varepsilon, \partial_x X_\varepsilon \circledcirc \partial_x Q_\varepsilon) \to \Xi,$$

in the appropriate topology. The limit random field h satisfies an equation which formally can be written as (recall that $C_\varepsilon \to +\infty$!)

$$\partial_t h = \partial_x^2 h + \frac{\lambda}{2}[(\partial_x h)^2 - \infty] + \xi. \tag{22}$$

By itself this equation is purely formal. We have to resort to a description of h based on our analysis above to make it precise. We know that

$$h = X + Y + Z + \lambda Q \circledcirc \partial_x H + H^\sharp \tag{23}$$

where $\Psi(\Xi) = (H, H^\sharp)$ and then we have

$$[(\partial_x h)^2 - \infty] = \lim_\varepsilon [(\partial_x h_\varepsilon)^2 - C_\varepsilon] = [\![(\partial_x X)^2]\!] + 2\partial_x X \partial_x (Y+Z) + 2(\partial_x Y)^2 + 2(\partial_x Z)^2$$
$$+ 2\partial_x (X + Y + Z)\partial_x(\lambda Q \circledcirc \partial_x H + H^\sharp) + (\lambda \partial_x (Q \circledcirc \partial_x H) + \partial_x H^\sharp)^2,$$

where all the objects in the r.h.s. are well defined. In particular, from the limiting procedure we see that we can understand the product $\partial_x X \partial_x (Q \circledcirc \partial_x H)$ via the commutator lemma as

$$\begin{aligned}\partial_x X \partial_x (Q \circledcirc \partial_x H) &= \partial_x X (\partial_x Q \circledcirc \partial_x H) + \cdots \\ &= \partial_x X \circledcirc (\partial_x Q \circledcirc \partial_x H) + \partial_x H (\partial_x X \circledcirc \partial_x Q) + \cdots .\end{aligned}$$

This representation gives a well defined meaning to the r.h.s. of the Equation (22) for all the functions of the form (23) for any choice of $(H, H^\sharp) \in C_t \mathcal{C}^{1/2-\kappa} \times C_t \mathcal{C}^{3/2-\kappa}$, not necessarily satisfying the equation. However remark that this definition of $[(\partial_x h)^2 - \infty]$ depends heavily on the enhancement Ξ which, as we have already seen above, cannot be in general determined by the noise ξ but carries information about the limiting procedure.

3.4 Convergence to KPZ for the growth model. We now have a description for a candidate limit to the Equation (15) or (17). We will stick to the weakly asymmetric regime (17) since the intermediate noise regime (15) can be treated with a similar but easier approach. The result of Hairer and Quastel is the following (Theorem 1.2 in Hairer and J. Quastel [2015]).

Theorem 3.1. *Let F be an even polinomial and $(\tilde{h}_\varepsilon)_\varepsilon$ a sequence of solutions of*

$$\partial_t \tilde{h}_\varepsilon = \partial_x^2 \tilde{h}_\varepsilon + \varepsilon^{-1} F(\varepsilon^{1/2} \partial_x \tilde{h}_\varepsilon) + \xi^{(\varepsilon)}.$$

where $\xi^{(\varepsilon)}$ is a regularization of the space–time white noise via a nice smoothing kernel ρ at scale ε, namely $\xi^{(\varepsilon)} = \rho_\varepsilon * \xi$ with $\rho_\varepsilon(t,x) = \varepsilon^{-3/2}\rho(t/\varepsilon^2, x/\varepsilon)$. Let C_0 be the constant

$$C_0 = \int\int (\partial_x P * \rho)(t,x) \mathrm{d}t \mathrm{d}x \tag{24}$$

where P is the heat kernel on $\mathbb{T}$, moreover let μ_{C_0} be the Gaussian measure on $\mathbb{R}$ with variance C_0 and define the constants

$$\lambda = \int F''(x)\mu_{C_0}(\mathrm{d}x), \qquad v = \int F(x)\mu_{C_0}(\mathrm{d}x). \tag{25}$$

Then there exists a further constant c such that random field

$$h_\varepsilon(t,x) = \tilde{h}_\varepsilon(t,x) - (v/\varepsilon + c)t \tag{26}$$

converges in law to the Hopf–Cole solution of the KPZ equation.

Let us remark that the Hopf–Cole transformation which was the key tool in Bertini and Giacomin analysis of the weakly asymmetric exclusion process is not applicable here, despite the fact that the theorem can be formulated in terms of Hopf–Cole solution. Indeed one can try to perform the change of variables $\phi_\varepsilon = \exp(\tilde{h}_\varepsilon)$ but the resulting equation for ϕ_ε is as difficult as the original equation.

In the rest of this section we give some ideas on how the estimates needed to establish Theorem 3.1 can be obtained in the paracontrolled setting described above. By performing the transformation (26) we see that h_ε is a solution to

$$\partial_t h_\varepsilon = \partial_x^2 h_\varepsilon + \varepsilon^{-1}[F(\varepsilon^{1/2}\partial_x h_\varepsilon) - v] + \eta_\varepsilon.$$

The naive approach of expanding the non-linearity around 0 does not really work since soon one realizes that there are no useful estimates for $\partial_x \tilde{h}_\varepsilon$ in L^∞. Even for the linear equation

$$\partial_t \tilde{X}_\varepsilon = \partial_x^2 \tilde{X}_\varepsilon + \eta_\varepsilon.$$

the best one can have (from stochastic considerations) is $\|\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon\|_{L^\infty} \lesssim \varepsilon^{-\kappa}$ for some arbitrarily small κ. We can however mimic the paracontrolled decomposition and let $h_\varepsilon = \tilde{X}_\varepsilon + \tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon + \tilde{H}_\varepsilon$ where $\tilde{Y}_\varepsilon, \tilde{Z}_\varepsilon, \tilde{H}_\varepsilon$ will be fixed below. Now we have the possibility to expand the non-linearity around the solution $\tilde{X}_\varepsilon$ of the linear equation, giving

$$\varepsilon^{-1}F(\varepsilon^{1/2}\partial_x h_\varepsilon) = \varepsilon^{-1}[F(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon) - v] + \varepsilon^{-1/2}F'(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon)\partial_x(\tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon + \tilde{H}_\varepsilon)$$

$$+\frac{1}{2}F''(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon)[\partial_x(\tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon + \tilde{H}_\varepsilon)]^2 + \mathcal{O}(\varepsilon^{1/2}F'''(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon)[\partial_x(\tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon + \tilde{H}_\varepsilon)]^3)$$

The terms $\tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon + \tilde{H}_\varepsilon$ will behave better than $\tilde{X}_\varepsilon$ and the Taylor remainder is now negligible in the limit thanks to the factor $\varepsilon^{1/2}$. The other terms can be cast in a form very similar to that used for the KPZ equation by letting

$$\Lambda_\varepsilon = F''(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon), \quad \partial_x \hat{X}_\varepsilon = \varepsilon^{-1/2}F'(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon), \quad \mathcal{L}\tilde{Y}_\varepsilon = \varepsilon^{-1}[F(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon) - v],$$

$$\mathcal{L}\tilde{Z}_\varepsilon = \partial_x \hat{X}_\varepsilon \partial_x (\tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon) + \frac{1}{2}\Lambda_\varepsilon (\partial_x \tilde{Y}_\varepsilon)^2 + \frac{1}{2}\Lambda_\varepsilon \partial_x \tilde{Y}_\varepsilon \partial_x \tilde{Z}_\varepsilon.$$

With these definition the equation for $\tilde{H}_\varepsilon$ becomes

$$\mathcal{L}\tilde{H}_\varepsilon = \left[\partial_x \hat{X}_\varepsilon + \frac{1}{2}\Lambda_\varepsilon(\partial_x \tilde{Y}_\varepsilon + \partial_x \tilde{Z}_\varepsilon)\right]\partial_x \tilde{H}_\varepsilon + \frac{1}{2}\Lambda_\varepsilon(\partial_x \tilde{H}_\varepsilon)^2 + \frac{1}{2}\Lambda_\varepsilon(\partial_x \tilde{Z}_\varepsilon)^2$$
$$+\mathcal{O}(\varepsilon^{1/2} F'''(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon)[\partial_x(\tilde{Y}_\varepsilon + \tilde{Z}_\varepsilon + \tilde{H}_\varepsilon)]^3).$$

Comparing this equation with Equation (21) one can argue that the convergence can be proven if we are able to show that

$$(\Lambda_\varepsilon, \tilde{X}_\varepsilon, \partial_x \hat{X}_\varepsilon, \tilde{Y}_\varepsilon, \tilde{Z}_\varepsilon) \to (\lambda, X, \lambda\partial_x X, Y, Z)$$

and some other relations coming from the paracontrolled Ansatz, needed to control the product $\partial_x \hat{X}_\varepsilon \partial_x \tilde{H}_\varepsilon$. All these conditions could be in principle be established via a tour de force of intricate computations involving Gaussian random fields. See Furlan and Gubinelli [2017] for similar estimate for weak universality in reaction diffusion equations via paracontrolled analysis or Hairer and Quastel Hairer and J. Quastel [2015] for the estimation of the stochastic terms in regularity structures.

Let us highlight the role of the constants λ, v defined in (25). The constant λ is the limit of the random field Λ_ε, indeed $\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon(t,x)$ is a Gaussian random variable whose asymptotic variance do not go to zero and converges to C_0 defined in Equation (24). From this is natural to deduce that the average of $\Lambda_\varepsilon(t,x)$ converges to

$$\mathbb{E}\Lambda_\varepsilon(t,x) = \mathbb{E}F''(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon(t,x)) \to \int F''(x)\mu_{C_0}(\mathrm{d}x) = \lambda.$$

Fluctuations around this average go to zero in $C_t\mathcal{C}^{-\kappa}$. As for v, its role is to center the random field $\varepsilon^{-1}F(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon)$ so that its average is zero. Stochastic analysis then shows that

$$\varepsilon^{-1}[F(\varepsilon^{1/2}\partial_x \tilde{X}_\varepsilon) - v] \to \frac{\lambda}{2}(\partial_x X)$$

in $C_t\mathcal{C}^{-1-\kappa}$ as $\varepsilon \to 0$.

3.5 Other weak universality results. Weak universality results in the context of KPZ equation have been proven using a variety of techniques. Discrete versions of the Hopf–Cole transformations allow to tackle the limit from the point of view of the SHE and prove weak universality for certain classes of weakly asymmetric exclusion processes Bertini and Giacomin [1997], Amir, Corwin, and J. Quastel [2011], Borodin and Corwin [2014], and Corwin and Tsai [2017] and for the free energy of directed random polymers in the intermediate disorder regime Alberts, Khanin, and J. Quastel [2010].

The KPZ equation strictly speaking does not have an invariant probability measure but it has an invariant measure given by the distribution of a two-sided geometric Brownian motion with a height shift given by Lebesgue measure Funaki and J. Quastel [2015]. Based on this invariant measure and on a stationary martingale problem formulation, Gonçalves and Jara Gonçalves and Jara [2014] introduced another notion of solution which they called *energy solution of KPZ*. This allowed to prove convergence to energy solutions for a large class of particle system for which the Hopf–Cole strategy was unavailable Gonçalves, Jara, and Sethuraman [2015], Blondel, Gonçalves, and Simon [2016], and Diehl, Gubinelli, and Perkowski [2017]. In Gubinelli and Jara [2013] this notion was refined and in Gubinelli and Perkowski [2017a] it has been shown to identify a unique solution which essentially coincide with the Hopf–Cole solution.

Weak universality has been investigated also in the context of reaction diffusion equations in $d = 2, 3$ dimensions in Shen and Weber [2016], Mourrat and Weber [2017a], Hairer and Xu [2016], Shen and Xu [2017], and Furlan and Gubinelli [2017], for $d = 2$ diffusion in random environment in Chouk, Gairing, and Perkowski [2017] and for the non-linear wave equation with additive noise in $d = 2$ dimensions Gubinelli, Koch, and Oh [2017].

4 Stochastic quantisation in three dimensions

The dynamical Φ^4_3 model has been the first serious application of regularity structures Hairer [2014]. This model corresponds formally to the SPDE

$$\partial_t \varphi - \Delta\varphi + \varphi^3 - \infty\varphi = \xi \tag{27}$$

in $\mathbb{T}^3$, where ξ is space–time white noise and Δ the Laplacian on $\mathbb{T}^3$. This equation is also called stochastic quantisation equation (SQE) for a $3d$ scalar field with quadratic interaction. It can be understood as the weak universal limit of certain reaction diffusion equations (see Sect. 3.5) or as a stochastic dynamics which is reversible with respect to the Φ^4_3 Euclidean quantum field theory. This latter object can be described formally as the probability measure μ given by

$$\mu(\mathrm{d}\phi) = Z^{-1} \exp\left[-\int_{\mathbb{T}^3} (\phi(x)^4 - \infty\phi(x)^2)\mathrm{d}x\right] \mu_0(\mathrm{d}\phi)$$

where μ_0 is the Gaussian measure on $\mathcal{S}'(\mathbb{T}^3)$ with covariance $(1-\Delta)^{-1}$. This formulation is formal since μ is not absolutely continuous wrt. μ_0 and has to be understood rigorously via a limiting procedure involving a regularised exponent in the exponential (the interaction). The construction of this measure has been one of the major successes of constructive QFT Glimm [1968], Glimm and Jaffe [1973], and Feldman [1974] and is considered one serious toy model to test constructive renormalization procedures ever since Rivasseau [1991], Benfatto, Cassandro, Gallavotti, Nicoló, Olivieri, Presutti, and Scacciatelli [1980, 1978], and Gallavotti [1985]. One of the simplest construction of this measure (still quite non-trivial) is given in Brydges, Fröhlich, and Sokal [1983].

The dynamical model is inspired by the idea of *stochastic quantization* introduced by Parisi and Wu Parisi and Wu [1981]: *define* the measure μ by constructing a stochastic dynamics evolving in a fictious additional time variable.

Stochastic quantisation has various unexpected advantages. Physically it provides a way to introduce a regularisation without breaking some fundamental symmetries of the model being studied (for example the gauge symmetry Jona-Lasinio and Parrinello [1988], Bertini, Jona-Lasinio, and Parrinello [1993], and Jona-Lasinio and Parrinello [1990]). Mathematically it provides a solid ground where to attempt a controlled perturbation theory (as we will see below). Indeed in the equation the random field ξ is exactly Gaussian while under the measure μ one can identify such "free" fields only resorting to renormalization group ideas Gallavotti [1985].

Stochastic quantisation has been rigorously studied in two dimensions by Jona–Lasinio and Mitter Jona-Lasinio and Mitter [1985], by Albeverio and Rœckner Albeverio and Röckner [1991] and by Da Prato and Debussche Da Prato and Debussche [2003]. In three dimensions solutions φ are distributions living in $C_t\mathcal{C}^{-1/2-\kappa}(\mathbb{T}^3)$. The definition of the non-linear terms is highly nontrivial and, unlike the KPZ equation, cannot be attacked with RP techniques. Hairer's solution of this problem (locally in time) showed the power and fexibility of these new methods. A bit later Catellier and Chouk Catellier and Chouk [2013] described an equivalent solution theory for (27) using a paracontrolled Ansatz. See also Mourrat, Weber, and Xu [2016] for a simplified approach to the construction of the stochastic terms. Kupiainen Kupiainen and Marcozzi [2017] showed that Wilsonian renormalization group can be adapted to deal with stochastic PDEs and provided yet another solution theory.

Important results on the SQE are those of Mourrat and Weber Mourrat and Weber [2016] which were able to extend the solution theory (using paracontrolled distributions) globally in time (but still on the torus $\mathbb{T}^3$) with a tour de force of estimates. They were able to leverage the strong drift given by the cubic term to show that solutions of the SQE "comes down from infinity" (that is, they forget the initial condition) in finite time. This opens the way to a rigorous implementation of *real* stochastic quatisation by attempting to prove that whatever the initial condition, solutions of (27) converges to μ as time goes to infinity. The problem is now quite well understood in the $d = 2$ case where Mourrat and Weber proved global space–time existence (i.e. in the full plane) for the dynamics Mourrat and Weber [2017b] and where we now have a quite good understanding of the spectral gap and the exponential convergence to equilibrium of the dynamics Tsatsoulis and Weber [2016]. The $d = 3$ case is less understood, due to the more intricate solution theory.

However the work of Mourrat and Weber shows that on the microscopic scale the dynamics is essentially dominated by perturbation theory around the linear equation and that the relevant non–linear features can be taken into account by the large scales. This decomposition is analogous to the approach put in place in constructive QFT to handle the "large field problem" Benfatto, Cassandro, Gallavotti, Nicoló, Olivieri, Presutti, and Scacciatelli [1980]. Dirichlet form description of the SQE has been investigated in $d = 2, 3$ R. Zhu and X. Zhu [2017] and Röckner, R. Zhu, and X. Zhu [2017a,b].

5 Other results

Many other results have been obtained in the last few year for other type of SSPDEs. In this section I will list some of the more interesting.

– Systems of KPZ–like equations have been studied by Funaki and Hoshino Funaki and Hoshino [2016].

– The dynamic version of the Sine–Gordon model in $d = 2$ has been studied by Hairer and Shen Hairer and Shen [2016]. In this model the regularity of the stochastic terms depend on the value of a parameter and singularities have relation with the phase transition of the $2d$ Coulomb gas.

– The techniques introduced to handle SSPDE can also be used to study unbounded operators which are formally not well defined. Allez and Chouk Allez and Chouk [2015] studied the Anderson Hamiltonian in $d = 2$, i.e. the unbounded operator on $L^2(\mathbb{T}^2)$ given by $H = -\Delta + \xi$ where ξ is a white noise in $\mathbb{T}^2$. They observe that the domain of this operator can be described quite effectively via a paracontrolled Ansatz. Cannizzaro and Chouk Cannizzaro and Chouk [2015] introduced a singular martingale problem to describe the law of a diffusion with a random singular drift. The generator of the martingale problem is a formal object which has to be understood via paracontrolled calculus.

– A complex Ginzburg–Landau model has been studied by Hoshino Hoshino, Inahama, and Naganuma [2017] and Hoshino [2017] using the ideas of Mourrat and Weber to obtain global in time solutions.

– Non-linear dispersive and hyperbolic singular SPDEs have been studied by Debussche and Weber Debussche and Weber [2016] on the torus $\mathbb{T}^2$ and then extended to the full space Debussche and Martin [2017] by Debussche and Martin. A nonlinear hyperbolic wave equation in $\mathbb{T}^2$ has been studied by Oh, Koch and the author in Gubinelli, Koch, and Oh [2017].

References

Tom Alberts, Kostya Khanin, and Jeremy Quastel (Aug. 2010). "Intermediate Disorder Regime for Directed Polymers in Dimension $1 + 1$". *Physical Review Letters* 105.9, p. 090603 (cit. on p. 2348).

S. Albeverio and M. Röckner (1991). "Stochastic differential equations in infinite dimensions: solutions via Dirichlet forms". *Probability Theory and Related Fields* 89.3, pp. 347–386. MR: 1113223 (cit. on p. 2350).

S. Alinhac and P. Gérard (1991). *Opérateurs pseudo-différentiels et théorème de Nash-Moser*. EDP Sciences (cit. on p. 2337).

Romain Allez and Khalil Chouk (Nov. 2015). "The continuous Anderson hamiltonian in dimension two". arXiv: 1511.02718 (cit. on p. 2351).

Gideon Amir, Ivan Corwin, and Jeremy Quastel (Apr. 2011). "Probability distribution of the free energy of the continuum directed random polymer in $1 + 1$ dimensions". *Communications on Pure and Applied Mathematics* 64.4, pp. 466–537 (cit. on pp. 2343, 2348).

Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin (Jan. 2011). *Fourier Analysis and Nonlinear Partial Differential Equations*. Springer (cit. on pp. 2337, 2338).

I. Bailleul and F. Bernicot (2016a). "Heat semigroup and singular PDEs". *Journal of Functional Analysis* 270.9, pp. 3344–3452. MR: 3475460 (cit. on pp. 2329, 2339, 2340).

– (Sept. 2016b). "Higher order paracontrolled calculus". arXiv: 1609.06966 (cit. on pp. 2339, 2340).

Ismael Bailleul, Arnaud Debussche, and Martina Hofmanová (Oct. 2016). "Quasilinear generalized parabolic Anderson model". arXiv: 1610.06726 (cit. on p. 2341).

M. Balázs, J. Quastel, and T. Seppäläinen (2011). "Fluctuation exponent of the KPZ stochastic Burgers equation". *Journal of the American Mathematical Society* 24.3, pp. 683–708 (cit. on p. 2343).

G. Benfatto, M. Cassandro, G. Gallavotti, F. Nicoló, E. Olivieri, E. Presutti, and E. Scacciatelli (1978). "Some probabilistic techniques in field theory". *Communications in Mathematical Physics* 59.2, pp. 143–166. MR: MR0491611 (cit. on p. 2350).

– (June 1980). "Ultraviolet stability in Euclidean scalar field theories". *Communications in Mathematical Physics* 71.2, pp. 95–130 (cit. on pp. 2350, 2351).

Lorenzo Bertini and Giambattista Giacomin (1997). "Stochastic Burgers and KPZ equations from particle systems". *Communications in Mathematical Physics* 183.3, pp. 571–607. MR: 1462228 (cit. on pp. 2343, 2348).

Lorenzo Bertini, Giovanni Jona-Lasinio, and Claudio Parrinello (Jan. 1993). "Stochastic Quantization, Stochastic Calculus and Path Integrals: Selected Topics". *Progress of Theoretical Physics Supplement* 111, pp. 83–113 (cit. on p. 2350).

Oriane Blondel, Patrícia Gonçalves, and Marielle Simon (2016). "Convergence to the stochastic Burgers equation from a degenerate microscopic dynamics". *Electronic Journal of Probability* 21 (cit. on p. 2349).

Jean-Michel Bony (1981). "Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires". *Annales scientifiques de l'École normale supérieure* 14.2, pp. 209–246 (cit. on pp. 2337, 2338).

Alexei Borodin and Ivan Corwin (Feb. 2014). "Macdonald processes". *Probability Theory and Related Fields* 158.1-2, pp. 225–400 (cit. on pp. 2343, 2348).

Yvain Bruned, Ajay Chandra, Ilya Chevyrev, and Martin Hairer (Nov. 2017). "Renormalising SPDEs in regularity structures". arXiv: 1711.10239 (cit. on pp. 2329, 2336).

Yvain Bruned, Martin Hairer, and Lorenzo Zambotti (Oct. 2016). "Algebraic renormalisation of regularity structures". arXiv: 1610.08468 (cit. on pp. 2329, 2336).

David C. Brydges, Jürg Fröhlich, and Alan D. Sokal (1983). "A new proof of the existence and nontriviality of the continuum φ_2^4 and φ_3^4 quantum field theories". *Communications in Mathematical Physics* 91.2, pp. 141–186. MR: 723546 (cit. on p. 2350).

Giuseppe Cannizzaro and Khalil Chouk (Jan. 2015). "Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential". arXiv: 1501.04751 (cit. on p. 2351).

Rémi Catellier and Khalil Chouk (Oct. 2013). "Paracontrolled Distributions and the 3-dimensional Stochastic Quantization Equation". arXiv: 1310.6869 (cit. on p. 2350).

Ajay Chandra and Martin Hairer (Dec. 2016). "An analytic BPHZ theorem for regularity structures". arXiv: 1612.08138 (cit. on pp. 2329, 2336).

J.-Y. Chemin (1988a). "Calcul paradifférentiel précisé et applications á des équations aux dérivées partielles non semilinéaires". *Duke Mathematical Journal* 56.3, pp. 431–469 (cit. on p. 2339).

– (1988b). "Interaction contrôlée dans les équations aux dérivées partielles non linéaires". *Bulletin de la Société Mathématique de France* 116.3, pp. 341–383 (cit. on p. 2339).

Khalil Chouk, Jan Gairing, and Nicolas Perkowski (Dec. 2017). "An invariance principle for the two-dimensional parabolic Anderson model with small potential". *Stochastics and Partial Differential Equations: Analysis and Computations* 5.4, pp. 520–558 (cit. on p. 2349).

Ivan Corwin (2012). "The Kardar-Parisi-Zhang equation and universality class". *Random Matrices Theory Appl.* 1.1, pp. 1130001, 76. MR: 2930377 (cit. on p. 2343).

Ivan Corwin and Li-Cheng Tsai (May 2017). "KPZ equation limit of higher-spin exclusion processes". *The Annals of Probability* 45.3, pp. 1771–1798 (cit. on p. 2348).

Giuseppe Da Prato and Arnaud Debussche (2003). "Strong solutions to the stochastic quantization equations". *The Annals of Probability* 31.4, pp. 1900–1916. MR: 2016604 (cit. on p. 2350).

A. M. Davie (2007). "Differential equations driven by rough paths: an approach via discrete approximation". *Applied Mathematics Research Express. AMRX* 2, Art. ID abm009, 40 (cit. on p. 2336).

Arnaud Debussche and Jörg Martin (July 2017). "Solution to the stochastic Schrödinger equation on the full space". arXiv: 1707.06431 (cit. on p. 2351).

Arnaud Debussche and Hendrik Weber (Dec. 2016). "The Schrödinger equation with spatial white noise potential". arXiv: 1612.02230 (cit. on p. 2351).

Joscha Diehl, Massimiliano Gubinelli, and Nicolas Perkowski (Sept. 2017). "The Kardar–Parisi–Zhang Equation as Scaling Limit of Weakly Asymmetric Interacting Brownian Motions". *Communications in Mathematical Physics* 354.2, pp. 549–589 (cit. on p. 2349).

Weinan E (Aug. 2011). *Principles of Multiscale Modeling*. 1 edition. Cambridge, UK ; New York: Cambridge University Press (cit. on p. 2330).

Lawrence C. Evans (Mar. 1990). *Weak Convergence Methods for Nonlinear Partial Differential Equations*. Providence, R.I: American Mathematical Society (cit. on p. 2331).

Joel Feldman (June 1974). "The $\lambda\phi_3^4$ field theory in a finite volume". *Communications in Mathematical Physics* 37.2, pp. 93–120 (cit. on p. 2349).

Denis Feyel and Arnaud De La Pradelle (Oct. 2006). "Curvilinear Integrals Along Enriched Paths". *Electronic Journal of Probability* 11 (cit. on p. 2336).

Denis Feyel, Arnaud de La Pradelle, and Gabriel Mokobodzki (2008). "A non-commutative sewing lemma". *Electronic Communications in Probability* 13, pp. 24–34. MR: 2372834 (cit. on p. 2336).

Peter K. Friz and Martin Hairer (Aug. 2014). *A Course on Rough Paths: With an Introduction to Regularity Structures*. Springer (cit. on pp. 2334–2336).

Peter Friz, Paul Gassiat, and Terry Lyons (2015). "Physical Brownian motion in a magnetic field as a rough path". *Transactions of the American Mathematical Society* 367.11, pp. 7939–7955. MR: 3391905 (cit. on p. 2340).

Peter Friz and Harald Oberhauser (2009). "Rough path limits of the Wong-Zakai type with a modified drift term". *Journal of Functional Analysis* 256.10, pp. 3236–3256. MR: 2504524 (cit. on p. 2340).

Tadahisa Funaki and Masato Hoshino (Nov. 2016). "A coupled KPZ equation, its two types of approximations and existence of global solutions". arXiv: 1611.00498 (cit. on p. 2351).

Tadahisa Funaki and Jeremy Quastel (June 2015). "KPZ equation, its renormalization and invariant measures". *Stochastic Partial Differential Equations: Analysis and Computations* 3.2, pp. 159–220 (cit. on p. 2349).

M. Furlan and M. Gubinelli (Oct. 2016). "Paracontrolled quasilinear SPDEs". arXiv: 1610.07886 (cit. on p. 2341).

Marco Furlan and Massimiliano Gubinelli (Aug. 2017). "Weak universality for a class of 3d stochastic reaction-diffusion models". arXiv: 1708.03118 (cit. on pp. 2348, 2349).

Giovanni Gallavotti (Apr. 1985). "Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods". *Reviews of Modern Physics* 57.2, pp. 471–562 (cit. on p. 2350).

James Glimm (1968). "Boson fields with the ϕ^4 interaction in three dimensions". *Communications in Mathematical Physics* 10, pp. 1–47. MR: 0231601 (cit. on p. 2349).

James Glimm and Arthur Jaffe (1973). "Positivity of the ϕ_3^4 Hamiltonian". *Fortschritte der Physik. Progress of Physics* 21, pp. 327–376. MR: 0408581 (cit. on p. 2349).

Patrícia Gonçalves and Milton Jara (May 2014). "Nonlinear Fluctuations of Weakly Asymmetric Interacting Particle Systems". *Archive for Rational Mechanics and Analysis* 212.2, pp. 597–644 (cit. on p. 2349).

Patrícia Gonçalves, Milton Jara, and Sunder Sethuraman (2015). "A stochastic Burgers equation from a class of microscopic interactions". *The Annals of Probability* 43.1, pp. 286–338. MR: 3298474 (cit. on p. 2349).

M. Gubinelli (2004). "Controlling rough paths". *Journal of Functional Analysis* 216.1, pp. 86–140. MR: 2091358 (cit. on p. 2336).

M. Gubinelli and M. Jara (2013). "Regularization by noise and stochastic Burgers equations". *Stochastic Partial Differential Equations. Analysis and Computations* 1.2, pp. 325–350. MR: 3327509 (cit. on p. 2349).

Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski (2015). "Paracontrolled distributions and singular PDEs". *Forum of Mathematics. Pi* 3, e6, 75. MR: 3406823 (cit. on pp. 2329, 2336, 2338, 2339).

Massimiliano Gubinelli, Herbert Koch, and Tadahiro Oh (Mar. 2017). "Renormalization of the two-dimensional stochastic nonlinear wave equation". arXiv: 1703.05461 (cit. on pp. 2349, 2351).

Massimiliano Gubinelli and Nicolas Perkowski (2015). *Lectures on singular stochastic PDEs*. Vol. 29. Ensaios Matemáticos [Mathematical Surveys]. Sociedade Brasileira de Matemática, Rio de Janeiro. MR: 3445609 (cit. on pp. 2337, 2339).

– (2017a). "Energy solutions of KPZ are unique". *Journal of the American Mathematical Society*, p. 1 (cit. on p. 2349).

– (Jan. 2017b). "KPZ Reloaded". *Communications in Mathematical Physics* 349.1, pp. 165–269 (cit. on pp. 2344, 2345).

M. Hairer (Mar. 2014). "A theory of regularity structures". *Inventiones mathematicae* 198.2, pp. 269–504 (cit. on pp. 2329, 2332, 2336, 2345, 2349).

Martin Hairer (June 2013). "Solving the KPZ equation". *Annals of Mathematics* 178.2, pp. 559–664 (cit. on pp. 2344, 2345).

Martin Hairer and Jan Maas (July 2012). "A spatial version of the Itô–Stratonovich correction". *The Annals of Probability* 40.4. Mathematical Reviews number (MathSciNet) MR2978135, pp. 1675–1714 (cit. on p. 2340).

Martin Hairer, Etienne Pardoux, and Andrey Piatnitski (Dec. 2013). "Random homogenisation of a highly oscillatory singular potential". *Stochastic Partial Differential Equations: Analysis and Computations* 1.4, pp. 571–605 (cit. on p. 2340).

Martin Hairer and Jeremy Quastel (Dec. 2015). "A class of growth models rescaling to KPZ". arXiv: 1512.07845 (cit. on pp. 2341, 2346, 2348).

Martin Hairer and Hao Shen (2016). "The dynamical sine-Gordon model". *Communications in Mathematical Physics* 341.3, pp. 933–989. MR: 3452276 (cit. on p. 2351).

Martin Hairer and Weijun Xu (Jan. 2016). "Large scale behaviour of 3D continuous phase coexistence models". arXiv: 1601.05138 (cit. on p. 2349).

Timothy Halpin-Healy and Yi-Cheng Zhang (Mar. 1995). "Kinetic roughening phenomena, stochastic growth, directed polymers and all that. Aspects of multidisciplinary statistical mechanics". *Physics Reports* 254.4, pp. 215–414 (cit. on p. 2341).

Masato Hoshino (Apr. 2017). "Global well-posedness of complex Ginzburg-Landau equation with a space-time white noise". arXiv: 1704.04396 (cit. on p. 2351).

Masato Hoshino, Yuzuru Inahama, and Nobuaki Naganuma (Feb. 2017). "Stochastic complex Ginzburg-Landau equation with space-time white noise". arXiv: 1702.07062 (cit. on p. 2351).

Svante Janson (June 1997). *Gaussian Hilbert Spaces*. Cambridge University Press (cit. on p. 2344).

G. Jona-Lasinio and P. K. Mitter (1985). "On the stochastic quantization of field theory". *Communications in Mathematical Physics (1965-1997)* 101.3, pp. 409–436 (cit. on p. 2350).

G. Jona-Lasinio and C. Parrinello (Nov. 1988). "On the stochastic quantization of gauge theories". *Physics Letters B* 213.4, pp. 466–470 (cit. on p. 2350).

– (1990). "Large deviations theory and stochastic gauge fixing". In: *Stochastic processes, physics and geometry (Ascona and Locarno, 1988)*. World Sci. Publ., Teaneck, NJ, pp. 478–484. MR: 1124225 (cit. on p. 2350).

Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang (Mar. 1986). "Dynamic Scaling of Growing Interfaces". *Physical Review Letters* 56.9, pp. 889–892 (cit. on pp. 2342, 2343).

Antti Kupiainen (Mar. 2016). "Renormalization Group and Stochastic PDEs". *Annales Henri Poincaré* 17.3, pp. 497–535 (cit. on pp. 2329, 2341).

Antti Kupiainen and Matteo Marcozzi (Feb. 2017). "Renormalization of Generalized KPZ Equation". *Journal of Statistical Physics* 166.3-4, pp. 876–902 (cit. on pp. 2341, 2350).

Terry Lyons (1998). "Differential equations driven by rough signals". *Revista Matemática Iberoamericana*, pp. 215–310 (cit. on p. 2336).

Terry J. Lyons, Michael J. Caruana, and Thierry Lévy (June 2007). *Differential Equations Driven by Rough Paths: Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004*. 1st ed. Springer (cit. on p. 2336).

Terry Lyons and Zhongmin Qian (2002). *System Control and Rough Paths*. Oxford University Press (cit. on p. 2336).

Konstantin Matetski, Jeremy Quastel, and Daniel Remenik (Dec. 2016). "The KPZ fixed point". arXiv: 1701.00018 (cit. on p. 2343).

E. J. McShane (1972). "Stochastic differential equations and models of random processes". *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pp. 263–294. MR: 0402921 (cit. on p. 2340).

G. Metivier (2008). *Para-differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems*. Pisa: Edizioni della Normale (cit. on p. 2337).

Yves Meyer (1981). "Remarques sur un théorème de J.-M. Bony". In: *Rendiconti del Circolo Matematico di Palermo. Serie II*, pp. 1–20 (cit. on p. 2337).

Jean-Christophe Mourrat and Hendrik Weber (Jan. 2016). "The dynamic Φ^4_3 model comes down from infinity". arXiv: 1601.01234 (cit. on p. 2350).

– (2017a). "Convergence of the two-dimensional dynamic Ising-Kac model to Φ^4_2". *Communications on Pure and Applied Mathematics* 70.4, pp. 717–812. MR: 3628883 (cit. on p. 2349).

– (2017b). "Global well-posedness of the dynamic Φ^4 model in the plane". *The Annals of Probability* 45.4, pp. 2398–2476. MR: 3693966 (cit. on p. 2350).

Jean-Christophe Mourrat, Hendrik Weber, and Weijun Xu (Oct. 2016). "Construction of Φ^4_3 diagrams for pedestrians". arXiv: 1610.08897 (cit. on p. 2350).

Felix Otto and Hendrik Weber (May 2016). "Quasilinear SPDEs via rough paths". arXiv: 1605.09744 (cit. on pp. 2329, 2333, 2341).

G. Parisi and Yong Shi Wu (1981). "Perturbation theory without gauge fixing". *Scientia Sinica. Zhongguo Kexue* 24.4, pp. 483–496. MR: 626795 (cit. on p. 2350).

J. D. Quastel (2014). "The Kardar-Parisi-Zhang equation and universality class". In: *XVIIth International Congress on Mathematical Physics*. World Sci. Publ., Hackensack, NJ, pp. 113–133. MR: 3204148 (cit. on p. 2343).

Jeremy Quastel (Mar. 2012). *Introduction to KPZ*. Lecture notes. Arizona School of Analysis and Mathematical Physics (cit. on p. 2343).

Daniel Revuz and Marc Yor (Dec. 2004). *Continuous Martingales and Brownian Motion*. 3rd. Springer (cit. on p. 2331).

V. Rivasseau (May 1991). *From Perturbative to Constructive Renormalization*. 2 edition. Princeton, N.J: Princeton University Press (cit. on p. 2350).

Michael Röckner, Rongchan Zhu, and Xiangchan Zhu (2017a). "Ergodicity for the stochastic quantization problems on the 2D-torus". *Communications in Mathematical Physics* 352.3, pp. 1061–1090. MR: 3631399 (cit. on p. 2351).

– (2017b). "Restricted Markov uniqueness for the stochastic quantization of $P(\Phi)_2$ and its applications". *Journal of Functional Analysis* 272.10, pp. 4263–4303. MR: 3626040 (cit. on p. 2351).

Hao Shen and Hendrik Weber (Aug. 2016). "Glauber dynamics of 2D Kac-Blume-Capel model and their stochastic PDE limits". arXiv: 1608.06556 (cit. on p. 2349).

Hao Shen and Weijun Xu (Oct. 2017). "Weak universality of dynamical Φ^4_3 : non-Gaussian noise". *Stochastics and Partial Differential Equations: Analysis and Computations* (cit. on p. 2349).

Herbert Spohn (Dec. 2011). *Large Scale Dynamics of Interacting Particles*. Softcover reprint of the original 1st ed. 1991 edition. Berlin, Heidelberg: Springer (cit. on p. 2343).

Héctor J. Sussmann (1991). "Limits of the Wong-Zakai type with a modified drift term". In: *Stochastic analysis*. Academic Press, Boston, MA, pp. 475–493. MR: 1119845 (cit. on p. 2340).

Terence Tao (2006). *Nonlinear Dispersive Equations: Local And Global Analysis*. American Mathematical Soc. (cit. on p. 2337).

Michael E. Taylor (July 2000). *Tools for Pde: Pseudodifferential Operators, Paradifferential Operators, and Layer Potentials*. Amer Mathematical Society (cit. on p. 2337).

Pavlos Tsatsoulis and Hendrik Weber (Sept. 2016). "Spectral Gap for the Stochastic Quantization Equation on the 2-dimensional Torus". arXiv: 1609.08447 (cit. on p. 2350).

John B. Walsh (1986). "An introduction to stochastic partial differential equations". In: *école d'été de probabilités de Saint-Flour, XIV—1984*. Vol. 1180. Lecture Notes in Math. Springer, Berlin, pp. 265–439. MR: 876085 (cit. on p. 2343).

S. Watanabe and N. Ikeda (Feb. 1981). *Stochastic Differential Equations and Diffusion Processes, Volume 24*. Amsterdam ; New York : Tokyo : New York, NY: North Holland (cit. on p. 2331).

L. C. Young (1936). "An inequality of the Hölder type, connected with Stieltjes integration". *Acta Mathematica* 67, pp. 251–282 (cit. on p. 2334).

Rongchan Zhu and Xiangchan Zhu (Mar. 2017). "Dirichlet form associated with the Φ^4_3 model". arXiv: 1703.09987 (cit. on p. 2351).

Received 2017-11-30.

Massimiliano Gubinelli
IAM & HCM
University of Bonn
Germany
gubinelli@iam.uni-bonn.de

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2357–2380)

ANALYTIC TOOLS FOR THE STUDY OF FLOWS AND INVERSE PROBLEMS

COLIN GUILLARMOU

Abstract

In this survey, we review recent results in hyperbolic dynamical systems and in geometric inverse problems using analytic tools, based on spectral theory and microlocal methods.

1 Introduction

We describe recent results in dynamical systems and inverse problems using analytic tools based on microlocal analysis. These tools are designed to understand the long time dynamics of hyperbolic dynamical systems through spectral theory, and to solve transport equations in certain functional spaces, even when the flow is not dissipative. They allow for example to prove meromorphic extension of dynamical zeta functions in the smooth setting (while it was only known in the real analytic setting before).

These tools can also be applied to geometric inverse problems such as geodesic X-ray tomography and the boundary rigidity or lens rigidity problem, where one wants to determine a Riemannian metric from the length of its closed geodesics in the closed case, or the Riemannian distance between boundary point in the case with boundary.

In Section 2, we review some recent results concerning the study of hyperbolic flows and Ruelle resonances, while in Section 3 we discuss the boundary/lens rigidity problem and the analysis of X-ray tomography in the curved setting.

2 Microlocal analysis for Anosov and Axiom A flows

2.1 Anosov and Axiom A flows. Consider X a non-vanishing smooth vector field on a compact smooth manifold $\mathfrak{M}$ (with or without boundary), generating a flow $\varphi_t : \mathfrak{M} \to \mathfrak{M}$.

MSC2010: 35-06.

We define the maps

$$\tau^{\pm} : \mathcal{M} \to \mathbb{R}^{\pm} \cup \{\pm\infty\}$$

by the condition that $(\tau^-(y), \tau^+(y))$ is the maximal interval of time where the flow $\varphi_t(y)$ is defined in $\mathcal{M}$ (we put $\tau^{\pm}(y) = 0$ if $\varphi_{\pm t}(y)$ is not defined for $t > 0$). We will call *trapped set* $\mathcal{K}$ the closed set of points where this interval is $\mathbb{R}$

$$\mathcal{K} := \{y \in \mathcal{M}; \tau^+(y) = +\infty, \tau^-(y) = -\infty\}$$

and we shall call *incoming tail* Γ_- and *outgoing tail* Γ_+ the sets

$$\Gamma_{\pm} := \{y \in \mathcal{M}; \tau^{\mp}(y) = \mp\infty\}.$$

Note that when $\partial\mathcal{M} = \varnothing$, we have $\Gamma_{\pm} = \mathcal{K} = \mathcal{M}$. We say that $\mathcal{K}$ is a hyperbolic set for the flow if there is a continuous flow-invariant splitting of $T\mathcal{M}$ over $\mathcal{K}$

$$T_{\mathcal{K}}\mathcal{M} = \mathbb{R}X \oplus E_s \oplus E_u$$

such that there are uniform constants $C > 0, \nu > 0$ satisfying

$$\begin{aligned} &\forall y \in \mathcal{K}, \forall \xi \in E_s(y), \forall t \geq 0, \quad ||d\varphi_t(y)\xi|| \leq Ce^{-\nu t}||\xi|| \\ &\forall y \in \mathcal{K}, \forall \xi \in E_u(y), \forall t \leq 0, \quad ||d\varphi_t(y)\xi|| \leq Ce^{-\nu |t|}||\xi||. \end{aligned} \tag{2-1}$$

Here the norm is with respect to any fixed Riemannian metric on $\mathcal{M}$. When $\mathcal{K} = \mathcal{M}$ and $\mathcal{M}$ is a closed manifold, we say that the flow of X is *Anosov*. When $\mathcal{K}$ is a compact set in the interior $\mathcal{M}^{\circ}$ of $\mathcal{M}$, we shall say that the flow is *Axiom A*, following the terminology of Smale [1967]. By the spectral decomposition of hyperbolic flows Katok and Hasselblatt [1995, Theorem 18.3.1 and Exercise 18.3.7], the non-wandering set $\Omega \subset \mathcal{K}$ of φ_t decomposes into finitely many disjoint invariant topologically transitive sets $\Omega = \cup_{i=1}^N \Omega_i$ for φ_t. By Katok and Hasselblatt [ibid., Corollary 6.4.20], the periodic orbits of the flow are dense in Ω. Each Ω_i is called a basic set in the terminology of hyperbolic dynamical systems. General Axiom A flows are defined by Smale [1967] and essentially consist in a finite union of basic sets and fixed points for the flow (in that case X would have to vanish at some points). For example, gradient flows of Morse-Smale type have finitely many fixed hyperbolic points and are Axiom A as well.

2.2 Solving transport equations and continuation of the resolvent. The classical important objects for a flow as above are the periodic orbits and their length, the topological entropy of the flow, the invariant measures, the ergodicity and mixing properties, and solving cohomological equations. In some sense, all these quantities or properties are related to the transport equation

$$(X - V)u = f \tag{2-2}$$

where $V \in C^0$ is a potential and u, f are functions or distributions.

For example, a periodic orbit γ gives rise to a Dirac distribution δ_γ given by $\langle \delta_\gamma, f\rangle = \int_\gamma f$ solving the equation

$$X\delta_\gamma = 0.$$

We say that δ_γ are invariant distributions for the flow (in fact they are invariant measures). The cohomological equation problem asks if $f \in C^\infty(\mathfrak{M})$ and $\langle w, f\rangle = 0$ for all $w \in \mathfrak{D}'(\mathfrak{M}) \cap \ker X$, then $f = Xu$ for some $u \in C^\infty$. The ergodicity and mixing of the flow with respect to a smooth measure can be read from the L^2 spectrum of X, and the entropy appears also as a leading eigenvalue of some operator $X - V$ for a well chosen potential V.

For an Anosov or Axiom A flow, we can then ask when the equation (2-2) can be solved, and in what spaces. A convenient way to analyse this is to view $P := -X + V$ as a first order differential operator and to define the resolvent

$$R_P(\lambda) := (P - \lambda)^{-1} : L^2(\mathfrak{M}) \to L^2(\mathfrak{M})$$

for $\mathrm{Re}(\lambda) \gg 1$. An explicit expression is given by the converging expression

$$R_P(\lambda) f(y) = -\int_{\tau^-(y)}^{0} e^{\lambda t + \int_t^0 V(\varphi_s(y))ds} f(\varphi_t(y))dt$$

if $\mathrm{Re}(\lambda) \gg 1$ is large enough. However this operator can not be extended in $\lambda \in \mathbb{C}$ on $L^2(\mathfrak{M})$ when we reach its L^2-spectrum. This is for example a problem in the study of the cohomological equation (say when $V = 0$ and μ is a smooth invariant measure for X) since the equation $Xu = f$ corresponds to the spectral value $\lambda = 0$ and X has essential spectrum on $i\mathbb{R}$.

In the case of an Anosov flow, a major step was first made by Butterley and Liverani [2007]. They proved that the resolvent of P admits a meromorphic extension to $\mathbb{C}$ on certain functional spaces and that P has only discrete spectrum on those spaces. Another proof of microlocal nature appeared later, first in the work of Faure and Sjöstrand [2011] and then of Dyatlov and Zworski [n.d.]. Before we summarise these results, let us introduce the dual Anosov decomposition

$$T^*\mathfrak{M} = E_0^* \oplus E_s^* \oplus E_u^*, \text{ with}$$
$$E_0^*(E_u \oplus E_s) = 0, \quad E_u^*(E_u \oplus \mathbb{R}X) = 0, \quad E_s^*(E_s \oplus \mathbb{R}X) = 0.$$

and mention that we denote by $H^s(\mathfrak{M})$ the usual L^2-based Sobolev space when $s \in \mathbb{R}$.

Theorem 1 (Butterley and Liverani [2007], Faure and Sjöstrand [2011], and Dyatlov and Zworski [n.d.]). *Let X be a smooth vector field generating an Anosov flow on a compact manifold $\mathfrak{M}$, let $V \in C^\infty(\mathfrak{M})$ and let $P = -X + V$ be the associated first-order differential operator.*
1) There exists $C_0 \geq 0$ such that the resolvent $R_P(\lambda) := (P - \lambda)^{-1} : L^2(\mathfrak{M}) \to L^2(\mathfrak{M})$ of P is defined for $\mathrm{Re}(\lambda) > C_0$ and extends meromorphically to $\lambda \in \mathbb{C}$ as a family of bounded operators $R_P(\lambda) : C^\infty(\mathfrak{M}) \to \mathfrak{D}'(M)$. The poles are called Ruelle resonances, the operator $\Pi_{\lambda_0} := -\mathrm{Res}_{\lambda_0} R_P(\lambda)$ at a pole λ_0 is a finite rank projector and there exists $p \geq 1$ such that $(P - \lambda_0)^p \Pi_{\lambda_0} = 0$. The distributions in $\mathrm{Ran}\,\Pi_{\lambda_0}$ are called generalized resonant states and those in $\mathrm{Ran}\,\Pi_{\lambda_0} \cap \ker(P - \lambda_0)$ are called resonant states.
2) There is $C_1 > 0$ depending only on the constant ν in (2-1) such that for each $N \in [0, \infty)$, there exists a Hilbert space $\mathcal{H}^N$ so that $C^\infty(\mathfrak{M}) \subset \mathcal{H}^N \subset H^{-N}(\mathfrak{M})$ and such that $R_P(\lambda) : \mathcal{H}^N \to \mathcal{H}^N$ is a meromorphic family of bounded operators in $\mathrm{Re}(\lambda) > C_0 - C_1 N$, and $(P - \lambda) : \mathrm{Dom}(P) \cap \mathcal{H}^N \to \mathcal{H}^N$ is an analytic family of Fredholm operators in that region with inverse given by $R_P(\lambda)$.
3) For a resonance λ_0, the wave-front set of each generalized resonant state $u \in \mathrm{Ran}(\Pi_{\lambda_0})$ is contained in E_u^*.

In Butterley and Liverani [2007] the space $\mathcal{H}^N$ is actually a Banach space, but we will focus here rather on the works Faure and Sjöstrand [2011] and Dyatlov and Zworski [n.d.] where $\mathcal{H}^N$ is indeed a Hilbert space defined by $\mathcal{H}^N = A_N(L^2(\mathfrak{M}))$ where A_N is a certain pseudo-differential operator in an exotic class. The operator A_N is constructed as $A_N = \mathrm{Op}(a_N)$ where Op denotes a standard quantization procedure (see e.g. Zworski [2012]) and $a_N \in C^\infty(T^*\mathfrak{M})$ is a symbol of the form $a_N(y, \xi) = \exp(m(y, \xi) \log |\xi|)$ for $|\xi| \geq 1$, and $m(y, \xi)$ is a homogeneous function of degree 0 in the fibers of $T^*\mathfrak{M}$, equal to -1 near E_s^* and $+1$ near E_u^*. Roughly speaking, a function in $\mathcal{H}^N$ is in the classical Sobolev space $H^N(\mathfrak{M})$ (microlocally) near E_s^* and in the classical Sobolev space $H^{-N}(\mathfrak{M})$ near E_u^*. The behaviour outside the characteristic set $\{(y, \xi) \in T^*\mathfrak{M}; \xi(X) = 0\} = E_u^* \oplus E_s^*$ of X has less importance. These $\mathcal{H}^N$ spaces are called *anisotropic Sobolev spaces*. Theorem 1 tells us that we can solve the transport equation $(-X + V - \lambda)u = f$ in a well-posed fashion provided f, u are in a good anisotropic Sobolev space, except for a discrete set of λ where f needs to be in the finite codimension range. These types of spaces were first introduced (or some Hölder version) for the case of hyperbolic diffeomorphisms, in the work of Blank, Keller, and Liverani [2002], Liverani [2005], Gouëzel and Liverani [2006], Baladi and Tsujii [2007] and with a microlocal approach in Faure, Roy, and Sjöstrand [2008]. We mention that there were previous important works on spectral approaches of hyperbolic dynamical systems by Ruelle, Fried, Pollicott, Rugh, Kitaev, etc, mostly in the case of real analytic diffeomorphisms and flows, but we won't focus on these aspects.

In the Axiom A case, we proved in Dyatlov and Guillarmou [2016] a result in the same spirit as Theorem 1. Our setting is a manifold $\mathfrak{M}$ with boundary and a non-vanishing vector field X with hyperbolic trapped set $\mathcal{K} \subset \mathfrak{M}^\circ$, with a convexity condition on $\partial\mathfrak{M}$ (the boundary is strictly convex with respect to the flow lines of X). We note that this convexity is not really necessary and can be removed by the argument of Guillarmou, Mazzucchelli, and Tzou [n.d., Section 2.2], for there is always a convex neighborhood of $\mathcal{K}$, even if $\partial\mathfrak{M}$ is not convex. For such flows, the stable space E_s over K extends continuously over Γ_- in a subbundle E_- satisfying hyperbolic estimates similar to E_s (i.e. those in (2-1)), while E_u extends to Γ_+ in a subbundle E_+ satisfying hyperbolicity estimates similar to E_u. In fact E_- is simply the union of tangent spaces to the stable manifolds of $\mathcal{K}$ while E_+ is those for the unstable manifolds of $\mathcal{K}$. We will also use the dual spaces $E_\pm^*$ over $\Gamma_\pm$ defined by

$$E_\pm^*(E_\pm \oplus \mathbb{R}X) = 0.$$

We will write below $\mathrm{WF}(u) \subset T^*\mathfrak{M}$ for the wave-front set of a distribution.

Theorem 2. *Dyatlov and Guillarmou [2016] Let $\mathfrak{M}$ be a manifold with boundary and X a smooth non vanishing vector field so that its trapped set $\mathcal{K}$ is a compact hyperbolic set in $\mathfrak{M}^\circ$. Then for each $V \in C^\infty(\mathfrak{M})$ the resolvent $R_P(\lambda) = (P-\lambda)^{-1}$ of $P := -X + V$ admits a meromorphic extension from $\mathrm{Re}(\lambda) \gg 1$ to $\lambda \in \mathbb{C}$ with poles of finite multiplicity as a map $C_c^\infty(\mathfrak{M}^\circ) \to \mathfrak{D}'(\mathfrak{M}^\circ)$. The poles are called resonances and the generalized eigenstates $u \in \mathrm{Ran}(\mathrm{Res}_{\lambda_0}(R_P(\lambda))$ satisfy the following properties*

$$\mathrm{supp}(u) \subset \Gamma_+, \quad \mathrm{WF}(u) \subset E_+^*.$$

Moreover, for $f \in C_c^\infty(\mathfrak{M}^\circ)$, we have $R_P(\lambda)f \in C^\infty(\mathfrak{M} \setminus \Gamma_+) \cap H^{-N}(\mathfrak{M})$ for some $N > 0$ depending only on $\mathrm{Re}(\lambda)$, and $\mathrm{WF}(R_P(\lambda)f) \subset E_+^*$.

Here again, the proof uses the construction of anisotropic Sobolev spaces, but new complications come from the fact that the hyperbolicity is only on a compact subset of $\mathfrak{M}$. In the proof, we extend the flow to a compact manifold with boundary and add some absorbing and elliptic operators outside $\mathfrak{M}$. We notice that geodesic flows on closed negatively curved manifolds are examples of Anosov flows. Similarly, examples of Axiom A flows are given by geodesic flows on negatively curved non-compact manifold (M, g) satisfying the following conditions: there exists a strictly convex region M_0 such that the map $\psi := \mathbb{R}^+ \times \partial M_0 \to M \setminus M_0$ given by $\psi(t,x) = \exp_x(t\nu_x)$ is a diffeomorphism if ν_x is the unit normal to ∂M_0 pointing outside M_0. Convex co-compact hyperbolic manifolds are such examples, but we can also consider asymptotically hyperbolic manifolds with hyperbolic trapped set that are not necessarily negatively curved.

For Morse-Smale gradient flows, Dang and Rivière [n.d.(b)] studied Ruelle resonances using also a Faure-Sjöstrand approach, this is another (simpler) case of Axiom A flows. In that case the Ruelle spectrum can be explicitly computed using a normal form for the flow near a hyperbolic fixed point. We finally mention a forthcoming work of Bonthonneau and Weich [n.d.] for cases where the flow is hyperbolic but the trapped set is not compact: they show meromorphic extension of the resolvent of the geodesic flow in the case of finite volume manifolds with hyperbolic cusps. They can therefore define Ruelle resonances also in that setting.

2.3 Localisation of the spectrum and decay of correlations. We now assume that $\mathfrak{M}$ is closed, X generates an Anosov flow and that there is a smooth invariant measure μ for the flow φ_t, that is $\mathcal{L}_X\mu = 0$. In that case we formally have $X^* = -X$ and, for $P = -X$, the resolvent $R_P(\lambda)$ is analytic in $\mathrm{Re}(\lambda) > 0$. The constant functions belong to $\ker X \cap C^\infty(\mathfrak{M})$, thus 0 is a resonance. It is easy to prove that there is no Jordan block at $\lambda = 0$. Ergodicity of μ with respect to φ_t is equivalent to the fact that the only resonant states with resonance $\lambda = 0$ are the constants. Mixing of φ_t is equivalent to the fact that 0 is the only resonance on the imaginary line $\mathrm{Re}(\lambda) = 0$ (corresponding to the L^2 spectrum of X). The *correlation functions* are defined for $f_1, f_2 \in C^\infty(\mathfrak{M})$ by

$$C(f_1, f_2, t) := \langle \varphi_t^* f_1, f_2\rangle_{L^2(\mathfrak{M},\mu)}.$$

Understanding the speed of mixing, when there is mixing, amounts to studying the behaviour of $C(f_1, f_2, t)$ as $t \to \pm\infty$ for each observables f_1, f_2. It is easy to check that the resolvent is related to the correlation functions via a Laplace transform:

$$\langle R_P(\lambda) f_1, f_2\rangle = -\int_{-\infty}^{0} e^{\lambda t} C(f_1, f_2, t)dt. \tag{2-3}$$

When the correlations have an asymptotic expansion of the form

$$C(f_1, f_2, t) = \langle f_1, f_2\rangle + \sum_{j=1}^{N}\sum_{k=0}^{k_j} e^{-\lambda_j t} t^k \alpha_{j,k}(f_1, f_2) + \mathcal{O}(e^{-\nu|t|}) \tag{2-4}$$

as $t \to -\infty$, for some $\lambda_j \in \mathbb{C}$ with $\mathrm{Re}(\lambda_j) \in (-\nu, 0)$ with $\nu > 0$ and $k_j \in \mathbb{N}$, one easily get from (2-3) that $\langle R_P(\lambda) f_1, f_2\rangle$ has only finitely many poles in $\mathrm{Re}(\lambda) > -\nu$ given by the λ_j (and λ_j is a pole of order $k_j + 1$). It is a bit more difficult but still true to prove that if $\langle R_P(\lambda) f_1, f_2\rangle$ has only finitely many poles in $\mathrm{Re}(\lambda) > -\nu$ for each $f_1, f_2 \in C^\infty(\mathfrak{M})$, with a polynomial bound

$$|\langle R_P(\lambda) f_1, f_2\rangle| \le C_{f_1,f_2}|\lambda|^p$$

for some constant C_{f_1,f_2} (depending bilinearly on f_1, f_2) and some $p \in \mathbb{R}$ independent of f_i, then an expansion of the form (2-4) holds true. This can be proved by some contour deformation when writing the operator $e^{tX} = \varphi_t^*$ in terms of resolvents, see for example Nonnenmacher and Zworski [2015, Corollary 5]. We will call the constant ν in (2-4) the *size of the essential spectral gap*.

Using representation theory, an exponential decay of correlations of mixing for geodesic flow on hyperbolic surfaces was proved by Ratner [1987] and it was extended to higher dimensions by Moore [1987]. In variable negative curvature for surfaces and more generally for Anosov flows with stable/unstable jointly non-integrable foliations, exponential decay of correlations was first shown by Dolgopyat [1998] and then by Liverani for contact flows Liverani [2004]. In these work, an essential gap of size $\epsilon > 0$ is shown, but $\epsilon > 0$ is not explicit.

Theorem 3 (Liverani [ibid.])**.** *For each contact Anosov flow, there is an essential spectral gap of positive size and the correlations decay exponentially fast.*

Later, the work of Tsujii [2010] gave a quantitative value for the size ν of the essential gap, and another proof appeared later in work of Nonnenmacher and Zworski [2015] (where they extended this result to general normally hyperbolic trapped sets).

Theorem 4 (Tsujii [2010, 2012] and Nonnenmacher and Zworski [2015])**.** *For contact Anosov flows, there is an essential gap of size ν for all $\nu < \nu_0$, where*

$$\nu_0 = \frac{1}{2}\Big(\liminf_{t\to\infty}\frac{1}{t}\inf_{y\in\mathfrak{M}}\log\det(d\varphi_t|_{E_u(y)})\Big)$$

More recently Tsujii proved that the contact assumption can be removed in dim 3, at least generically.

Theorem 5 (Tsujii [n.d.])**.** *On a manifold of dimension 3 admitting an Anosov flow, for $r \geq 3$ there is an open dense set in C^r of volume preserving Anosov flows that have an essential gap, and thus are exponentially mixing.*

For the billiard flow associated with a two-dimensional finite horizon Lorentz Gas (the Sinai billiard flow with finite horizon), Baladi, Demers, and Liverani [2018] recently proved an exponential decay of correlations and the existence of a non-explicit essential gap. In the Axiom A case, we note the result of Naud [2005] for hyperbolic convex co-compact surfaces and Stoyanov [2011, 2013] for more general cases proving an essential spectral gap; both results use Dolgopyat method. The recent work of Bourgain and Dyatlov [n.d.] gives an essential gap of size $1/2 + \epsilon$ for some $\epsilon > 0$ on all convex co-compact hyperbolic surfaces.

For contact Anosov flows with pinching of the Lyapunov exponents (holding for example in pinched negative curvature), Faure and Tsujii [2013, 2017] established the striking fact that the Ruelle resonance spectrum has a band structure.

Theorem 6 (Faure and Tsujii [2013, 2017]). *Let X be a contact Anosov flow on manifold $\mathfrak{M}$. There is $C > 0$ such that for each $\epsilon > 0$ the resonance spectrum in $|\mathrm{Im}(\lambda)| > C$ is contained in the union over $k \in \mathbb{N}_0$ of the bands*

$$B_k := \{\lambda \in \mathbb{C}; \mathrm{Re}(\lambda) \in [\gamma_k^- - \epsilon, \gamma_k^+ + \epsilon]\}$$

where

$$\gamma_k^+ := \lim_{t\to\infty} \sup_{y\in\mathfrak{M}} \frac{1}{t}\Big(-\frac{1}{2}\int_0^t \mathrm{div}(X|_{E_u})(\varphi_s(y))ds - k\log\big|\big|(d\varphi_t|_{E_u(y)})^{-1}\big|\big|^{-1}\Big),$$

$$\gamma_k^- := \lim_{t\to\infty} \inf_{y\in\mathfrak{M}} \frac{1}{t}\Big(-\frac{1}{2}\int_0^t \mathrm{div}(X|_{E_u})(\varphi_s(y))ds - k\log\big|\big|d\varphi_t|_{E_u(y)}\big|\big|\Big)$$

and $\mathrm{div}(X|_{E_u}) := (\partial_t \log\det d\varphi_t|_{E_u})|_{t=0} > 0$.

This band structure was first observed in related (but different) settings in the works by Faure [2007], Faure and Tsujii [2015] and by Dyatlov [2015]. We note that only finitely many bands B_k do not intersect except for geodesic flows in constant negative curvature where the Lyapunov exponents are constants: in curvature -1, $\gamma_k^- = \gamma_k^+ = -\frac{n}{2} - k$ if the dimension of the Riemannian manifold is $n+1$. Actually, in that setting, the manifold M is a quotient of hyperbolic space $\mathbb{H}^{n+1}$ by a co-compact group Γ, and the Ruelle resonance spectrum for the flow on SM has been (almost) completely characterised by Dyatlov, Faure, and Guillarmou [2015]: there is a one-to-one correspondence between the Ruelle resonances/resonant eigenstates with the spectrum/eigenfunctions of some Bochner Laplacians on certain bundles over M.

Theorem 7 (Dyatlov, Faure, and Guillarmou [ibid.]). *Let M be a compact hyperbolic manifold of dimension $n+1 \geq 2$. Assume that $\lambda \in \mathbb{C} \setminus \big(-\frac{n}{2} - \frac{1}{2}\mathbb{N}_0\big)$. Denote by $m_X(\lambda)$ the multiplicity of $\lambda \in \mathbb{C}$ as a Ruelle resonance for the geodesic flow X on SM, and let $\Delta_k = \nabla^*\nabla$ be the rough Laplacian on the space of trace-free divergence-free symmetric tensors of order k. Then for $\lambda \notin -2\mathbb{N}$, we have*

$$m_X(\lambda) = \sum_{m\geq 0}\sum_{\ell=0}^{\lfloor m/2\rfloor} \dim\ker\Big(\Delta_{m-2\ell} + (\lambda + m + \frac{n}{2})^2 - \frac{n^2}{4} - m + 2\ell\Big)$$

and for $\lambda \in -2\mathbb{N}$, we have

$$m_X(\lambda) = \sum_{\substack{m\geq 0\\ m\neq -\lambda}}\sum_{\ell=0}^{\lfloor m/2\rfloor} \dim\ker\Big(\Delta_{m-2\ell} + (\lambda + m + \frac{n}{2})^2 - \frac{n^2}{4} - m + 2\ell\Big).$$

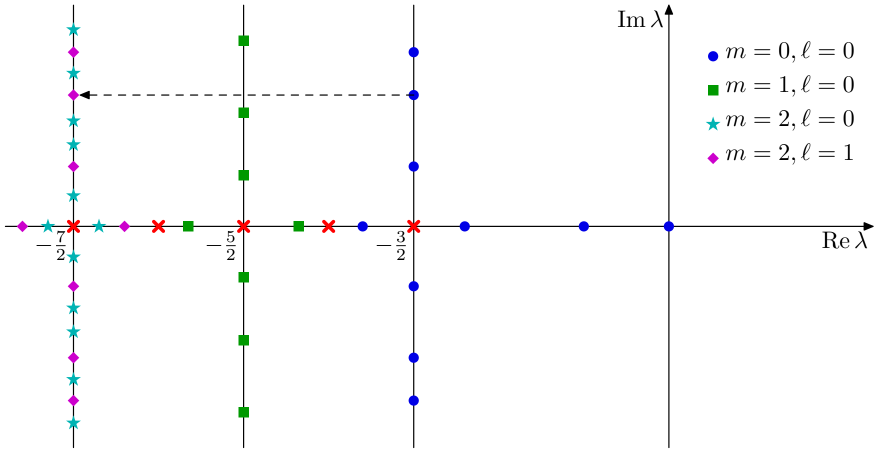


Figure 1: An illustration of Theorem 7 for $n=3$. The red crosses mark exceptional points where the theorem does not apply. Note that the points with $m=2,\ell=1$ are simply the points with $m=0,\ell=0$ shifted by -2 (modulo exceptional points), as illustrated by the arrow.

The first band (actually line) of Ruelle resonances appear at $\mathrm{Re}(\lambda)=-\frac{n}{2}+i\mathbb{R}$, they correspond to the spectrum of the Laplacian on functions ($m=\ell=0$ in Theorem 7). In dimension $n+1=2$, i.e. for surfaces, the statement is simpler since the space of trace-free divergence-free tensors is finite dimensional. Then the resonance spectrum is simply

$$\Big(-\mathbb{N}_0+\bigcup_{1/4+r_j^2\in\sigma(\Delta)}(-\tfrac{1}{2}+ir_j)\Big)\bigcup(-\mathbb{N})$$

where $\sigma(\Delta_0)$ denotes the spectrum of the Laplacian Δ_0 acting on functions on $\Gamma\backslash\mathbb{H}^2$ (for the analysis of the special points $-\mathbb{N}/2$, see Guillarmou, Hilgert, and Weich [n.d.(a)]). In fact, in Dyatlov, Faure, and Guillarmou [2015], we show an explicit correspondence between the resonant states and the eigenfunctions of $\Delta_{m-2\ell}$ on M: for example, for the first band $m=0,\ell=0$, the correspondence is given by the pushforward map (integration in the fibers of SM)

$$\pi_{0*}:\ker_{\mathcal{H}^N}(-X-\lambda)\to\ker(\Delta_0+\lambda(n+\lambda)),\quad \pi_{0*}u(x):=\int_{S_xM}u(x,v)dv.$$

where $\mathcal{H}^N$ is an anisotropic Sobolev space as in Theorem 1 with $N\gg 1$ large enough. We call this a *classical-quantum correspondence* between the eigenspaces. A partial generalisation to all compact rank-1 locally symmetric spaces is done by Guillarmou, Hilgert, and Weich [n.d.(b)]. The case of the flow acting on sections of certain bundles is worked out by Küster and Weich [n.d.].

In the case of convex co-compact hyperbolic manifolds, where the flow is Axiom A and the Laplacian has continuous spectrum, the description of the Ruelle resonance spectrum and the classical-quantum correspondence has been done completely in dimension 2 by Guillarmou, Hilgert, and Weich [n.d.(a)], and outside the special points $-\frac{n}{2}-\frac{1}{2}\mathbb{N}$ in higher dimension by Hadfield [n.d.].

In the analysis of the first band of resonances and resonant states for hyperbolic manifolds, we strongly use a differential operator $U_- : C^\infty(SM) \to C^\infty(SM; E_s^*)$ that is a covariant derivative in the direction of the unstable space:

$$\xi \in E_u(y), \quad U_- f(y)\xi := df(y)\xi,$$

recall that $E_s^*(E_s \oplus \mathbb{R}X) = 0$ thus E_s^* is a dual space to E_u. We prove that the resonant states associated to the first band are characterised as the solutions $u \in \mathfrak{D}'(SM)$, $Xu = -\lambda u$ with $U_- u = 0$ for the compact case, and with the additional condition $\operatorname{supp}(u) \subset \Gamma_+$ for the convex co-compact case. Such distributions u can be lifted to $S\mathbb{H}^{n+1}$ and are in one-to-one correspondence with distributions on $\partial\mathbb{H}^{n+1} = \mathbb{S}^n$ that have a particular conformal covariance with respect to the group Γ. The correspondence is done via pullback through the backward endpoint map $B_- : S\mathbb{H}^{n+1} \to \partial\mathbb{H}^{n+1}$ defined by $B_-(y) := \lim_{t\to+\infty} \exp_x(-tv)$. Applying the Poisson transform to those distributions we find eigenfunctions for Δ_0 on $\mathbb{H}^{n+1}$ that are Γ-equivariant, thus descend to $\Gamma\backslash\mathbb{H}^{n+1}$.

In variable curvature or more generally contact flows, a similar covariant derivative U_- can be defined, but the stable/unstable bundles are only Hölder continuous thus applying U_- to $\mathfrak{D}'(SM)$ does not make sense. In dimension $\dim(SM) = 3$, for contact flows, the operator U_- has $C^{2-\epsilon}(SM)$ coefficients for all $\epsilon > 0$ by Hurder and Katok [1990]. The first band of resonances has resonant states that are regular enough to apply U_- and we show with Faure that the rigidity $U_- u = 0$ of resonant states in constant curvature still holds in variable curvature.

Theorem 8 (Faure and Guillarmou [n.d.]). *Let $\mathfrak{M}$ be a smooth 3-dimensional compact oriented manifold and let X be a smooth vector field generating a contact Anosov flow. Assume that the unstable bundle is orientable. If λ_0 is a resonance of X with $\operatorname{Re}(\lambda_0) > -\mu_{\min}$ and if u is a generalised resonant state of P with resonance λ_0, then $U_- u = 0$. Here $\mu_{\min}$ is the minimal expansion rate given by*

$$\mu_{\min} := \lim_{t\to+\infty} \inf_{z\in\mathfrak{M}} -\frac{1}{t}\log\left|d\varphi_t(z)|_{E_s(z)}\right| = \lim_{t\to+\infty} \inf_{z\in\mathfrak{M}} -\frac{1}{t}\log\left|d\varphi_{-t}(z)|_{E_u(z)}\right|.$$

We also remark that for Morse-Smale flows, i.e. with finitely many hyperbolic fixed points and finitely many hyperbolic periodic orbits, the Ruelle spectrum has been computed explicitly by Dang and Rivière [n.d.(a)] (among other things dealt with in the article).

2.4 Dynamical zeta functions. Consider a smooth vector field X on $\mathcal{M}$ and $\mathbf{X} : C^\infty(\mathcal{M};\mathcal{E}) \to C^\infty(\mathcal{M};\mathcal{E})$ a first order differential operator on a bundle $\mathcal{E}$ satisfying

$$\forall f \in C^\infty(\mathcal{M}), \forall u \in C^\infty(\mathcal{M};\mathcal{E}),\ \mathbf{X}(fu) = (Xf)u + f(\mathbf{X}u).$$

Define the vector bundle $\mathcal{E}_0$ by

$$\mathcal{E}_0(x) = \{\eta \in T_x^*\mathcal{M} \mid \langle X(x), \eta\rangle = 0\}, \quad x \in \mathcal{M}$$

and the *linearized Poincaré map* by

$$\mathcal{P}_{x,t} : \mathcal{E}_0(x) \to \mathcal{E}_0(\varphi_t(x)), \quad \mathcal{P}_{x,t} = (d\varphi_t(x)^{-1})^T|_{\mathcal{E}_0(x)}.$$

Next, the parallel transport $\alpha_{x,t} : \mathcal{E}(x) \to \mathcal{E}(\varphi_t(x))$ is defined as follows: for each $\mathbf{u} \in C^\infty(\mathcal{M};\mathcal{E})$, we put $\alpha_{x,t}(\mathbf{u}(x)) = e^{-t\mathbf{X}}\mathbf{u}(\varphi_t(x))$. Now, assume that $\gamma(t) = \varphi^t(x_0)$ is a closed trajectory, that is $\gamma(T) = \gamma(0)$ for some $T > 0$. (We call T the period of γ, and regard the same γ with two different values of T as two different closed trajectories. The minimal positive $T^\sharp$ such that $\gamma(T^\sharp) = \gamma(0)$ is called the *primitive period*.) For such closed orbit γ, we define $\mathcal{P}_\gamma = \mathcal{P}_{x,T}$ where x is any point on the closed orbit γ, and similarly $\alpha_\gamma = \alpha_{x,T}$. We can note that $\mathrm{Tr}(\alpha_\gamma)$ and $|\det(1 - \mathcal{P}_\gamma)|$ are independent of the choice of x on the closed orbit.

Giulietti, Liverani, and Pollicott [2013], and then Dyatlov and Zworski [n.d.], show the meromorphic extension of the zeta function for the flow and of the Ruelle zeta function for Anosov flows. In the Axiom A case, this is proved by Dyatlov and Guillarmou [2016].

Theorem 9 (Giulietti, Liverani, and Pollicott [2013], Dyatlov and Zworski [n.d.], and Dyatlov and Guillarmou [2016]). *1) Define for* $\mathrm{Re}\,\lambda \gg 1$*, the dynamical zeta function for* $\mathbf{X}$

$$Z_{\mathbf{X}}(\lambda) := \sum_\gamma \frac{e^{-\lambda T_\gamma}\, T_\gamma^\sharp\, \mathrm{Tr}(\alpha_\gamma)}{|\det(I - \mathcal{P}_\gamma)|} \tag{2-5}$$

where the sum is over all closed trajectories γ inside $\mathcal{M}$ (resp. inside $\mathcal{K}$) in the Anosov case (resp. in the Axiom A case), $T_\gamma > 0$ is the period of γ, and $T_\gamma^\sharp$ is the primitive period. Then $Z_{\mathbf{X}}(\lambda)$ extends meromorphically to $\lambda \in \mathbb{C}$. The poles of $Z_{\mathbf{X}}(\lambda)$ are the

Ruelle resonances of $\mathbf{X}$ *and the residue at a pole* λ_0 *is equal to* $\mathrm{Rank}(\mathrm{Res}_{\lambda_0}(-\mathbf{X}-\lambda)^{-1})$.
2) The Ruelle zeta function defined by

$$\zeta(\lambda) := \prod_{\gamma^\sharp}\big(1-\exp(-T_{\gamma^\sharp}(\lambda))\big), \quad \mathrm{Re}\,\lambda \gg 1.$$

admits a meromorphic continuation to $\mathbb{C}$ *both in the Anosov and the Axiom A case.*

These results answer positively a conjecture of Smale. In the work Faure and Tsujii [2015], Faure and Tsujii define a Gutzwiller-Voros dynamical zeta function associated to the operator $-X+\frac{1}{2}\mathrm{div}(X|_{E_u})$ and show that its zeros in a band $\mathrm{Re}(\lambda) > -C$ for some $C>0$ are located asymptotically close to the imaginary line as $\mathrm{Im}(\lambda)\to\infty$, using their band structure results of Theorem 6 (with the suitable potential added). This zeta function is the natural generalisation of Selberg's zeta function in variable curvature.

3 Boundary rigidity and X-ray tomography problems

The boundary and lens rigidity problems are inverse problems consisting in determining a Riemannian manifold (M,g) with boundary from boundary measurements on the geodesic flow. As boundary data, we employ the *boundary distance function*

$$\beta_g := d_g|_{\partial M\times\partial M}, \tag{3-1}$$

where $d_g : M\times M \to [0,\infty)$ is the Riemannian distance, and the *lens data*

$$\tau_g^+ : \partial SM \to [0,\infty], \qquad \sigma_g : \partial SM\setminus\Gamma_- \to \partial SM.$$

Here, SM denotes the unit tangent bundle, $\Gamma_- := \{y\in\partial SM \mid \tau_g^+(y)=+\infty\}$, the *exit time* $\tau_g^+(x,v)$ is the maximal non-negative time of existence of the geodesic $\gamma_{x,v}(t)=\exp_x(tv)$, and the *scattering map* $\sigma_g(x,v) := (\gamma_{x,v}(\tau_g^+(x,v)), \dot\gamma_{x,v}(\tau_g^+(x,v)))$ gives the exit position and "angle" of $\gamma_{x,v}$. When τ_g^+ is everywhere finite, (M,g) is said to be *non-trapping*. The *boundary rigidity problem* asks whether the boundary distance β_g determine (M,g) up to diffeomorphisms fixing ∂M. Analogously, the *lens rigidity problem* asks whether the lens data (τ_g^+,σ_g) determine (M,g) up to diffeomorphisms fixing ∂M. For *simple* Riemannian manifolds, that is, compact Riemannian balls with strictly convex boundary and without conjugate points, these two rigidity problems are equivalent, since the boundary distance and the lens data can be easily recovered from each another. When the manifold has non-empty trapped set, non convex boundary or conjugate points, this equivalence is not in general true. There are easy counter examples to boundary rigidity when there are non-minimizing length geodesics (see C. B. Croke [1991]). We will say that a metric is *deformation lens rigid* if any one-parameter family of metrics with the same lens data are isometric.

In the closed setting, there is a corresponding problem that consists in determining a metric from the length of its closed geodesics, called the *length spectrum*. Here again there are counter examples due to Vignéras [1980] of non-isometric hyperbolic surfaces with same length spectrum. It is therefore more appropriate to ask wether the *marked length spectrum* determine the metric, where the marking of the geodesics is made using free homotopy classes (recall that in negative curvature there is a unique closed geodesic in each free homotopy class).

One way to attack these problems is the consider the linearised problem, which consists in analysing the kernel of the geodesic X-ray transform on symmetric tensors or order 2. More generally, the X-ray on the bundle $\otimes_S^m T^*M$ of symmetric tensors of order $m \in \mathbb{N}_0$ is defined in the case with boundary as

$$I_m : C^0(M; \otimes_S^m T^*M) \to L^\infty_{\mathrm{loc}}(\mathcal{G}), \quad I_m(f)(\gamma) := \int_0^{\ell_\gamma} f(\gamma(t))(\otimes^m \dot\gamma(t))dt$$

where $\mathcal{G}$ denotes the set of geodesics γ in SM with endpoints on the boundary ∂SM, thus having finite length $\ell(\gamma) < \infty$. In the closed case, a similar definition holds with $\mathcal{G}$ being the set of closed geodesics and $I_m(f)(\gamma)$ is defined as above but normalized by $1/\ell(\gamma)$ so that it is an $L^\infty(\mathcal{G})$ function. It is easy to check that $I_m(Df) = 0$ if $f \in C^1(M; \otimes_S^{m-1}T^*M)$ satisfies $f|_{\partial M} = 0$ and D is the symmetrised Levi-Civita covariant derivative (the condition $f|_{\partial M} = 0$ is obviously removed in the closed manifold setting). In general, the best one can get is injectivity of I_m on $\ker D^*$, i.e. divergence-free tensors, which is called *solenoidal injectivity*.

3.1 Simple metrics. Simple manifolds were introduced by Michel [1981/82] and can be defined as manifolds (M, g) that are a topological ball with strictly convex boundary and so that g has no conjugate points. Their exponential map is a diffeomorphism at each point $x \in M$. In particular, there is a unique geodesic between each pair of points $x, x' \in M$, and this geodesic has length $d_g(x, x')$. Michel made the conjecture that two simple manifolds (M, g_1) and (M, g_2) with same boundary distance $\beta_{g_1} = \beta_{g_2}$ verify that there is $\psi : M \to M$ such that $\psi^* g_2 = g_1$ and $\psi|_{\partial M} = \mathrm{Id}$. As mentionned above, the boundary rigidity and lens rigidity questions are equivalent in that setting.

For negatively curved and non-positively curved cases, it was shown by Otal [1990b] and C. B. Croke [1990] that the conjecture holds in dimension 2.

Theorem 10 (C. B. Croke [1990] and Otal [1990b]). *Two non-positively curved simple surfaces with the same boundary distance are isometric via an isometry fixing the boundary.*

In fact, in these works, we notice that the boundary is even allowed to be non-convex. The full conjecture in dimension 2 was later proved by L. Pestov and Uhlmann [2005], using earlier works of Muhometov [1981] which allow to recover a conformal factor from the boundary distance.

Theorem 11 (L. Pestov and Uhlmann [2005]). *Two simple surfaces with the same boundary distance are isometric via an isometry fixing the boundary. Moreover the scattering map σ_g determines the conformal class of a simple manifold.*

In higher dimension, little was known until recently. Burago and Ivanov [2010] proved that metrics close to flat simple metrics are boundary rigid and Stefanov and Uhlmann [2005] showed that generic simple metrics are boundary rigid. A more recent work due to Stefanov, Uhlmann, and Vasy [n.d.(b)] solves Michel's conjecture in the category of non-positively curved simple metrics, they even show a local result near boundary points.

Theorem 12 (Stefanov, Uhlmann, and Vasy [ibid.]). *Two simple manifolds in dimension $n \geq 3$ which are non-positively curved and with the same boundary distance are isometric via an isometry fixing the boundary. Moreover the boundary distance near a point $p \in \partial M$ determines the metric near p in M.*

In Stefanov, Uhlmann, and Vasy [ibid.], the condition for rigidity is weaker than non-positive curvature: it is asked that the manifolds are foliated by strictly convex hypersurfaces.

The analysis of X-ray transform is the main tool in the proof of Theorem L. Pestov and Uhlmann [2005] and Stefanov, Uhlmann, and Vasy [n.d.(b)]. These proofs are essentially of analytic nature, contrary to C. B. Croke [1990], Otal [1990b], and Burago and Ivanov [2010] where the method is more geometric. Let us quickly review some known results on this linearised problem, that is the injectivity of the X-ray transform. For simple metrics, I_0 and I_1 are known to be solenoidal injective, this was proved by Muhometov [1981] for I_0 and Anikonov and Romanov [1997] for I_1. In dimension 2, the injectivity on I_2 follows from L. Pestov and Uhlmann [2005] and the injectivity of I_m for $m > 2$ was only proved recently by Paternain, Salo, and Uhlmann [2013]. In dimension $n > 2$ and for $m \geq 2$, the injectivity of I_m in non-positive curvature was proved by L. N. Pestov and Sharafutdinov [1988]. The main tool that is used in these cases is an energy identity called *Mukhometov-Pestov identity*. We will review it quickly in the next section. We also notice that a local injectivity result (i.e. we consider the X-ray transform of a tensor only on an open subset of geodesics) has been recently proved by Uhlmann and Vasy [2016] for I_0 and Stefanov, Uhlmann, and Vasy [n.d.(a)] for I_1 and I_2 using new microlocal methods.

Theorem 13 (Uhlmann and Vasy [2016] and Stefanov, Uhlmann, and Vasy [n.d.(a)]). *Let (M, g) be a Riemannian manifold of dimension $n \geq 3$, and assume that $p \in \partial M$ is such that ∂M is strictly convex at p.*
1) Let $f \in C^\infty(M)$ and assume that $I_0(f)(\gamma) = 0$ for all γ passing through a small neighborhood of $T_p\partial M$, i.e. γ are short geodesics that are almost tangent to ∂M. Then $f = 0$ near p.
2) Let $f \in C^\infty(M; \otimes_S^m T^*M)$ with $m \in \{1, 2\}$ such that $f = u + Dv$ with $v|_{\partial M} = 0$. If $I_m(f)(\gamma) = 0$ for all γ passing through a small neighborhood of $T_p\partial M$, then $u = 0$ near p.

This is the local result that allows Stefanov, Uhlmann, and Vasy [n.d.(b)] to prove Theorem 12 through a layer stripping method. The proof uses the scattering calculus of Melrose to analyse the normal operator $I_m^* I_m$. An artificial boundary is put near p in order to make the local analysis a global one on a new manifold, and the normal operator is somehow replaced by a localised one $I_m^* \chi I_m$ for some well chosen function $\chi(\gamma)$. For simple manifolds, the normal operator $I_m^* I_m$ is a pseudo-differential operator of order -1 that is elliptic on $\ker D^*$, this is a quite helpful fact to analyse the Fredholm properties and closed range properties of the operators of interest. The presence of conjugate points would ruin this property. In Uhlmann and Vasy [2016], the localisation using the χ in $I_m^* \chi I_m$ allows for example to avoid conjugate points since the geodesics almost tangent to ∂M are short and thus free of conjugate points, showing that $I_m^* \chi I_m$ is also pseudo-differential. In dimension 3 there are enough directions to get ellipticity of $I_m^* \chi I_m$, which is not the case in dimension $n = 2$, and in fact Uhlmann and Vasy [ibid.] show that this is a strong enough ellipticity to obtain injectivity (full ellipticity in the scattering calculus of Melrose).

3.2 Cases with trapped set, conjugate points or non-convex boundary. There are three different ways a manifold can be not simple: it has non-empty trapped set, it has non-convex boundary or pairs of conjugate points.

Trapped case. First, let us mention some recent results for the case with trapped set. In Guillarmou [2017b], we address the case where the trapped set is a hyperbolic set for the geodesic flow. For example, this condition is always satisfied in negative curvature. We consider a manifold (M, g) with strictly convex boundary, hyperbolic trapped set and no conjugate points. The simplest example is a hyperbolic cylinder with one closed geodesic. We are able to show injectivity of the X-ray transform on tensors.

Theorem 14 (Guillarmou [ibid.]). *Let (M, g) be a manifold with strictly convex boundary, hyperbolic trapped set and no conjugate points. The ray transforms I_0 and I_1 are* *solenoidal injective. If in addition the curvature of g is non-positive, I_m is solenoidal injective for all $m \geq 2$. Such a manifold is deformation lens rigid.*

This result shows in particular that all negatively curved manifold with strictly convex boundary have solenoidal injective ray transform for all tensors and are deformation lens rigid. The proof uses two steps, one is purely of dynamical system nature and is a Livsic type result (although Livsic theorem is usually for integration on closed orbits). Here $\mathfrak{M}$ will typically be the unit tangent bundle SM of M, where the geodesic flow lives.

Theorem 15 (Guillarmou [2017b]). *Let $\mathfrak{M}$ be a manifold with boundary and X a non-vanishing smooth vector field with trapped set $\mathfrak{K} \subset \mathfrak{M}^\circ$ that is hyperbolic, and assume that $\partial\mathfrak{M}$ is strictly convex for the flow of X. If $f \in C^\infty(\mathfrak{M})$ vanishes to infinite order at $\partial\mathfrak{M}$ and satisfies $\int_\gamma f = 0$ for all integral curve γ with endpoints on $\partial\mathfrak{M}$, then there exists $u \in C^\infty(\mathfrak{M})$ such that $Xu = f$ and $u|_{\partial M} = 0$.*

If $I_0 f = 0$ we deduce from some classical argument using the short geodesics near ∂M that f vanishes at ∂M to infinite order, we can then apply Theorem 15 to get $u \in C^\infty(SM)$ such that $Xu = \pi_0^* f$ with $u|_{\partial SM} = 0$, where $\pi_0^* : C^\infty(M) \to C^\infty(SM)$ is the pull-back by the projection $\pi_0 : SM \to M$ on the base of the fibration. The Mukhometov-Pestov identity is the following identity: if $\dim(M) = n$, for each $w \in H^2(SM) \cap H_0^1(SM)$

$$||\nabla^v X w||^2_{L^2(SM)} = ||X\nabla^v w||^2_{L^2(sM)} + (n-1)||Xw||^2_{L^2(SM)} - \langle \mathfrak{R}\nabla^v w, \nabla^v w\rangle_{L^2(SM)}.$$

Here $\nabla^v w = P^v \nabla w$ where ∇ is the gradient for the Sasaki metric and P^v is the orthogonal projection on the vertical space $\ker d\pi_0$ with respect to the same metric, $\mathfrak{R}$ is a natural operator made from the Riemann curvature tensor. Applying to $w = u$, the left hand side is 0 since $\nabla^v \pi_0^* = 0$ and the quantity $||X\nabla^w||^2_{L^2} - \langle \mathfrak{R}\nabla^v w, \nabla^v w\rangle_{L^2(SM)} \geq 0$ when there are no conjugate points, using the index theory for the energy functional of curves. This implies $Xu = 0$, thus $f = 0$. A similar argument works for I_1, and also for higher order tensors provided the curvature is non-positive.

We notice that a surface containing a flat cylinder is such that I_0 has infinite dimensional kernel (at least if I_0 maps to the space of geodesics γ with endpoints on the boundary), thus the hyperbolicity condition on the trapped set is somehow a condition that might be difficult to remove to get injectivity of X-ray in other trapped situations.

Using Theorem 14, we are able to show a "Pestov-Uhlmann" type result for surfaces.

Theorem 16 (Guillarmou [ibid.]). *Let (M_1, g_1) and (M_2, g_2) be two Riemannian surfaces with strictly convex boundary, hyperbolic trapped set and no conjugate points. Assume that $\partial M_1 = \partial M_2$ and that their scattering maps agree, i.e. $\sigma_{g_1} = \sigma_{g_2}$, then there is a diffeomorphism $\psi : M_1 \to M_2$ such that $\psi^* g_2 = e^\rho g_1$ for some $\rho \in C^\infty(M_1)$ vanishing at ∂M_1.*

So far we are not able to prove that the lens data allows to determine the remaining conformal factor ρ in Theorem 16, although we believe it does. However, in a work with Guillarmou and Mazzucchelli [2016], we show a marked lens rigidity result for the same class of Riemannian surface, that is the lens data in the universal cover (or equivalently the boundary distance in the universal cover) determine the metric.

The proof of Theorem 16 is quite complicated and uses the approach of L. Pestov and Uhlmann [2005], that mainly reduces the problem to showing solenoidal injectivity of I_1 and surjectviity of I_0^*, the dual transform to I_0 with respect to some natural measure on the set $\mathcal{G}$ of geodesics. We already know from Theorem 14 that I_0, I_1 are solenoidal injective. To prove surjectivity of I_0^*, the strategy is to prove that $I_0^* I_0$ is a Fredholm operator. We can check that

$$I_0^* I_0 = -2\pi_{0*} R_X(0) \pi_0^*$$

where $R_X(\lambda) = (-X - \lambda)^{-1}$ is the resolvent of the flow studied in Theorem 2, π_0^* is as above and π_{0*} is its adjoint consisting in integration in fibers. In the paper Dyatlov and Guillarmou [2016] with Dyatlov, we actually characterised the wave-front set of the Schwartz kernel of $R_X(\lambda)$ using propagation of singularities with radial points. Basically, writing $R_X(0) = -\int_{-\infty}^{0} e^{tX} dt$, and using that $e^{tX} = \varphi_t^*$ is a Fourier integral operator with well-known wave-front set, we already see that the conormal to the diagonal is in the wave-front set (the contribution of $t = 0$ in the integral) as well as the graph of the symplectic flow $\Phi_t = (d\varphi_t^{-1})^T$ on $T^*(SM)$. Another component appears from long time propagation, and that is where the propagation with radial point shows up, it is given by $E_+^* \times E_-^*$. Using standard rules for composition of wave-front sets, applying the push-forward $\pi_{0*} \otimes \pi_{0*}$ to the Schwartz kernel of $R_X(0)$, everything in the wave-front disappears except the conormal to the diagonal: this is a consequence of the no-conjugate points assumption and the fact that $E_\pm$ is transversal to the vertical space $\ker d\pi_0 \subset T(SM)$ in the characteristic set $\{\xi \in T^*(SM); \xi(X) = 0\}$.

There are a couple of other rigidity results in the trapped case, due to C. B. Croke and Herreros [2016] and C. Croke [2014]: in C. B. Croke and Herreros [2016] it is shown that a 2-dimensional negatively curved or flat cylinder with convex boundary is lens rigid, and C. Croke [2014] proved that the flat product metric on $B_n \times S^1$ is scattering rigid if B_n is the unit ball in $\mathbb{R}^n$.

Non-convex boundary. When the boundary is non-convex, there are also complications: the boundary distance is not a priori directly related to the lens data. In fact, it is shown to be the case for simply connected surfaces with no conjugate points by Guillarmou, Mazzucchelli, and Tzou [n.d.]. It is probably not true in higher dimension due to the fact that there are simply connected manifolds with boundary having geodesics with endpoints on ∂M and that are not length minimizing. The determination of the C^∞-jet in that case is also more complicated since there does not exist small geodesics near points in ∂M where ∂M is concave. This determination has however been proved by Stefanov and Uhlmann [2009] in the non-trapping with no-conjugate points case (and certain trapped cases). The injectivity of the X-ray transform for non-trapping manifolds with no-conjugate points has been proved by Dairbekov [2006] and extended by Guillarmou, Mazzucchelli, and Tzou [n.d.] to the case where the trapped set is a hyperbolic set not intersecting the boundary.

Theorem 17 (Dairbekov [2006] and Guillarmou, Mazzucchelli, and Tzou [n.d.])**.** *Assume that (M, g) has no conjugate points and that its trapped set $\mathcal{K}$ does not intersect ∂SM, then I_0 and I_1 are solenoidal injective. Moreover I_m is solenoidal injective if in addition the curvature is non-positive.*

Recall that the boundary rigidity results of Otal [1990b], C. B. Croke [1990] and Burago and Ivanov [2010] do not involve convexity of the boundary. In Guillarmou, Mazzucchelli, and Tzou [n.d.], we are recently able to extend L. Pestov and Uhlmann [2005], Otal [1990b], and C. B. Croke [1990] to non-trapping manifolds with no conjugate points, a class that is more general than simple manifolds.

Theorem 18 (Guillarmou, Mazzucchelli, and Tzou [n.d.])**.** *1) Let M be a simply connected compact surface with boundary. If g_1 and g_2 are two Riemannian metrics on M without conjugate points such that $\beta_{g_1} = \beta_{g_2}$, then there is a diffeomorphism $\psi : M \to M$ such that $\psi|_{\partial M} = \mathrm{Id}$ and $\psi^* g_2 = g_1$.*
2) Let (M_1, g_1) and (M_2, g_2) be two non-trapping, oriented compact Riemannian surfaces with boundary, without conjugate points, and with the same lens data. Then there exists a diffeomorphism $\psi : M_1 \to M_2$ such that $\psi^* g_2 = g_1$.

This result uses the method of Pestov-Uhlmann and a careful analysis near the glancing trajectories to be able to show that I_0^* is a surjective operator. Working with the normal operator $I_0^* I_0$ in order to prove this property would not be a very good idea since this operator has problematic singularities due to glancing geodesics: it is not a pseudo-differential operator anymore as in the simple manifold case. We thus have to consider a modified normal operator that separates the glancing trajectories from the non-glancing ones. We then show that the scattering map σ_g determines (M, g) up to conformal diffeomorphism. The lens rigidity result in the non-simply connected case in Theorem 18 also uses some unpublished work of Zhou [2011] done in his PhD thesis under Croke's direction; this work is based on a result of C. Croke [2005] on lens rigidity for finite quotients. We finally conjecture that Theorem 18 should be true also in higher dimension; this would be a more general result than Michel's conjecture.

Conjugate points. Very little is known in cases with conjugate points. It is conjectured that non-trapping manifolds should have injective X-ray transform, even when there are conjugate points, but so far this conjecture remains open. There has been some recent analysis of the normal operator $I_0^* I_0$ by Stefanov and Uhlmann [2012], Monard, Stefanov, and Uhlmann [2015], Bao and Zhang [2014] and Holman and Uhlmann [n.d.] who prove that this is a Fourier integral operator under certain assumptions on the type of conjugate points. In certain cases, this implies in dimension $n \geq 3$ that the kernel of I_0 is finite dimensional. We also note that the results of Uhlmann and Vasy [2016] deal with certain cases with conjugate points in dimension $n \geq 3$, under the foliation by convex hypersurfaces condition.

3.3 Closed manifolds.

For closed Riemannian manifolds with Anosov geodesic flow, the main result is due to Otal [1990a] and C. B. Croke [1991] who proved the following:

Theorem 19 (Otal [1990a] and C. B. Croke [1991]). *Two closed Riemannian surface with negative curvature and with the same marked length spectrum are isometric.*

The method of proof is purely geometric: Otal proves first that the geodesic flows for the two metrics are conjugate, and that the conjugation preserves the Liouville measure. Then he uses a sequence of clever arguments based on Gauss-Bonnet formula for triangles to show that the conjugation of the flows comes from an isometry. An extension to certain manifolds with non-positive curvature has been obtained by C. Croke, Fathi, and Feldman [1992].

For manifolds conformal one to each other, the fact that the marked length spectrum determines the conformal factor has been proved by Katok [1988]; the proof is in dimension 2 but extends to higher dimension.

Maybe the first works on this topic were done by Guillemin and Kazhdan [1980a] and Guillemin and Kazhdan [1980b], where they proved deformation rigidity of the length spectrum in negative curvature for surfaces. This was extended by C. B. Croke and Sharafutdinov [1998] in higher dimension and in the Anosov setting for surfaces by Paternain, Salo, and Uhlmann [2014] and Guillarmou [2017a].

Theorem 20 (Guillemin and Kazhdan [1980a], C. B. Croke and Sharafutdinov [1998], Paternain, Salo, and Uhlmann [2014], and Guillarmou [2017a]). *1) Let g_s be a one-parameter family of negatively curved metrics on a closed manifold M. If g_s have the same length spectrum for all small $s \in (-\epsilon, \epsilon)$, then $g_s = \psi_s^* g_0$ for some smooth family of isometries ψ_s for s small.*
2) Let g_s be a one-parameter family of metrics with Anosov flows on a closed surface M. If g_s have the same length spectrum for all small $s \in (-\epsilon, \epsilon)$, then $g_s = \psi_s^* g_0$ for some smooth family of isometries ψ_s for s small.

These results are direct consequences of the solenoidal injectivity of the X-ray transform I_2, which is proved in negative curvature using Livsic theorem and a Mukhometov-Pestov energy identity in the same spirit as what we explained above for manifolds with boundary. For the Anosov case without negative (or non-positive) curvature assumption, the methods of Paternain, Salo, and Uhlmann [2014] and Guillarmou [2017a] use surjectivity of I_1^* and the complex structure of Riemann surfaces. The surjectivity of I_1^* consists in the construction of invariant distribution by the flow that have prescribed first Fourier coefficient in the Fourier decomposition in the fibers (that are circles). It is shown in Paternain, Salo, and Uhlmann [2014] that surjectivity of I_1^* follows from solenoidal injectivity of I_1, and that it implies surjectivity of I_2^* using Max Noether theorem, which in turn implies solenoidal injectivity of I_2. The argument is extended in Guillarmou [2017a] to prove solenoidal injectivity of I_m for all m for surfaces with Anosov geodesic flows.

Theorem 21 (Guillarmou [ibid.])**.** *If (M, g) is a closed surface with Anosov geodesic flow, I_m is solenoidal injective.*

The microlocal approach for hyperbolic flows of Faure and Sjöstrand [2011] and Dyatlov and Zworski [n.d.] and Theorem 1 is strongly used by Guillarmou [2017a] to show that the invariant distributions constructed in the surjectivity of I_1^* can be multiplied, through a careful analysis of their wave-front sets (shown to be contained in $E_s^* \cup E_u^*$).

We also mention that in the work Guillarmou [ibid.], we obtain new direct proofs of the regularity theory for the Livsic cohomological equation $Xu = f$ of Anosov flows, including in Sobolev spaces (which was not done), extending some results of de la Llave, Marco, and Moriyón [1986] and Journé [1986].

References

Yu. E. Anikonov and V. G. Romanov (1997). "On uniqueness of determination of a form of first degree by its integrals along geodesics". *J. Inverse Ill-Posed Probl.* 5.6, 487–490 (1998). MR: 1623603 (cit. on p. 2370).

Viviane Baladi, Mark F. Demers, and Carlangelo Liverani (2018). "Exponential decay of correlations for finite horizon Sinai billiard flows". *Invent. Math.* 211.1, pp. 39–177. MR: 3742756 (cit. on p. 2363).

Viviane Baladi and Masato Tsujii (2007). "Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms". *Ann. Inst. Fourier (Grenoble)* 57.1, pp. 127–154. MR: 2313087 (cit. on p. 2360).

Gang Bao and Hai Zhang (2014). "Sensitivity analysis of an inverse problem for the wave equation with caustics". *J. Amer. Math. Soc.* 27.4, pp. 953–981. MR: 3230816 (cit. on p. 2375).

Michael Blank, Gerhard Keller, and Carlangelo Liverani (2002). "Ruelle–Perron–Frobenius spectrum for Anosov maps". *Nonlinearity* 15.6, pp. 1905–1973. MR: 1938476 (cit. on p. 2360).

Y. Bonthonneau and T. Weich (n.d.). *Pollicott-Ruelle spectrum for manifolds with hyperbolic cusps*. Preprint (cit. on p. 2362).

Jean Bourgain and Semyon Dyatlov (n.d.). "Spectral gaps without the pressure condition". arXiv: 1612.09040 (cit. on p. 2363).

Rufus Bowen and David Ruelle (1975). "The ergodic theory of Axiom A flows". *Invent. Math.* 29.3, pp. 181–202. MR: 0380889.

Dmitri Burago and Sergei Ivanov (2010). "Boundary rigidity and filling volume minimality of metrics close to a flat one". *Ann. of Math. (2)* 171.2, pp. 1183–1211. MR: 2630062 (cit. on pp. 2370, 2374).

Oliver Butterley and Carlangelo Liverani (2007). "Smooth Anosov flows: correlation spectra and stability". *J. Mod. Dyn.* 1.2, pp. 301–322. MR: 2285731 (cit. on pp. 2359, 2360).

C. Croke, A. Fathi, and J. Feldman (1992). "The marked length-spectrum of a surface of nonpositive curvature". *Topology* 31.4, pp. 847–855. MR: 1191384 (cit. on p. 2375).

Christopher Croke (2005). "Boundary and lens rigidity of finite quotients". *Proc. Amer. Math. Soc.* 133.12, pp. 3663–3668. MR: 2163605 (cit. on p. 2374).

– (2014). "Scattering rigidity with trapped geodesics". *Ergodic Theory Dynam. Systems* 34.3, pp. 826–836. MR: 3199795 (cit. on p. 2373).

Christopher B. Croke (1990). "Rigidity for surfaces of nonpositive curvature". *Comment. Math. Helv.* 65.1, pp. 150–169. MR: 1036134 (cit. on pp. 2369, 2370, 2374).

– (1991). "Rigidity and the distance between boundary points". *J. Differential Geom.* 33.2, pp. 445–464. MR: 1094465 (cit. on pp. 2368, 2375).

Christopher B. Croke and Pilar Herreros (2016). "Lens rigidity with trapped geodesics in two dimensions". *Asian J. Math.* 20.1, pp. 47–57. MR: 3460758 (cit. on p. 2373).

Christopher B. Croke and Vladimir A. Sharafutdinov (1998). "Spectral rigidity of a compact negatively curved manifold". *Topology* 37.6, pp. 1265–1273. MR: 1632920 (cit. on p. 2375).

Nurlan S. Dairbekov (2006). "Integral geometry problem for nontrapping manifolds". *Inverse Problems* 22.2, pp. 431–445. MR: 2216407 (cit. on p. 2374).

Nguyen Viet Dang and Gabriel Rivière (n.d.[a]). "Spectral analysis of morse-smale flows ii: resonances and resonant states". arXiv: 1703.08038 (cit. on p. 2367).

– (n.d.[b]). "Spectral analyss of Morse-Smale gradient flows". Preprint (cit. on p. 2362).

Dmitry Dolgopyat (1998). "On decay of correlations in Anosov flows". *Ann. of Math. (2)* 147.2, pp. 357–390. MR: 1626749 (cit. on p. 2363).

Semyon Dyatlov (2015). "Resonance projectors and asymptotics for r-normally hyperbolic trapped sets". *J. Amer. Math. Soc.* 28.2, pp. 311–381. MR: 3300697 (cit. on p. 2364).

Semyon Dyatlov, Frédéric Faure, and Colin Guillarmou (2015). "Power spectrum of the geodesic flow on hyperbolic manifolds". *Anal. PDE* 8.4, pp. 923–1000. MR: 3366007 (cit. on pp. 2364, 2365).

Semyon Dyatlov and Colin Guillarmou (2016). "Pollicott-Ruelle resonances for open systems". *Ann. Henri Poincaré* 17.11, pp. 3089–3146. MR: 3556517 (cit. on pp. 2361, 2367, 2373).

Semyon Dyatlov and Maciej Zworski (n.d.). *Mathematical theory of scattering resonances, book in preparation* (cit. on pp. 2359, 2360, 2367, 2376).

Frédéric Faure (2007). "Prequantum chaos: resonances of the prequantum cat map". *J. Mod. Dyn.* 1.2, pp. 255–285. MR: 2285729 (cit. on p. 2364).

Frederic Faure and Colin Guillarmou (n.d.). "Horocyclic invariance of Ruelle resonant states for contact Anosov flows in dimension 3". arXiv: 1705.07965 (cit. on p. 2366).

Frédéric Faure, Nicolas Roy, and Johannes Sjöstrand (2008). "Semi-classical approach for Anosov diffeomorphisms and Ruelle resonances". *Open Math. J.* 1, pp. 35–81. MR: 2461513 (cit. on p. 2360).

Frédéric Faure and Johannes Sjöstrand (2011). "Upper bound on the density of Ruelle resonances for Anosov flows". *Comm. Math. Phys.* 308.2, pp. 325–364. MR: 2851145 (cit. on pp. 2359, 2360, 2376).

Frédéric Faure and Masato Tsujii (2013). "Band structure of the Ruelle spectrum of contact Anosov flows". *C. R. Math. Acad. Sci. Paris* 351.9-10, pp. 385–391. MR: 3072166 (cit. on p. 2364).

– (2015). "Prequantum transfer operator for symplectic Anosov diffeomorphism". *Astérisque* 375, pp. ix+222. MR: 3461553 (cit. on pp. 2364, 2368).

– (2017). "The semiclassical zeta function for geodesic flows on negatively curved manifolds". *Invent. Math.* 208.3, pp. 851–998. MR: 3648976 (cit. on p. 2364).

P. Giulietti, C. Liverani, and M. Pollicott (2013). "Anosov flows and dynamical zeta functions". *Ann. of Math. (2)* 178.2, pp. 687–773. MR: 3071508 (cit. on p. 2367).

Sébastien Gouëzel and Carlangelo Liverani (2006). "Banach spaces adapted to Anosov systems". *Ergodic Theory Dynam. Systems* 26.1, pp. 189–217. MR: 2201945 (cit. on p. 2360).

Colin Guillarmou (2017a). "Invariant distributions and X-ray transform for Anosov flows". *J. Differential Geom.* 105.2, pp. 177–208. MR: 3606728 (cit. on pp. 2375, 2376).

– (2017b). "Lens rigidity for manifolds with hyperbolic trapped sets". *J. Amer. Math. Soc.* 30.2, pp. 561–599. MR: 3600043 (cit. on pp. 2371, 2372).

Colin Guillarmou, Joachim Hilgert, and Tobias Weich (n.d.[a]). "Classical and quantum resonances for hyperbolic surfaces". To appear in *Math. Annalen.* arXiv: 1605.08801 (cit. on pp. 2365, 2366).

– (n.d.[b]). *First band of Ruelle resonances for compact locally symmetric spaces of rank one*. Preprint (cit. on p. 2365).

Colin Guillarmou and Marco Mazzucchelli (2016). "Marked boundary rigidity for surfaces". *Ergodic Theory and Dynamical Systems*, pp. 1–20 (cit. on p. 2373).

Colin Guillarmou, Marco Mazzucchelli, and Leo Tzou (n.d.). "Boundary and lens rigidity for non-convex manifolds". arXiv: 1711.10059 (cit. on pp. 2361, 2373, 2374).

V. Guillemin and D. Kazhdan (1980a). "Some inverse spectral results for negatively curved 2-manifolds". *Topology* 19.3, pp. 301–312. MR: 579579 (cit. on p. 2375).

Victor Guillemin and David Kazhdan (1980b). "Some inverse spectral results for negatively curved n-manifolds". In: *Geometry of the Laplace operator (Proc. Sympos. Pure Math., Univ. Hawaii, Honolulu, Hawaii, 1979)*. Proc. Sympos. Pure Math., XXXVI. Amer. Math. Soc., Providence, R.I., pp. 153–180. MR: 573432 (cit. on p. 2375).

Charles Hadfield (n.d.). "Ruelle and Quantum Resonances for Open Hyperbolic Manifolds". arXiv: 1708.01200 (cit. on p. 2366).

Sean Holman and Gunther Uhlmann (n.d.). *On the microlocal analysis of the geodesic x-ray transform with conjugate points* (cit. on p. 2375).

S. Hurder and A. Katok (1990). "Differentiability, rigidity and Godbillon-Vey classes for Anosov flows". *Inst. Hautes Études Sci. Publ. Math.* 72, 5–61 (1991). MR: 1087392 (cit. on p. 2366).

Jean-Lin Journé (1986). "On a regularity problem occurring in connection with Anosov diffeomorphisms". *Comm. Math. Phys.* 106.2, pp. 345–351. MR: 855316 (cit. on p. 2376).

Anatole Katok (1988). "Four applications of conformal equivalence to geometry and dynamics". *Ergodic Theory Dynam. Systems* 8*.Charles Conley Memorial Issue, pp. 139–152. MR: 967635 (cit. on p. 2375).

Anatole Katok and Boris Hasselblatt (1995). *Introduction to the modern theory of dynamical systems*. Vol. 54. Encyclopedia of Mathematics and its Applications. With a supplementary chapter by Katok and Leonardo Mendoza. Cambridge University Press, Cambridge, pp. xviii+802. MR: 1326374 (cit. on p. 2358).

Benjamin Küster and Tobias Weich (n.d.). "Quantum-classical correspondence on associated vector bundles over locally symmetric spaces". arXiv: 1710.04625 (cit. on p. 2366).

Carlangelo Liverani (2004). "On contact Anosov flows". *Annals of Math. (2)* 159.3, pp. 1275–1312. MR: 2113022 (cit. on p. 2363).

– (2005). "Fredholm determinants, Anosov maps and Ruelle resonances". *Discrete Contin. Dyn. Syst.* 13.5, pp. 1203–1215. MR: 2166265 (cit. on p. 2360).

R. de la Llave, J. M. Marco, and R. Moriyón (1986). "Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation". *Ann. of Math. (2)* 123.3, pp. 537–611. MR: 840722 (cit. on p. 2376).

René Michel (1981/82). "Sur la rigidité imposée par la longueur des géodésiques". *Invent. Math.* 65.1, pp. 71–83. MR: 636880 (cit. on p. 2369).

François Monard, Plamen Stefanov, and Gunther Uhlmann (2015). "The geodesic ray transform on Riemannian surfaces with conjugate points". *Comm. Math. Phys.* 337.3, pp. 1491–1513. MR: 3339183 (cit. on p. 2375).

Calvin C. Moore (1987). "Exponential decay of correlation coefficients for geodesic flows". In: *Group representations, ergodic theory, operator algebras, and mathematical physics (Berkeley, Calif., 1984)*. Vol. 6. Math. Sci. Res. Inst. Publ. Springer, New York, pp. 163–181. MR: 880376 (cit. on p. 2363).

R. G. Muhometov (1981). "On a problem of reconstructing Riemannian metrics". *Sibirsk. Mat. Zh.* 22.3, pp. 119–135, 237. MR: 621466 (cit. on p. 2370).

Frédéric Naud (2005). "Expanding maps on Cantor sets and analytic continuation of zeta functions". *Ann. Sci. École Norm. Sup. (4)* 38.1, pp. 116–153. MR: 2136484 (cit. on p. 2363).

Stéphane Nonnenmacher and Maciej Zworski (2015). "Decay of correlations for normally hyperbolic trapping". *Invent. Math.* 200.2, pp. 345–438. MR: 3338007 (cit. on p. 2363).

Jean-Pierre Otal (1990a). "Le spectre marqué des longueurs des surfaces à courbure négative". *Ann. of Math. (2)* 131.1, pp. 151–162. MR: 1038361 (cit. on p. 2375).

– (1990b). "Sur les longueurs des géodésiques d'une métrique à courbure négative dans le disque". *Comment. Math. Helv.* 65.2, pp. 334–347. MR: 1057248 (cit. on pp. 2369, 2370, 2374).

Gabriel P. Paternain, Mikko Salo, and Gunther Uhlmann (2013). "Tensor tomography on surfaces". *Invent. Math.* 193.1, pp. 229–247. MR: 3069117 (cit. on p. 2370).

– (2014). "Spectral rigidity and invariant distributions on Anosov surfaces". *J. Differential Geom.* 98.1, pp. 147–181. MR: 3263517 (cit. on pp. 2375, 2376).

L. N. Pestov and V. A. Sharafutdinov (1988). "Integral geometry of tensor fields on a manifold of negative curvature". *Sibirsk. Mat. Zh.* 29.3, pp. 114–130, 221. MR: 953028 (cit. on p. 2370).

Leonid Pestov and Gunther Uhlmann (2005). "Two dimensional compact simple Riemannian manifolds are boundary distance rigid". *Ann. of Math. (2)* 161.2, pp. 1093–1110. MR: 2153407 (cit. on pp. 2370, 2373, 2374).

Marina Ratner (1987). "The rate of mixing for geodesic and horocycle flows". *Ergodic Theory Dynam. Systems* 7.2, pp. 267–288. MR: 896798 (cit. on p. 2363).

S. Smale (1967). "Differentiable dynamical systems". *Bull. Amer. Math. Soc.* 73, pp. 747–817. MR: 0228014 (cit. on p. 2358).

P. Stefanov, G. Uhlmann, and A. Vasy (n.d.[a]). "Inverting the local geodesic X-ray transform on tensors". To appear in *J. Anal. Math.* (cit. on pp. 2370, 2371).

Plamen Stefanov and Gunther Uhlmann (2005). "Boundary rigidity and stability for generic simple metrics". *J. Amer. Math. Soc.* 18.4, pp. 975–1003. MR: 2163868 (cit. on p. 2370).

– (2009). "Local lens rigidity with incomplete data for a class of non-simple Riemannian manifolds". *J. Differential Geom.* 82.2, pp. 383–409. MR: 2520797 (cit. on p. 2374).

– (2012). "The geodesic X-ray transform with fold caustics". *Anal. PDE* 5.2, pp. 219–260. MR: 2970707 (cit. on p. 2375).

Plamen Stefanov, Gunther Uhlmann, and Andras Vasy (n.d.[b]). "Local and global boundary rigidity and the geodesic X-ray transform in the normal gauge". arXiv: 1702.03638 (cit. on pp. 2370, 2371).

Luchezar Stoyanov (2011). "Spectra of Ruelle transfer operators for axiom A flows". *Nonlinearity* 24.4, pp. 1089–1120. MR: 2776112 (cit. on p. 2363).

– (2013). "Pinching conditions, linearization and regularity of axiom A flows". *Discrete Contin. Dyn. Syst.* 33.2, pp. 391–412. MR: 2975118 (cit. on p. 2363).

Masato Tsujii (n.d.). "Exponential mixing for generic volume-preserving Anosov flows in dimension three". arXiv: 1601.00063 (cit. on p. 2363).

– (2010). "Quasi-compactness of transfer operators for contact Anosov flows". *Nonlinearity* 23.7, pp. 1495–1545. MR: 2652469 (cit. on p. 2363).

– (2012). "Contact Anosov flows and the Fourier-Bros-Iagolnitzer transform". *Ergodic Theory Dynam. Systems* 32.6, pp. 2083–2118. MR: 2995886 (cit. on p. 2363).

Gunther Uhlmann and András Vasy (2016). "The inverse problem for the local geodesic ray transform". *Invent. Math.* 205.1, pp. 83–120. MR: 3514959 (cit. on pp. 2370, 2371, 2375).

Marie-France Vignéras (1980). "Variétés riemanniennes isospectrales et non isométriques". *Ann. of Math. (2)* 112.1, pp. 21–32. MR: 584073 (cit. on p. 2369).

Xiaochen Zhou (2011). *The C^∞-jet of Non-Concave Manifolds and Lens Rigidity of Surfaces*. Thesis (Ph.D.)–University of Pennsylvania. ProQuest LLC, Ann Arbor, MI, p. 58. MR: 2995947 (cit. on p. 2374).

Maciej Zworski (2012). *Semiclassical analysis*. Vol. 138. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xii+431. MR: 2952218 (cit. on p. 2360).

Received 2017-12-19.

Colin Guillarmou
CNRS
Université Paris-Sud
Département de Mathématiques
91400 Orsay
France
colin.guillarmou@math.u-psud.fr
cguillar@math.cnrs.fr

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2381–2408)

SMALL SCALES AND SINGULARITY FORMATION IN FLUID DYNAMICS

Alexander Kiselev

Abstract

We review recent advances in understanding singularity and small scales formation in solutions of fluid dynamics equations. The focus is on the Euler and surface quasi-geostrophic (SQG) equations and associated models.

1 Introduction

Fluids are all around us, and attempts to mathematically understand fluid motion go back many centuries. Anyone witnessing a dramatic phenomena like tornado or hurricane or even an everyday river flow or ocean wave breaking can easily imagine the complexity of the task. There has been tremendous accumulation of knowledge in the field, yet it is remarkable that some of the fundamental properties of key equations of fluid mechanics remain poorly understood.

A special role in fluid mechanics is played by the incompressible Euler equation, first formulated in 1755 Euler [1755]. Amazingly, it appears to be the second partial differential equation ever derived (the first one is wave equation derived by D'Alembert 8 years earlier). The incompressible Euler equation describes motion of an inviscid, volume preserving fluid; fluid with such properties is often called "ideal". The Euler equation is a nonlinear and nonlocal system of PDE, with dynamics near a given point depending on the flow field over the entire region filled with fluid. This makes analysis of these equations exceedingly challenging, and the array of mathematical methods applied to their study has been very broad.

The basic purpose of an evolution PDE is solution of Cauchy problem: given initial data, find a solution that can then be used for prediction of the modelled system. This is

This work has been supported by the NSF-DMS award 1712294.
MSC2010: primary 35Q35; secondary 76B03, 35L67.
Keywords: Euler equation, SQG equation, two-dimensional incompressible flow, small scale creation, vorticity gradient growth, singularity formation, hyperbolic flow.

exactly how weather forecasting works, or how new car and airplane shapes are designed. Therefore, one of the first questions one can ask about a PDE is existence and uniqueness of solutions in appropriate functional spaces. The PDE is called globally regular if there exists a unique, sufficiently smooth solution for reasonable classes of initial data. On the other hand, singularity formation - meaning that some quantities associated with fluid motion become infinite - can indicate spontaneous creation of intense fluid motion. Singularities are also important to understand since they may indicate potential breakdown of the model, may lead to loss of uniqueness and predictive power, and are very hard to resolve computationally. More generally, one can ask a related and broader question about creation of small scales in fluids - coherent structures that vary sharply in space and time, and contribute to phenomena such as turbulence (see e.g. Eyink [2008] for further references).

For the Euler equation, the global regularity vs finite time blow up story is very different depending on the dimension. Let us recall that the incompressible Euler equation in a domain $D \subset \mathbb{R}^d$, $d = 2$ or 3 with natural no penetration boundary conditions is given by

$$\partial_t u + (u \cdot \nabla)u = \nabla p, \quad \nabla \cdot u = 0, \quad u \cdot n|_{\partial D} = 0, \tag{1}$$

along with the initial data $u(x,0) = u_0$. Here $u(x,t)$ is the vector field describing fluid velocity, p is pressure, and n is the normal at the boundary ∂D. The equation (1) is just the second Newton's law written for ideal fluid. The difference between dimensions becomes clear if we rewrite the equation in vorticity $\omega = \mathrm{curl} u$:

$$\partial_t \omega + (u \cdot \nabla)\omega = (\omega \cdot \nabla)u, \quad \omega(x,0) = \omega_0(x), \tag{2}$$

along with Biot-Savart law which allows to recover u from ω. For instance, in the case of a smooth domain $D \subset \mathbb{R}^2$ one gets $u = \nabla^\perp(-\Delta_D)^{-1}\omega$ - where $\nabla^\perp = (\partial_{x_2}, -\partial_{x_1})$ and Δ_D is the Dirichlet Laplacian. In the form (2), one can observe that the term $(\omega \cdot \nabla)u$ on the right hand side vanishes in dimension two. The resulting equation conserves the L^∞ norm and in fact any L^p norm of a regular solution, which helps prove global regularity. This result has been known since 1930s works by Wolibner [1933] and Hölder [1933]. We will focus on the 2D Euler equation in Section 3 below. We note that another feature of the Euler equation made obvious by the vorticity representation (2) (in any dimension) is nonlocality: the inverse Laplacian in the Biot-Savart law is a manifestly nonlocal operator, involving integration over the entire domain.

In three dimensions, the "vortex-stretching" term $(\omega \cdot \nabla)u$ is present and can affect the intensity of vorticity. Local well-posedness results in a range of natural spaces are well known; one can consult Majda and Bertozzi [2002] or Marchioro and Pulvirenti [1994] for proofs and further references. However, the global regularity vs finite time singularity formation question remains open. In fact, this problem for the Euler equation is a close relative of the celebrated Clay Institute Millennium problem on the 3D Navier-Stokes equation C. L. Fefferman [2006]. Indeed, the Navier-Stokes equation only differs from (1) by the presence of a regularizing, linear term Δu on the right hand side modeling viscosity (and by different boundary conditions if boundaries are present). The nonlinearity - the principal engine of possible singular growth - is identical in both equations.

A host of numerical experiments sought to discover a scenario for singularity formation in solutions of 3D Euler equation (see e.g. Bhattacharjee, Ng, and Wang [1995], Boratav and R. B. Pelz [1994], E and Shu [1994], Grauer and Sideris [1991], Hou and R. Li [2006], Kerr [1993], Larios, Petersen, Titi, and Wingate [n.d.], Ohkitani and Gibbon [2000], R. Pelz and Gulak [1997], and Pumir and Siggia [1992], or a detailed review by Gibbon [2008] where more references can be found). In the analytic direction, a complete review would be too broad to attempt here. Let us mention the classical work of Beale, Kato, and Majda [1984] on criteria for global regularity, papers Caflisch [1993] and Siegel and Caflisch [2009] on singularities for complex-valued solutions to Euler equations, as well as regularity criteria by Constantin, C. Fefferman, and Majda [1996] and by Hou and collaborators Hou and R. Li [2006], Hou and C. Li [2005], and Deng, Hou, and Yu [2005] which involve more subtle geometric conditions sufficient for regularity. See Constantin [2007] for more history and analytical aspects of this problem.

There have been several recent developments in classical problems on regularity and solution estimates for the fundamental equations of fluid mechanics. First, Hou and Luo produced a new set of careful numerical experiments suggesting finite time singularity formation for solutions of 3D Euler equation Luo and Hou [2014]. The scenario of Hou and Luo is axi-symmetric. Very fast vorticity growth is observed at a ring of hyperbolic stagnation points of the flow located on the boundary of a cylinder. None of the available regularity criteria such as Constantin, C. Fefferman, and Majda [1996], Hou and R. Li [2006], Deng, Hou, and Yu [2005], and Hou and C. Li [2005] seem to apply to scenario's geometry. The scenario has a close analog for the 2D inviscid Boussinesq system, for which the question of global regularity is also open; it is listed as one of the "eleven great problems of hydrodynamics" by V. I. Yudovich [2003]. More details on the scenario will be provided in Section 2.

The hyperbolic stagnation point on the boundary scenario also leads to interesting phenomena in solutions of the 2D Euler equation. The best known upper bound on the growth of the gradient of vorticity, as well as higher order Sobolev norms, is double exponential in time:

$$\|\nabla\omega(\cdot,t)\|_{L^\infty} \leq (1+\|\nabla\omega_0\|_{L^\infty})^{\exp(C\|\omega_0\|_{L^\infty}t)}. \tag{3}$$

This result has appeared explicitly in V. I. Yudovich [1962], though related bounds can be traced back to Wolibner [1933]. The question whether such upper bounds are sharp has been open for a long time. Kiselev and Šverák [2014] provided an example of smooth initial data in the disk such that the corresponding solution exhibits double exponential growth in the gradient of vorticity for all times, establishing qualitative sharpness of (3). The construction is based on the hyperbolic point at the boundary scenario, and will be described in more detail in Section 3.

Further attempts to rigorously understand the Hou-Luo scenario involved construction of 1D and 2D models retaining some of the analytical structure of the original problem. We will discuss some of these models in Section 4.

The surface quasi-geostrophic (SQG) equation is similar to the 2D Euler equation in vorticity form, but is more singular:

$$\partial_t \omega + (u \cdot \nabla)\omega = 0, \;\; u = \nabla^{\perp}(-\Delta)^{-1+\alpha}\omega, \;\; \alpha = 1/2, \;\; \omega(x,0) = \omega_0(x). \tag{4}$$

The value $\alpha = 0$ corresponds to the 2D Euler equation, while $0 < \alpha < \frac{1}{2}$ is called the modified SQG range. The SQG and modified SQG equations come from atmospheric science. They model evolution of temperature near the surface of a planet and can be derived by formal asymptotic analysis from a larger system of rotating 3D Navier-Stokes equations coupled with temperature equation through buoyancy force Held, Pierrehumbert, Garner, and Swanson [1995], Majda [2003], Pedlosky [1987], and Pierrehumbert, Held, and Swanson [1994]. In mathematical literature, the SQG equation was first considered by Constantin, Majda, and Tabak [1994], where a parallel between the structure of the SQG equation and the 3D Euler equation was drawn. The SQG and modified SQG equations are perhaps simplest looking equations of fluid mechanics for which the question of global regularity vs finite time blow up remains open. The equation (4) can be considered with smooth initial data, but another important class of initial data is patches, where $\theta_0(x)$ equals linear combination of characteristic functions of some disjoint domains $\Omega_j(0)$. The resulting evolution yields time dependent regions $\Omega_j(t)$. The regularity question in this context addresses the regularity class of the boundary $\partial\Omega_j(t)$ and lack of contact between different components. Existence and uniqueness of patch solution for 2D Euler equation follows from Yudovich theory Judovič [1963], Majda and Bertozzi [2002], and Marchioro and Pulvirenti [1994]. The global regularity question has been settled affirmatively by Chemin [1993] (Bertozzi and Constantin [1993] provided a different proof). For the SQG and modified SQG equations patch dynamics is harder to set up. Local well-posedness has been shown by Rodrigo in C^∞ class Rodrigo [2005] and by Gancedo in Sobolev spaces Gancedo [2008] in the whole plane setting. Numerical simulations by Córdoba, Fontelos, Mancho, and Rodrigo [2005] suggest that finite time singularities – in particular formation of corners and different components touching each other – is possible, but rigorous understanding of this phenomena remained missing. In Kiselev, Ryzhik, Yao, and Zlatoš [2016], Kiselev, Yao, and Zlatoš [2017], we considered modified SQG and 2D Euler patches in half-plane, with the no penetration boundary conditions. The initial patches are regular and do not touch each other but may touch the boundary. We proved a kind

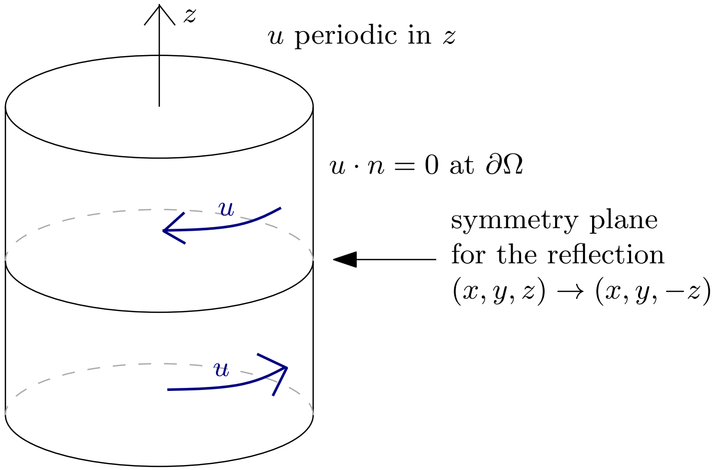


Figure 1: The initial data for Hou-Luo scenario

of phase transition in this setting: the 2D Euler patches stay globally regular, while for a range of small $\alpha > 0$ some initial data lead to blow up in finite time. The blow up scenario again involves a stagnation hyperbolic point of the flow on the boundary. This result will be described in more detail in Section 5.

2 The 3D Euler equation and the 2D Boussinesq system: the hyperbolic scenario

In Luo and Hou [2014] the authors study 3D axi-symmetric solutions of incompressible Euler equation with roughly the initial configuration shown on Figure 1: only swirl u^ϕ is initially non-zero, and it is odd and periodic in z variable.

One of the standard forms of the axi-symmetric Euler equations in the usual cylindrical coordinates (r, ϕ, z) is

$$\partial_t \left(\frac{\omega^\phi}{r}\right) + u^r \partial_r \left(\frac{\omega^\phi}{r}\right) + u^z \partial_z \left(\frac{\omega^\phi}{r}\right) = \partial_z \left(\frac{(ru^\phi)^2}{r^4}\right) \tag{5a}$$

$$\partial_t (ru^\phi) + u^r \partial_r (ru^\phi) + u^z \partial_z (ru^\phi) = 0\,, \tag{5b}$$

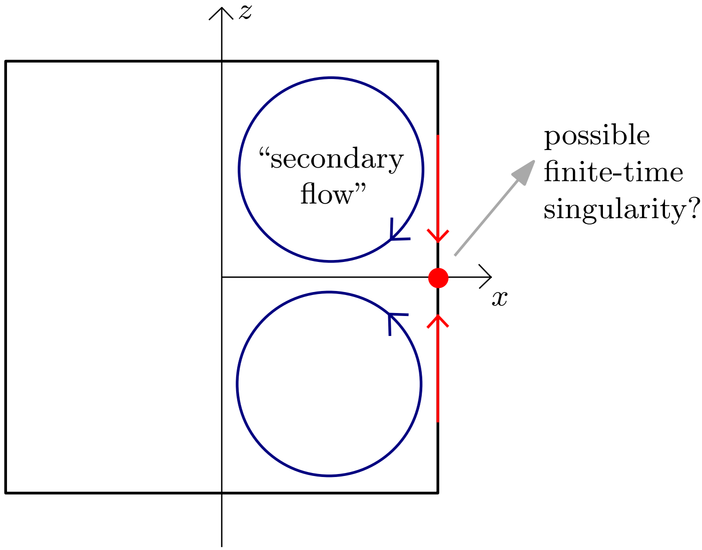


Figure 2: The secondary flows in fixed ϕ section

with the understanding that u^r, u^z are given from ω^ϕ via the Biot-Savart law which in the setting of Hou-Luo scenario takes form

$$u^r = -\frac{\partial_z \psi}{r}, \quad u^z = \frac{\partial_r \psi}{r}, \quad L\psi = \frac{\omega^\phi}{r}, \quad L\psi = -\frac{1}{r}\partial_r \left(\frac{1}{r}\partial_r \psi\right) - \frac{1}{r^2}\partial^2_{zz}\psi.$$

From (5), it is clear that the swirl will spontaneously generate toroidal rolls corresponding to non-zero ω^ϕ. These are the so-called "secondary flows", Prandtl [1952]; its effect on river flows was studied by Einstein [1926]. Thus the initial condition leads to the (schematic) picture in the xz–plane shown on Figure 2, in which we also indicate the point where a conceivable finite-time singularity (or at least an extremely strong growth of vorticity) is observed numerically. In the three-dimensional picture, the points with very fast growth form a ring on the boundary of the cylinder.

A somewhat similar scenario can be considered for the 2D inviscid Boussinesq system in a half-space $\mathbb{R}^+ = \{(x, y) \in \mathbb{R} \times (0, \infty)\}$ (or in a flat half-cylinder $\mathbf{S}^1 \times (0, \infty)$), which we will write in the vorticity form:

(6a) $$\partial_t \omega + u_1 \partial_x \omega + u_2 \partial_y \omega = \partial_x \theta$$

(6b) $$\partial_t \theta + u_1 \partial_x \theta + u_2 \partial_y \theta = 0\,.$$

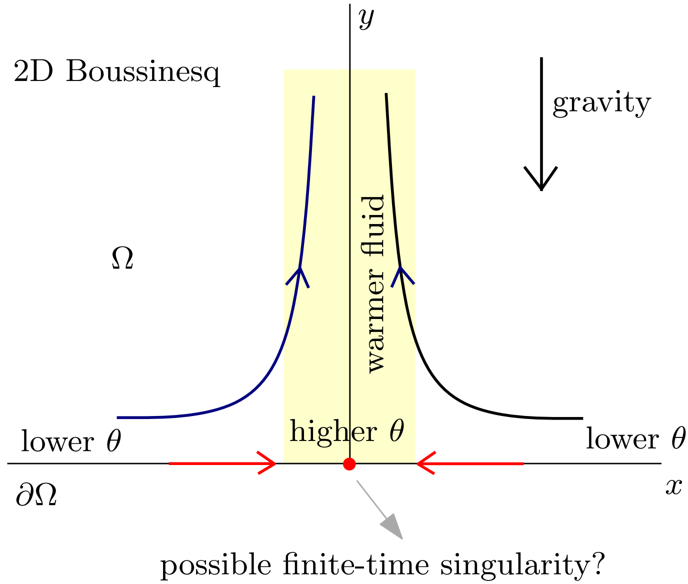


Figure 3: The 2D Boussineq singularity scenario

Here $u = (u_1, u_2)$ is obtained from ω by the usual Biot-Savart law $u = \nabla^{\perp}(-\Delta)^{-1}\omega$, with appropriate boundary conditions on Δ, and θ represents the fluid temperature or density.

It is well-known (see e.g. Majda and Bertozzi [2002]) that this system has properties similar to the 3D axi-symmetric Euler (5), at least away from the rotation axis. Indeed, comparing (5) with (6), we see that θ essentially plays the role of the square of the swirl component ru^{ϕ} of the velocity field u, and ω replaces ω^{ϕ}/r. The real difference between the two systems only emerges near the axis of rotation, where the factors of r can conceivably change the nature of dynamics. For the purpose of comparison with the axi-symmetric flow, the last picture should be rotated by $\pi/2$, after which it resembles the picture relevant for (6), see Figure 3.

In both the 3D axi-symmetric Euler case and in the 2D Boussinesq system case the best chance for possible singularity formation seems to be at the points of symmetry at the boundary, which numerical simulations suggest are fixed hyperbolic points of the flow.

3 The 2D Euler equation

A reasonable first step to understand the hyperbolic stagnation point on the boundary blow up scenario is to consider the case of constant density in the Boussinesq system first. Of course, this reduces the system to the 2D Euler equation:

$$\partial_t \omega + (u \cdot \nabla)\omega = 0, \;\; u = \nabla^{\perp}(-\Delta_D)^{-1}\omega, \;\; \omega(x,0) = \omega_0(x). \tag{7}$$

Here Δ_D stands for Dirichlet Laplacian; such choice of the boundary condition corresponds to no-penetration property $u \cdot n|_{\partial D} = 0$. Of course, for the 2D Euler equation solutions are globally regular. Let us state this result, going back to 1930s Wolibner [1933] and Hölder [1933].

Theorem 3.1. *Let $D \subset \mathbb{R}^2$ be a compact domain with smooth boundary, and $\omega_0(x) \in C^1(D)$Then there exists a unique smooth solution $\omega(x,t)$ of the equation (7) corrsponding to the initial data ω_0, which moreover satisfies*

$$1 + \log\left(1 + \frac{\|\nabla\omega(x,t)\|_{L^\infty}}{\|\omega_0\|_{L^\infty}}\right) \le \left(1 + \log\left(1 + \frac{\|\nabla\omega_0\|_{L^\infty}}{\|\omega_0\|_{L^\infty}}\right)\right)\exp(C\|\omega_0\|_{L^\infty} t) \tag{8}$$

for some constant C which may depend only on the domain D.

A key ingredient of the proof is the Kato inequality Kato [1986]: for every $1 > \alpha > 0$, we have

$$\|\nabla u(x,t)\|_{L^\infty} \le C(\alpha, D)\|\omega_0\|_{L^\infty}\left(1 + \log\frac{\|\omega(x,t)\|_{C^\alpha}}{\|\omega_0\|_{L^\infty}}\right). \tag{9}$$

Note that the matrix ∇u consists of double Riesz transforms $\partial^2_{ij}(-\Delta_D)^{-1}\omega$. Riesz transforms are well known to be bounded on L^p, $1 < p < \infty$ (see e.g. Stein [1970]), but the bound fails at the endpoints and we have to pay a logarithm of the higher order norm to obtain a correct bound. It is exactly the extra log in (9) that leads to the double exponential upper bound as opposed to the single one (see e.g. Kiselev and Šverák [2014] for more details).

The question of whether such upper bounds are sharp has been open for a long time. Judovič [1974] and V. I. Yudovich [2000] provided an example showing infinite growth of the vorticity gradient at the boundary of the domain, by constructing an appropriate Lyapunov-type functional. These results were further improved and generalized in Morgulis, Shnirelman, and V. Yudovich [2008], leading to description of a broad class of flows with infinite growth in their vorticity gradient. Nadirashvili [1991] proved a more quantitative linear in time lower bound for a “winding” flow in an annulus. A variant of the example due to Bahouri and Chemin [1994] provides singular stationary solution of the 2D

Euler equation defined on $\mathbb{T}^2$ with fluid velocity which is just log-Lipschitz in spacial variables. Namely, if we set $\mathbb{T}^2 = [-\pi, \pi) \times [-\pi, \pi)$, the solution is equal to -1 in the first quadrant $[0, \pi) \times [0, \pi)$ and is odd with respect to both coordinate axes. Note that the solution is just L^∞ but existence and uniqueness of solutions in this class is provided essentially by Yudovich theory Judovič [1963]. The origin is a fixed hyperbolic point of the fluid velocity, with x_1 being the contracting direction, and the velocity satisfies $u_1(x_1, 0) = \frac{4}{\pi} x_1 \log x_1 + O(x_1)$ for small x_1. The trajectory starting at a point $(x_1, 0)$ on a horizontal separatrix will therefore converge to the origin at a double exponential rate in time. If a smooth passive scalar ψ initially supported away from the origin is advected by a flow generated by singular cross, $\partial_{x_1}\psi$ will grow at a double exponential rate in time if $\psi(x_1, 0)$ is not identically zero. Of course, derivative growth does not make sense for the singular cross solution itself since it is stationary and already singular. But this solution shows a blueprint of how double exponential growth can be conceivably generated in smooth solution: it needs to approach the discontinuous configuration similar to the singular cross, while at the same time the solution should be nonzero on the contracting direction. This turns out to be hard to implement, especially without boundary.

In recent years, there has been a series of works by Denisov on this problem. In Denisov [2009], he constructed an example with superlinear growth in vorticity gradient of the solution in the periodic case. In Denisov [2015], he showed that for any time T, one can arrange smooth initial data so that the corresponding solution will experience double exponential burst of growth over $[0, T]$. The example is based on smoothing out Bahouri-Chemin example, abandoning odd symmetry to put a ripple on a separatrix, and controlling the resulting solution over finite time interval.

In Kiselev and Šverák [2014], we proved

Theorem 3.2. *Consider two-dimensional Euler equation on a unit disk D. There exists a smooth initial data ω_0 with $\|\nabla\omega_0\|_{L^\infty}/\|\omega_0\|_{L^\infty} > 1$ such that the corresponding solution $\omega(x,t)$ satisfies*

$$\frac{\|\nabla\omega(x,t)\|_{L^\infty}}{\|\omega_0\|_{L^\infty}} \geq \left(\frac{\|\nabla\omega_0\|_{L^\infty}}{\|\omega_0\|_{L^\infty}}\right)^{c\exp(c\|\omega_0\|_{L^\infty} t)} \tag{10}$$

for some $c > 0$ and for all $t \geq 0$.

The theorem shows that double exponential growth in the gradient of vorticity can actually happen for all times, so the double exponential upper bound is sharp. As in Hou-Luo blow up scenario, growth happens near a hyperbolic fixed point of the flow at the boundary. The result has been generalized to the case of any compact sufficiently regular domain with symmetry axis by Xu [2016]. The question of whether double exponential growth can happen in the bulk of the fluid remains open; Zlatoš [2015] has improved the techniques behind Theorem 3.2 to construct examples of smooth solutions with exponential growth of $\|\nabla^2\omega\|_{L^\infty}$ in periodic setting. The question of whether double exponential growth is at all possible in the bulk of the fluid remains wide open.

A key step in the proof is understanding the structure of fluid velocity near the hyperbolic point. Let

$$D^+ = \{x \in D \,|x_1 \geq 0.\}$$

We will choose the initial data that is odd in x_1, and $-1 \leq \omega_0(x) < 0$ for $x \in D^+$. Let us set the origin of our coordinate system at the bottom of the disc, where interesting things will be happening. Given the symmetry of ω, we have

$$u(x,t) = -\nabla^\perp \int_D G_D(x,y)\omega(y,t)\,dy \tag{11}$$

$$= -\frac{1}{2\pi}\nabla^\perp \int_{D^+} \log\left(\frac{|x-y||\tilde{x}-\bar{y}|}{|x-\bar{y}||\tilde{x}-y|}\right)\omega(y,t)\,dy, \tag{12}$$

where G_D is the Green's function of Dirichlet Laplacian and $\tilde{x} = (-x_1, x_2)$. For each point $(x_1, x_2) \in D^+$, let us introduce the region

$$Q(x_1,x_2) = \{(y_1,y_2) \in D^+ : \;\; x_1 \leq y_1, \;\; x_2 \leq y_2\},$$

and set

$$\Omega(x_1,x_2,t) = -\frac{4}{\pi}\int_{Q(x_1,x_2)} \frac{y_1y_2}{|y|^4}\omega(y,t)\,dy_1dy_2. \tag{13}$$

Finally, for any $0 < \gamma < \pi/2$, let ϕ be the usual polar angle coordinate of point x, and denote

$$D_1^\gamma = \{x \in D^+ \,|0 \leq \phi \leq \pi/2-\gamma\}, \;\; D_2^\gamma = \{x \in D^+ \,|\gamma \leq \phi \leq \pi/2\}.$$

The following Lemma is crucial for the proof of Theorem 3.2.

Lemma 3.3. *Suppose that ω_0 is odd with respect to x_1. Fix a small $\gamma > 0$. There exists $\delta > 0$ so that for all $x \in D_1^\gamma$ such that $|x| \leq \delta$ we have*

$$u_1(x_1,x_2,t) = -x_1\Omega(x_1,x_2,t) + x_1B_1(x_1,x_2,t), \tag{14}$$

where $|B_1(x_1,x_2,t)| \leq C(\gamma)\|\omega_0\|_{L^\infty}$.

Similarly, for all $x \in D_2^\gamma$ such that $|x| \leq \delta$ we have

$$u_2(x_1,x_2,t) = x_2\Omega(x_1,x_2,t) + x_2B_2(x_1,x_2,t), \tag{15}$$

where $|B_2(x_1,x_2,t)| \leq C(\gamma)\|\omega_0\|_{L^\infty}$.

Proof of the lemma is based on analysis of (11), full details can be found in Kiselev and Šverák [2014]. The terms on the right hand sides of (14), (15) involving Ω can be thought of as main terms in certain regimes. Indeed, observe that as support of the set where $\omega(x,t) \geq c > 0$ approaches the origin, the size of Ω given by (13) may grow as a logarithm of this distance. Thus Lemma 3.3 provides a sort of quantitative version of Bahouri-Chemin log-Lipschitz singular flow for smooth setting. In the regime where the Ω terms dominate, the flow trajectories described by (14), (15) are close to precise hyperbolas. Another interesting feature of the formulas (14), (15) is a hidden comparison principle: the influence region $Q(x_1,x_2)$ tends to be larger for points closer to the origin. The comparison principle is not precise, but it turns out to be true up to Lipschitz errors. This feature is key in the construction of the example with double exponential growth.

Let us now sketch the construction of such example. Fix small $\gamma > 0$, and take $\epsilon > 0$ smaller than the corresponding value of δ from Lemma 3.3. We also take ϵ sufficiently small so that $\log \epsilon^{-1}$ is much larger than the constant $C(\gamma)$ from Lemma 3.3. Take the initial data ω_0 such that $\omega_0(x_1,x_2) = -1$ if $x_1 \geq \epsilon^{10}$, odd with respect to x_1, and satisfying $-1 \leq \omega_0(x) \leq 0$ for $x \in D^+$. This leaves certain ambiguity in how we define ω_0 in the region close to x_2 axis. As we will see, it does not matter for the construction exactly how ω_0 is defined there.

The first observation is that with such choice of the initial data,

$$\Omega(x,t) \geq C \log \epsilon^{-1} \tag{16}$$

for every $x \in D^+$ with $|x| \leq \epsilon$. Indeed, due to incompressibility of the flow the measure of the points $x \in D^+$ where $\omega(x,t) > -1$ does not exceed 2ϵ for any time t. It is then straightforward to show that even if all these points are pushed by the dynamics into the region where the size of the kernel in (13) is maximal, the estimate (16) still holds.

Next, for $0 < x_1' < x_1'' < 1$ we denote

$$\mathcal{O}(x_1', x_1'') = \left\{(x_1,x_2) \in D^+\,,\ x_1' \leq x_1 \leq x_1''\,,\ x_2 \leq x_1\right\}\,. \tag{17}$$

We would like to analyze the evolution in time of the region $\mathcal{O}_{\epsilon^{10}\epsilon}$. For this purpose, for $0 < x_1 < 1$ we let

$$\underline{u}_1(x_1,t) \quad = \quad \min_{(x_1,x_2)\in D^+,\, x_2 \leq x_1} u_1(x_1,x_2,t) \tag{18}$$

and

$$\overline{u}_1(x_1,t) \quad = \quad \max_{(x_1,x_2)\in D^+,\, x_2 \leq x_1} u_1(x_1,x_2,t)\,. \tag{19}$$

Define $a(t)$ by

$$a' = \overline{u}_1(a,t)\,, \quad a(0) = \epsilon^{10} \tag{20}$$

and $b(t)$ by

$$b' = \underline{u}_1(b,t), \quad b(0) = \epsilon. \tag{21}$$

Let

$$\mathcal{O}_t = \mathcal{O}(a(t), b(t)). \tag{22}$$

We claim that $\omega(x,t) = -1$ for every $x \in \mathcal{O}_t$, and every $t \geq 0$. Indeed, if this is not the case, then there must exist some time s and a trajectory $\Phi_s(y)$ such that $y \notin \mathcal{O}_0$, and $\Phi_s(y) \in \partial\mathcal{O}_s$ for the first time. But the trajectory $\Phi_s(y)$ cannot enter $\mathcal{O}_s$ through the boundary of D due to the no-penetration boundary condition. It is also not hard to see it cannot enter at the $x_1 = a(s)$ or $x_1 = b(s)$ pieces of $\partial\mathcal{O}_s$ due to the definition of $a(t), b(t)$ and $\overline{u}$, $\underline{u}$. This leaves the diagonal $x_1 = x_2$. However, due to Lemma 3.3 and the estimate (16), we have that

$$\frac{C\log\epsilon^{-1} - C(\gamma)}{C\log\epsilon^{-1} + C(\gamma)} \leq \frac{-u_1(x_1,x_1,t)}{u_2(x_1,x_1,t)} \leq \frac{C\log\epsilon^{-1} + C(\gamma)}{C\log\epsilon^{-1} - C(\gamma)} \tag{23}$$

for every $t \geq 0$ and $x_1 \leq \epsilon$. Due to choice of ϵ, we have that the ratio $-u_1/u_2$ on the diagonal part of the boundary of $\mathcal{O}_s$ is close to 1. Thus the vector field u points outside of the region $\mathcal{O}_s$ on the diagonal part of the boundary at all times and so the trajectory cannot enter through the diagonal either.

Next, we are going to estimate how quickly $a(t)$ approaches the origin. The constant C below depends only on γ and may change from line to line. By Lemma 3.3, we have

$$\underline{u}_1(b(t),t) \geq -b(t)\,\Omega(b(t), x_2(t)) - Cb(t),$$

for some $x_2(t) \leq b(t)$, $(x_2(t), b(t)) \in D^+$ as $\|\omega(x,t)\|_{L^\infty} \leq 1$ by our choice of the initial datum ω_0. A straightforward calculation shows that

$$\Omega(b(t), x_2(t)) \leq \Omega(b(t), b(t)) + C.$$

Thus we get

$$\underline{u}_1(b(t),t) \geq -b(t)\,\Omega(b(t), b(t)) - Cb(t). \tag{24}$$

Similarly,

$$\overline{u}_1(a(t),t) \leq -a(t)\,\Omega(a(t), \tilde{x}_2(t)) \leq -a(t)\,\Omega(a(t), 0) + Ca(t).$$

These estimates establish a form of comparison principle, up to Lipschitz errors, of the fluid velocities of the front and back of the region $\mathcal{O}_{a(t),b(t)}$.

Now observe that

$$\Omega(a(t),0) \geq -\frac{4}{\pi}\int_{\mathcal{O}_t} \frac{y_1 y_2}{|y|^4}\omega(y,t)\,dy_1 dy_2 + \Omega(b(t),b(t)). \tag{25}$$

Since $\omega(y,t) = -1$ on $\mathcal{O}_t$, a direct estimate shows that the integral in (25) is bounded from below by $\kappa(-\log a(t) + \log b(t))$ for some $\kappa > 0$. Applying the above estimates to evolution of $a(t)$ and $b(t)$ we obtain

$$\frac{d}{dt}(\log a(t) - \log b(t)) \leq \kappa\,(\log a(t) - \log b(t)) + C. \tag{26}$$

Applying Gronwall lemma and choosing ϵ small enough leads to $\log a(t) \leq \exp(\kappa t)\log\epsilon$. To arrive at (10), it remains to note that we can arrange $\|\nabla\omega_0\|_{L^\infty} \lesssim \epsilon^{-10}$.

4 The one-dimensional models

One-dimensional models in fluid mechanics have a long history. We briefly review some of the results most relevant to our narrative. In the context of modeling finite time blow and global regularity issues, Constantin, Lax, and Majda [1985] considered the model

$$\partial_t\omega = \omega H\omega, \;\; \omega(x,0) = \omega_0(x), \tag{27}$$

where H is the Hilbert transform, $H\omega(x,t) = \frac{1}{\pi}P.V.\int_{\mathbb{R}} \frac{\omega(y,t)}{x-y}\,dy$. The equation (27) is designed to model the vortex stretching term on the right hand side of (2); the advection term is omitted. Surprisingly, the model (27) is explicitly solvable due to special properties of Hilbert transform. Finite time blow up happens for a broad class of initial data - specifically, near the points where $\omega_0(x)$ vanishes and the real part of $H\omega_0$ has the right sign.

A more general model has been proposed by De Gregorio [1990], De Gregorio [1996]:

$$\partial_t\omega + u\partial_x\omega = \omega H\omega, \;\; u_x = H\omega, \;\; \omega(x,0) = \omega_0(x). \tag{28}$$

This model includes the advection term. Amazingly, the question of whether the solutions to (28) are globally regular or can blow up in finite time is currently open. Numerical simulations appear to suggest global regularity Okamoto, Sakajo, and Wunsch [2008], but the mechanism for it is not well understood. Recently, global regularity near a manifold of equilibria as well as other interesting features of the solutions of (28) have been shown in Jia, Stewart, and Sverak [n.d.]. Variants of (28) and other related models appear in for example Bauer, Kolev, and Preston [2016], Castro and Córdoba [2010], Elgindi and Jeong [n.d.(c)], Escher and Kolev [2014], and Wunsch [2011], where further references can be found.

Already in Luo and Hou [2014], Hou and Luo proposed a simplified one-dimensional model specifically designed to gain insight into the singularity formation process in the scenario described in Section 2. This model is given by

$$\begin{aligned} &\partial_t \omega + u \partial_x \omega = \partial_x \theta, \\ &\partial_t \theta + u \partial_x \theta = 0, \;\; u_x = H\omega. \end{aligned} \tag{29}$$

Here as above H is the Hilbert transform, and the setting can be either periodic or the entire axis with some decay of the initial data. The model (29) can be thought of as an effective equation on the $x_2 = 0$ axis in the Boussinesq case (see (6) and Figure 3) or on the boundary of the cylinder in the 3D axi-symmetric Euler case. The model can be derived from the full equations under certain boundary layer assumption: that $\omega(x,t)$ is concentrated in a boundary layer of width a near $x_2 = 0$ axis and is independent of x_2, that is $\omega(x_1, x_2, t) = \omega(x_1, t)\chi_{[0,a]}(x_2)$. Such assumption is necessary to close the equation and reduces the half-plane Biot-Savart law to $u_x = H\omega$ in the main order; the parameter a enters into the additional term that is non-singular and is dropped from (29). See Luo and Hou [ibid.], Choi, Hou, Kiselev, Luo, Sverak, and Yao [2017] for more details. We will call the system (29) the HL model.

The HL model is still fully nonlocal. A further simplification was proposed in Choi, Kiselev, and Yao [2015], where the Biot-Savart law has been replaced with

$$u(x,t) = -x \int_x^1 \frac{\omega(y,t)}{y}\, dy. \tag{30}$$

Here the most natural setting is on an interval $[0,1]$ with smooth initial data supported away from the endpoints. The law (30) is motivated by the velocity representation in Lemma 3.3 above, as it is the simplest one dimensional analog of such representation. This law is "almost local" - if one divides u by x and differentiates, one gets local expression. We will call the model (30) the CKY model.

For both HL and CKY models, local well-posedness in a reasonable family of spaces (such as sufficiently regular Sobolev spaces) is not difficult to obtain. In Choi, Kiselev, and Yao [ibid.], finite time blow up has been proved for the CKY model. The proof used analysis of the trajectories and of the nonlinear feedback loop generated by the forcing term $\partial_x \theta$. We will sketch a very similar argument below. The proof does not provide a detailed blow picture. In a later work Hou and Liu [2015], more precise picture of blow up was established with aid of computer assisted proof. It shows self-similar behavior near the origin properly matched with the outside region. For the original HL model, finite time blow up has been proved in Choi, Hou, Kiselev, Luo, Sverak, and Yao [2017]. For the model including the additional term obtained from the boundary layer assumption into Biot-Savart law, finite time blow up has been proved in Do, Kiselev, and Xu [n.d.]. We now sketch a variant of the blow up proof for the HL model (29).

Let us consider an HL model on $[0, L]$ with periodic boundary conditions. In this setting, using the expression for periodic Hilbert transform, the Biot-Savart law becomes

$$u_x(x) = H\omega(x) = \frac{1}{L} P.V. \int_0^L \omega(y) \cot[\mu(x-y)]\, dy,$$

where $\mu = \pi/L$. Integration leads to

$$u(x) = \frac{1}{\pi} \int_0^L \omega(y) \log|\sin[\mu(x-y)]|\, dy. \tag{31}$$

The initial data will be chosen as follows: ω_0 is odd, which together with periodicity implies that it is also odd with respect to $x = L/2$, and satisfies $\omega_0(x) \geq 0$ if $x \in [0, L/2]$. The initial density θ_0 is even with respect to both 0 and $L/2$, and satisfies $\theta_0' \geq 0$ for $x \in [0, L/2]$. The solution, while it exists, will satisfy the same properties. The symmetry assumptions on ω lead to the following version of the Biot-Savart law, which can be verified by direct computation.

Lemma 4.1. *Let ω be periodic with period L and odd at $x = 0$ and let u be defined by (31). Then for any $x \in [0, \frac{1}{2}L]$,*

$$u(x)\cot(\mu x) = -\frac{1}{\pi} \int_0^{L/2} K(x,y)\omega(y)\cot(\mu y)\, dy, \tag{32}$$

where

$$K(x,y) = s \log\left|\frac{s+1}{s-1}\right| \qquad \text{with} \qquad s = s(x,y) = \frac{\tan(\mu y)}{\tan(\mu x)}. \tag{33}$$

Furthermore, the kernel $K(x,y)$ has the following properties:

1. *$K(x,y) \geq 0$ for all $x,\ y \in (0, \frac{1}{2}L)$ with $x \neq y$;*
2. *$K(x,y) \geq 2$ and $K_x(x,y) \geq 0$ for all $0 < x < y < \frac{1}{2}L$;*

The key observation is a certain positivity property that will help us control the behavior of trajectories Choi, Hou, Kiselev, Luo, Sverak, and Yao [2017].

Lemma 4.2. *Let the assumptions in Lemma 4.1 be satisfied and assume in addition that $\omega \geq 0$ on $[0, \frac{1}{2}L]$. Then for any $a \in [0, \frac{1}{2}L]$,*

$$\int_a^{L/2} \omega(x)\big[u(x)\cot(\mu x)\big]_x\, dx \geq 0. \tag{34}$$

With these two lemmas, the rest of the proof proceeds as follows. Towards a contradiction, let us assume that there exists a global solution (θ, ω) to (29) with the initial data as described in the beginning of this section and denote $A := \theta_0(\frac{1}{2}L) > 0$. Since θ_0 is assumed to be increasing on $[0, \frac{1}{2}L]$, we can choose a decreasing sequence x_n in $(0, \frac{1}{2}L)$ with $n \geq 0$ such that $\theta_0(x_n) = [2^{-1} + 2^{-(n+2)}]A$. Note that $x_0 < \frac{1}{2}L$ since $\theta_0(x_0) < \theta_0(\frac{1}{2}L)$.

For x_n defined as above, let $\Phi_n(t)$ denote the characteristics of (29) originating from x_n, that is, let $\frac{d}{dt}\Phi_n(t) = u(\Phi_n(t), t)$ with $\Phi_n(0) = x_n$. Lemma 4.1 then implies the following estimate on the evolution of Φ_n:

$$\frac{d}{dt}\Phi_n(t) = u(\Phi_n(t), t) \leq -\frac{2}{\pi}\tan(\mu\Phi_n(t))\int_{\Phi_n(t)}^{L/2} \omega(y,t)\cot(\mu y)\,dy \tag{35}$$

$$\leq -\frac{2\mu}{\pi}\,\Phi_n(t)\Omega_n(t), \tag{36}$$

where for simplicity we have written

$$\Omega_n(t) := \int_{\Phi_n(t)}^{L/2} \omega(y,t)\cot(\mu y)\,dy. \tag{37}$$

Introducing the new variable $\psi_n(t) := -\log \Phi_n(t)$, we may write (36) as

$$\frac{d}{dt}\psi_n(t) \geq \frac{2\mu}{\pi}\,\Omega_n(t). \tag{38}$$

Then for each $n \geq 1$, we have

$$\begin{aligned}
\frac{d}{dt}\Omega_n(t) &= \int_{\Phi_n(t)}^{L/2} \omega(y,t)\big[u(y,t)\cot(\mu y)\big]_y\,dy + \int_{\Phi_n(t)}^{L/2} \theta_y(y,t)\cot(\mu y)\,dy \\
&\geq \int_{\Phi_n(t)}^{\Phi_{n-1}(t)} \theta_y(y,t)\cot(\mu y)\,dy \\
&\geq \cot(\mu\Phi_{n-1}(t))\big[\theta_0(x_{n-1}) - \theta_0(x_n)\big] = 2^{-(n+2)}A\cot(\mu\Phi_{n-1}(t)),
\end{aligned}$$

where in the second step we have used Lemma 4.2 and the fact that $\theta_x \geq 0$ on $[0, \frac{1}{2}L]$. To find a lower bound for the right hand side, note that for any fixed $z \in (0, \frac{1}{2}\pi)$, there exists some constant $c > 0$ depending only on z such that $\cot(x) > cx^{-1}$ for any $x \in (0, z]$. In our situation, we have $\mu\Phi_{n-1}(t) \leq \mu\Phi_0(t) \leq \mu x_0 < \frac{1}{2}\pi$, and as a result there exists some constant $c_0 > 0$ depending only on μ and x_0 such that $\cot(\mu\Phi_{n-1}(t)) \geq c_0[\Phi_{n-1}(t)]^{-1}$. This leads to the estimate

$$\frac{d}{dt}\Omega_n(t) \geq 2^{-(n+2)}c_0Ae^{\psi_{n-1}(t)}. \tag{39}$$

Once we have (38), (39), the proof of finite time blow is fairly straightforward. Details of a similar argument can be found in Choi, Kiselev, and Yao [2015]. We can choose A large enough to show inductively that $\psi_n(t_n) \geq bn + a$ for some suitably chosen $a \in \mathbb{R}$, $b > 0$ and an increasing sequence $t_n \to T < \infty$. This implies that θ has to develop a shock at $x = 0$ by time T - unless blow up happens before that in some other fashion (invalidating regularity assumptions underlying our estimates such as integration by parts). In fact, the informal flavor of bounds (38), (39) is that of $F'' \geq ce^{cF}$ differential inequality, leading to dramatically fast growth.

5 The SQG patch problem: a blow up blueprint

Part of the difficulty in securing rigorous understanding of the Hou-Luo blow up scenario for 3D axi-symmetric Euler or 2D inviscid Boussinesq equation lies in growth of ω, which destroys the estimate of error terms in Lemma 3.3. The error terms may no longer be of smaller order than the main term. In fact, heuristic computations taking ω that behaves approximately like some inverse power of x_1 in a certain region near origin - an ansatz that appears to be in agreement with the numerical simulations - indicate that the error terms will now be of the same order as the main term obtained by integration over the bulk. In this section, we discuss a different setting in which this situation is the case - the portion of the Biot-Savart integral pushing the solution towards blow up has the same order as the part pushing in the opposing direction. Nevertheless, the conclusion is finite time blow, essentially due to presence of a parameter that can be used to overcome the error term. The setting is that of modified SQG patch solutions in the half-plane.

Namely, let us in this section set $D = \mathbb{R}^2_+ = \{(x_1, x_2) \,|\, x_2 \geq 0\}$. The Bio-Savart law for the patch evolution on the half-plane $D := \mathbb{R} \times \mathbb{R}^+$ is

$$u = \nabla^\perp (-\Delta_D)^{-1+\alpha} \omega,$$

with Δ_D being the Dirichlet Laplacian on D, which can also be written as

$$u(x,t) := \int_D \left(\frac{(x-y)^\perp}{|x-y|^{2+2\alpha}} - \frac{(x-\bar{y})^\perp}{|x-\bar{y}|^{2+2\alpha}} \right) \omega(y,t)dy \tag{40}$$

The case $\alpha = 0$ corresponds to the 2D Euler equation, while $\alpha = 1/2$ to the SQG equation; the range $0 < \alpha < 1$ is called modified SQG. Note that u is divergence free and tangential to the boundary. A traditional approach to the 2D Euler ($\alpha = 0$) vortex patch evolution, going back to Yudovich (see Marchioro and Pulvirenti [1994] for an exposition) is via the corresponding flow map. The active scalar ω is advected by u from Equation (40) via

$$\omega(x,t) = \omega\left(\Phi_t^{-1}(x), 0\right), \tag{41}$$

where

$$\frac{d}{dt}\Phi_t(x) = u\left(\Phi_t(x), t\right) \qquad \text{and} \qquad \Phi_0(x) = x. \tag{42}$$

The initial condition ω_0 for (40)-(42) is patch-like,

$$\omega_0 = \sum_{k=1}^{N} \theta_k \chi_{\Omega_{0k}}, \tag{43}$$

with $\theta_1, \ldots, \theta_N \neq 0$ and $\Omega_{01}, \ldots, \Omega_{0N} \subseteq D$ bounded open sets, whose closures $\overline{\Omega_{0,k}}$ are pairwise disjoint and whose boundaries $\partial\Omega_{0k}$ are simple closed curves. The question of regularity of solution in patch setting becomes the question of the conservation of regularity class of the patch boundary, as well as lack of self-intersection or collisions between different patches.

One reason the Yudovich theory works for the 2D Euler equations is that for ω which is (uniformly in time) in $L^1 \cap L^\infty$, the velocity field u given by Equation (40) with $\alpha = 0$ is log-Lipschitz in space, and the flow map Φ_t is everywhere well-defined (see e.g. Majda and Bertozzi [2002] and Marchioro and Pulvirenti [1994]). In our situation, when ω is a patch solution and $\alpha > 0$, the flow u from Equation (40) is smooth away from the patch boundaries $\partial\Omega_k(t)$ but is only $1-\alpha$ Hölder constinuous at $\partial\Omega_k(t)$, which is exactly where one needs to use the flow map. This creates significant technical difficulties in proving local well-posedness of patch evolution in some reasonable functional space. For the case without boundaries, local well-posedness has been proved in Rodrigo [2005] for C^∞ patches and for Sobolev H^3 patches in Gancedo [2008] for $0 < \alpha \leq 1$. A naive intuition on why patch evolution can be locally well-posed for $\alpha > 0$ is that the below-Lipschitz loss of regularity only affects the tangential component of the fluid velocity at patch boundary. The normal to patch component, that intuitively should determine the evolution of the patch, retains stronger regularity.

In presence of boundaries, the problem is harder. Intuitively, one reason for the difficulties can be explained as follows. In the simplest case of half-plane the reflection principle implies that the boundary can be replaced by a reflected patch (or patches) of the opposite sign. If the patch is touching the boundary, then the reflected and original patch are touching each other, and the low regularity tangential component of the velocity field generated by the reflected patch has strong influence on the boundary of the original patch near touch points. Even in the 2D Euler case, the global regularity for patches in general domains with boundaries is currently open (partial results for patches not touching the boundary or with loss of regularity can be found in Depauw [1999], Dutrifoy [2003]). In the half-plane, a global regularity result has been recently established in Kiselev, Ryzhik, Yao, and Zlatoš [2016]:

Theorem 5.1. *Let $\alpha = 0$ and $\gamma \in (0, 1]$. Then for each $C^{1,\gamma}$ patch-like initial data ω_0, there exists a unique global $C^{1,\gamma}$ patch solution ω to (41), (40), (42) with $\omega(\cdot, 0) = \omega_0$.*

In the case $\frac{1}{24} > \alpha > 0$ with boundary, even local well-posedness results are highly non-trivial. The following result has been proved in Kiselev, Yao, and Zlatoš [2017] for the half-plane.

Theorem 5.2. *If $\alpha \in (0, \frac{1}{24})$, then for each H^3 patch-like initial data ω_0, there exists a unique local H^3 patch solution ω with $\omega(\cdot, 0) = \omega_0$. Moreover, if the maximal time T_ω of existence of ω is finite, then at T_ω a singularity forms: either two patches touch, or a patch boundary touches itself or loses H^3 regularity.*

We note that one has to be careful in the definition of solutions in this case as trajectories (42) may not be unique. Solutions can be defined in a weak sense by pairing with a test function, or in an appealing geometric way by specifying evolution of patch boundary with velocity (40) in the sense of Hausdorff distance; see Kiselev, Yao, and Zlatoš [ibid.] for more details. The constraint $\alpha < \frac{1}{24}$ appears due to estimates near boundary; it is not clear if it is sharp.

On the other hand, in Kiselev, Ryzhik, Yao, and Zlatoš [2016], it was proved that for any $\alpha > 0$, there exist patch-like initial data leading to finite time blow up.

Theorem 5.3. *Let $\alpha \in (0, \frac{1}{24})$. Then there are H^3 patch-like initial data ω_0 for which the unique local H^3 patch solution ω with $\omega(\cdot, 0) = \omega_0$ becomes singular in finite time (i.e., its maximal time of existence T_ω is finite).*

Together, Theorems 5.1 and 5.3 give rigorous meaning to calling the 2D Euler equation critical. In the half-plane patch framework $\alpha = 0$ is the exact threshold for phase transition from global regularity to possibility of finite time blow up.

In what follows, we will sketch proof of the blow up Theorem Theorem 5.3. We concentrate on the main ideas only; full details can be found in Kiselev, Ryzhik, Yao, and Zlatoš [ibid.]. Let us describe the initial data set up.Denote $\Omega_1 := (\varepsilon, 4) \times (0, 4)$, $\Omega_2 := (2\varepsilon, 3) \times (0, 3)$, and let $\Omega_0 \subseteq D^+ \equiv \mathbb{R}^+ \times \mathbb{R}^+$ be an open set whose boundary is a smooth simple closed curve and which satisfies $\Omega_2 \subseteq \Omega_0 \subseteq \Omega_1$. Here ϵ is a small parameter depending on α that will be chosen later.

Let $\omega(x, t)$ be the unique H^3 patch solution corresponding to the initial data

$$\omega(x, 0) := \chi_{\Omega_0}(x) - \chi_{\tilde{\Omega}_0}(x) \tag{44}$$

with maximal time of existence $T_\omega > 0$. Here, $\tilde{\Omega}_0$ is the reflection of Ω_0 with respect to the x_2-axis. Then

$$\omega(x, t) = \chi_{\Omega(t)}(x) - \chi_{\tilde{\Omega}(t)}(x) \tag{45}$$

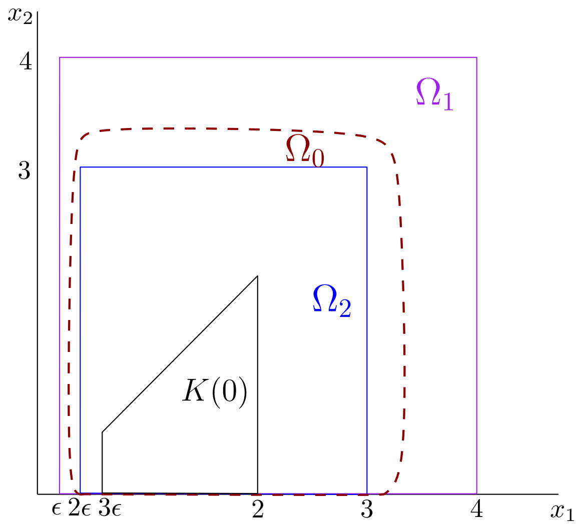


Figure 4: The domains $\Omega_1, \Omega_2, \Omega_0$, and $K(0)$ (with $\omega_0 = \chi_{\Omega_0} - \chi_{\tilde{\Omega}_0}$).

for $t \in [0, T_\omega)$, with $\Omega(t) := \Phi_t(\Omega_0)$. It can be seen from (40) that the rightmost point of the left patch on the x_1-axis and the leftmost point of the right patch on the x_1-axis will move toward each other. In the case of the 2D Euler equations $\alpha = 0$, Theorem 5.1 shows that the two points never reach the origin. When $\alpha > 0$ is small, however, it is possible to control the evolution sufficiently well to show that — unless the solution develops another singularity earlier — both points will reach the origin in a finite time. The argument yielding such control is fairly subtle, and the estimates do not extend to all $\alpha < \frac{1}{2}$, even though one would expect singularity formation to persist for more singular equations. This situation is not uncommon in the field: there is plenty of examples with the infinite in time growth of derivatives for the smooth solutions of 2D Euler equation, while none are available for the more singular SQG equation Kiselev and Nazarov [2012].

To show finite time blow up, we will deploy a barrier argument. Define

$$K(t) := \{x \in D^+ \ : \ x_1 \in (X(t), 2) \text{ and } x_2 \in (0, x_1)\} \tag{46}$$

for $t \in [0, T]$, with $X(0) = 3\epsilon$. Clearly, $K(0) \subset \Omega(0)$. Set the evolution of the barrier by

$$X'(t) = -\frac{1}{100\alpha} X(t)^{1-2\alpha}. \tag{47}$$

Then $X(T) = 0$ for $T = 50(3\epsilon)^{2\alpha}$. So if we can show that $K(t)$ stays inside $\Omega(t)$ while the patch solution stays regular, then we obtain that singularity must form by time T : the different patch components will touch at the origin by this time unless regularity is lost before that.

The key step in the proof involves estimates of the velocity near origin. In particular, u_1 needs to be sufficiently negative to exceed the barrier speed (47); u_2 needs to be sufficiently positive in order to ensure that $\Omega(t)$ cannot cross the barrier along its diagonal part. Note that it suffices to consider the part of the barrier that is very close to the origin, on the order $\sim \epsilon^{2\alpha}$. Indeed, the time T of barrier arrival at the origin has this order, and the fluid velocity satisfies uniform L^∞ bound that follows by a simple estimate which uses only $\alpha < 1/2$. Thus the patch $\Omega(t)$ has no time to reach more distant boundary points of the barrier before formation of singularity.

Let us focus on the estimates for u_1. For $y = (y_1, y_2) \in \bar{D}^+ = \mathbb{R}^+ \times \mathbb{R}^+$, we denote $\bar{y} := (y_1, -y_2)$ and $\tilde{y} := (-y_1, y_2)$. Due to odd symmetry, (40) becomes (we drop t from the notation in this sub-section)

$$u_1(x) = -\int_{D^+} K_1(x,y)\omega(y)dy, \tag{48}$$

where

$$K_1(x,y) = \underbrace{\frac{y_2 - x_2}{|x-y|^{2+2\alpha}}}_{K_{11}(x,y)} - \underbrace{\frac{y_2 - x_2}{|x-\tilde{y}|^{2+2\alpha}}}_{K_{12}(x,y)} - \underbrace{\frac{y_2 + x_2}{|x+y|^{2+2\alpha}}}_{K_{13}(x,y)} + \underbrace{\frac{y_2 + x_2}{|x-\bar{y}|^{2+2\alpha}}}_{K_{14}(x,y)}, \tag{49}$$

Analyzing (49), it is not hard to see that we can split the region of integration in the Biot-Savart law according to whether it helps or opposes the bounds we seek. Define

$$u_1^{bad}(x) := -\int_{\mathbb{R}^+\times(0,x_2)} K_1(x,y)\omega(y)dy \tag{50}$$

$$u_1^{good}(x) := -\int_{\mathbb{R}^+\times(x_2,\infty)} K_1(x,y)\omega(y)dy. \tag{51}$$

The following two lemmas contain key estimates.

Lemma 5.4 (Bad part). *Let $\alpha \in (0, \frac{1}{2})$ and assume that ω is odd in x_1 and $0 \le \omega \le 1$ on D^+. If $x \in \overline{D^+}$ and $x_2 \le x_1$, then*

$$u_1^{bad}(x) \le \frac{1}{\alpha}\left(\frac{1}{1-2\alpha} - 2^{-\alpha}\right) x_1^{1-2\alpha}. \tag{52}$$

The proof of this lemma uses (49) and after cancellations leads to the bound

$$u_1^{bad}(x) \le -\int_{(0,2x_1)\times(0,x_2)} \frac{y_2 - x_2}{|x-y|^{2+2\alpha}} dy, \tag{53}$$

which gives (52)

In the estimate of the good part, we need to use a lower bound on ω that will be provided by the barrier. Define

$$A(x) := \{y \,:\, y_1 \in (x_1, x_1+1) \text{ and } y_2 \in (x_2, x_2 + y_1 - x_1)\}\,. \tag{54}$$

Lemma 5.5 (Good part)**.** *Let $\alpha \in (0, \frac{1}{2})$ and assume that ω is odd in x_1 and for some $x \in \overline{D^+}$ we have $\omega \geq \chi_{A(x)}$ on D^+, with $A(x)$ from (54). There exists $\delta_\alpha \in (0,1)$, depending only on α, such that the following holds.*

If $x_1 \leq \delta_\alpha$, then

$$u_1^{good}(x) \leq -\frac{1}{6\cdot 20^\alpha \alpha} x_1^{1-2\alpha}.$$

Here analysis of (49) leads to

$$u_1^{good}(x) \leq \underbrace{-\int_{A_1} \frac{y_2 - x_2}{|x-y|^{2+2\alpha}}dy}_{T_1} + \underbrace{\int_{A_2} \frac{y_2 - x_2}{|x-y|^{2+2\alpha}}dy}_{T_2},$$

with the domains

$$\begin{aligned} A_1 &:= \{y \,:\, y_2 \in (x_2, x_2+1) \text{ and } y_1 \in (x_1 + y_2 - x_2, 3x_1 + y_2 - x_2)\}\,, \\ A_2 &:= (x_1+1, 3x_1+1) \times (x_2, x_2+1)\,. \end{aligned}$$

The term T_2 can be estimated by Cx_1, since the region of integration A_2 lies at a distance ~ 1 from the singularity. A relatively direct estimate of the term T_1 leads to the result of the Lemma.

A distinctive feature of the problem is that estimates for the “bad” and “good” terms appearing in Lemmas 5.4 and 5.5 above have the same order of magnitude $x_1^{1-2\alpha}$. This is unlike the 2D Euler double exponential growth construction, where we were able to isolate the main term. To understand the balance in the estimates for the “bad” and “good” terms, note that the “bad” term estimate comes from integration of the Biot-Savart kernel over rectangle $(0, 2x_1) \times (0, x_2)$, while the good term estimate from integration of the same kernel over the region A_1 above. When α is close to zero, the kernel is longer range, and the more extended nature of the region A_1 makes the “good” term dominate. In particular, the coefficient $\frac{1}{\alpha}\left(\frac{1}{1-2\alpha} - 2^{-\alpha}\right)$ in front of $x_1^{1-2\alpha}$ in Lemma 5.4 converges to to finite limit as $\alpha \to 0$, while the coefficient $\frac{1}{6\cdot 20^\alpha \alpha}$ in Lemma 5.5 tends to infinity. On the other hand, when $\alpha \to \frac{1}{2}$, the singularity in the Biot-Savart kernel is strong and getting close to non-integrable. Then it becomes important that the “bad” term integration region contains an angle π range near the singularity, while the “good” region only $\frac{\pi}{4}$. For this reason, controlling the "bad" term for larger values of α is problematic - although there is no reason why there cannot be a different, more clever argument achieving this goal.

It is straightforward to check that the dominance of the "good" term over "bad" one extends to the range $\alpha \in (0, \frac{1}{24})$, and that in this range we get as a result

$$u_1(x,t) \le -\frac{1}{50\alpha} x_1^{1-2\alpha} \tag{55}$$

for $x = (x_1, x_2)$ such that $x_1 \le \delta_\alpha$ and $x_1 \ge x_2$. A similar bound can be proved showing that

$$u_2(x,t) \ge \frac{1}{50\alpha} x_2^{1-2\alpha} \tag{56}$$

for $x = (x_1, x_2)$ such that $x_2 \le \delta_\alpha$ and $x_1 \le x_2$.

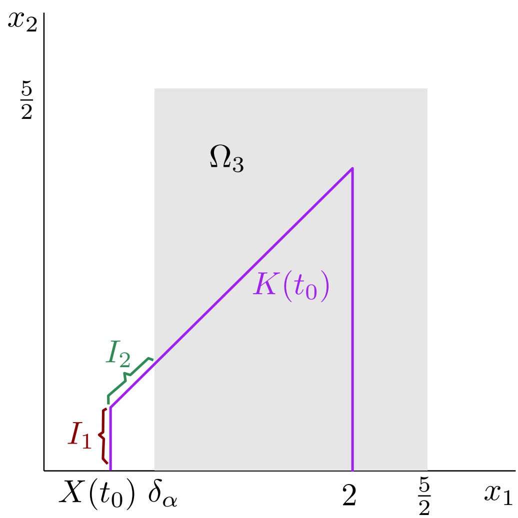


Figure 5: The segments I_1 and I_2 and the sets Ω_3 and $K(t_0)$.

The proof is completed by a contradiction argument, where we assume that the barrier $K(t)$ catches up with the patch $\Omega(t)$ at some time $t = t_0 < T$ of first contact. Taking ϵ sufficiently small compared to δ_α from Lemma 5.5, we can make sure the contact can only happen on the intervals I_1 and I_2 along the boundary of the barrier $K(t_0)$ appearing on Figure 5. But then bounds (55), (56) and the evolution of the barrier prescription (47) lead to the conclusion that the barrier should have been crossed at $t < t_0$, yielding a contradiction; full details can be found in Kiselev, Ryzhik, Yao, and Zlatoš [2016]).

6 Discussion

There are a few more recent papers that have contributed towards understanding the hyperbolic point blow up scenario. Two-dimensional simplified models of the 2D Boussinesq system have been considered in Hoang, Orcan-Ekmekci, Radosz, and Yang [n.d.] and in Kiselev and Tan [2018]. In both cases, the derivative forcing term in (6) is replaced by a simpler sign-definite approximation $\frac{\theta}{x_1}$, and the Biot-Savart law is replaced by a simpler version $u = (-x_1\Omega(x,t), x_2\Omega(x,t))$. In Hoang, Orcan-Ekmekci, Radosz, and Yang [n.d.], Ω takes form similar to the 2D Euler example (13). In Kiselev and Tan [2018], Ω is closely related but is also chosen to keep u incompressible. Both papers prove finite time blow up, Hoang, Orcan-Ekmekci, Radosz, and Yang [n.d.] by a sort of barrier argument while the argument Kiselev and Tan [2018] deploys an appropriate Lyapunov-type functional.

A very interesting recent work by Elgindi and Jeong takes a different approach Elgindi and Jeong [n.d.(b)], Elgindi and Jeong [n.d.(a)]. In Elgindi and Jeong [n.d.(b)], they look at a class of scale invariant solutions for the 2D Boussinesq system that satisfy $\frac{1}{\lambda}u(\lambda x, t) = u(x,t)$ and $\frac{1}{\lambda}\theta(\lambda x, t) = \theta(x,t)$. Observe that this class allows velocity and density that grow linearly at infinity. Also, the solution is not regular at the origin: for example the vorticity is just L^∞. The setting is a sector which has size $\frac{\pi}{2}$ (and some results can be generalized to other angles $< \pi$). First, they prove a local well-posedness theorem in a class of solutions that includes scale invariant solutions; additional symmetry assumptions are needed for this result. Secondly, for such solutions, they obtain an effective one-dimensional equation, some solutions of which are shown to blow up in finite time. These are not the first examples of infinite energy solutions (see Childress, Ierley, Spiegel, and Young [1989], Constantin [2000], and Sarria and Wu [2015]). However in Elgindi and Jeong [n.d.(b)] a procedure to cut off the solution at infinity while maintaining finite time blow up property is carried out. This yields finite energy solutions leading to finite time blow up - in the sense that $\int_0^T \|\nabla u(\cdot,t)\|_{L^\infty}\,dt \to \infty$ at blow up time. The vertex of the sector is a hyperbolic stagnation point of the flow, making connection to the Hou-Luo scenario. In Elgindi and Jeong [n.d.(a)], related results are announced and partly proved in the 3D axi-symmetric Euler case; here the domain is given by $z^2 \le c(|x|^2 + |y|^2)$ with a sufficiently small c. The finite time singularity formation in this setting remains open as the effective one-dimensional system turns out to be more complex.

The main challenge to analyzing smooth solutions to 2D Boussinesq and 3D axi-symmetric Euler equations in this context stems from difficulties estimating the velocity produced by Biot-Savart law with growing vorticity. It does not appear that there is a clear main term, as in 2D Euler example, or a clear small parameter to play in the same order of magnitude opposing terms, as in modified SQG patches. On the other hand, all model problems point to finite time blow up outcome in the original Hou-Luo scenario. The challenge is finding enough controllable structures to carry through rigorous analysis.

References

H. Bahouri and J.-Y. Chemin (1994). "Équations de transport relatives á des champs de vecteurs non-lipschitziens et mécanique des fluides". *Arch. Rational Mech. Anal.* 127.2, pp. 159–181. MR: 1288809 (cit. on p. 2388).

Martin Bauer, Boris Kolev, and Stephen C. Preston (2016). "Geometric investigations of a vorticity model equation". *J. Differential Equations* 260.1, pp. 478–516. MR: 3411680 (cit. on p. 2393).

J. T. Beale, T. Kato, and Andrew J. Majda (1984). "Remarks on the breakdown of smooth solutions for the 3-D Euler equations". *Comm. Math. Phys.* 94.1, pp. 61–66. MR: 763762 (cit. on p. 2383).

A. L. Bertozzi and P. Constantin (1993). "Global regularity for vortex patches". *Comm. Math. Phys.* 152.1, pp. 19–28. MR: 1207667 (cit. on p. 2384).

A. Bhattacharjee, C. S. Ng, and Xiaogang Wang (1995). "Finite-time vortex singularity and Kolmogorov spectrum in a symmetric three-dimensional spiral model". *Phys. Rev. E (3)* 52.5, part B, pp. 5110–5123. MR: 1385908 (cit. on p. 2383).

Oluş N Boratav and Richard B Pelz (1994). "Direct numerical simulation of transition to turbulence from a high-symmetry initial condition". *Physics of Fluids* 6.8, pp. 2757–2784 (cit. on p. 2383).

Russel E. Caflisch (1993). "Singularity formation for complex solutions of the 3D incompressible Euler equations". *Phys. D* 67.1-3, pp. 1–18. MR: 1234435 (cit. on p. 2383).

A. Castro and D. Córdoba (2010). "Infinite energy solutions of the surface quasi-geostrophic equation". *Adv. Math.* 225.4, pp. 1820–1829. MR: 2680191 (cit. on p. 2393).

Jean-Yves Chemin (1993). "Persistance de structures géométriques dans les fluides incompressibles bidimensionnels". *Ann. Sci. École Norm. Sup. (4)* 26.4, pp. 517–542. MR: 1235440 (cit. on p. 2384).

S. Childress, G. R. Ierley, E. A. Spiegel, and W. R. Young (1989). "Blow-up of unsteady two-dimensional Euler and Navier-Stokes solutions having stagnation-point form". *J. Fluid Mech.* 203, pp. 1–22. MR: 1002875 (cit. on p. 2404).

Kyudong Choi, Thomas Y. Hou, Alexander Kiselev, Guo Luo, Vladimir Sverak, and Yao Yao (2017). "On the finite-time blowup of a one-dimensional model for the three-dimensional axisymmetric Euler equations". *Comm. Pure Appl. Math.* 70.11, pp. 2218–2243. MR: 3707493 (cit. on pp. 2394, 2395).

Kyudong Choi, Alexander Kiselev, and Yao Yao (2015). "Finite time blow up for a 1D model of 2D Boussinesq system". *Comm. Math. Phys.* 334.3, pp. 1667–1679. MR: 3312447 (cit. on pp. 2394, 2397).

P. Constantin, P. D. Lax, and Andrew J. Majda (1985). "A simple one-dimensional model for the three-dimensional vorticity equation". *Comm. Pure Appl. Math.* 38.6, pp. 715–724. MR: 812343 (cit. on p. 2393).

Peter Constantin (2000). "The Euler equations and nonlocal conservative Riccati equations". *Internat. Math. Res. Notices* 9, pp. 455–465. MR: 1756944 (cit. on p. 2404).

– (2007). "On the Euler equations of incompressible fluids". *Bull. Amer. Math. Soc. (N.S.)* 44.4, pp. 603–621. MR: 2338368 (cit. on p. 2383).

Peter Constantin, Charles Fefferman, and Andrew J. Majda (1996). "Geometric constraints on potentially singular solutions for the 3-D Euler equations". *Comm. Partial Differential Equations* 21.3-4, pp. 559–571. MR: 1387460 (cit. on p. 2383).

Peter Constantin, Andrew J. Majda, and Esteban Tabak (1994). "Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar". *Nonlinearity* 7.6, pp. 1495–1533. MR: 1304437 (cit. on p. 2384).

Diego Córdoba, Marco A. Fontelos, Ana M. Mancho, and Jose L. Rodrigo (2005). "Evidence of singularities for a family of contour dynamics equations". *Proc. Natl. Acad. Sci. USA* 102.17, pp. 5949–5952. MR: 2141918 (cit. on p. 2384).

Salvatore De Gregorio (1990). "On a one-dimensional model for the three-dimensional vorticity equation". *J. Statist. Phys.* 59.5-6, pp. 1251–1263. MR: 1063199 (cit. on p. 2393).

– (1996). "A partial differential equation arising in a 1D model for the 3D vorticity equation". *Math. Methods Appl. Sci.* 19.15, pp. 1233–1255. MR: 1410208 (cit. on p. 2393).

Jian Deng, Thomas Y. Hou, and Xinwei Yu (2005). "Geometric properties and nonblowup of 3D incompressible Euler flow". *Comm. Partial Differential Equations* 30.1-3, pp. 225–243. MR: 2131052 (cit. on p. 2383).

Sergey A. Denisov (2009). "Infinite superlinear growth of the gradient for the two-dimensional Euler equation". *Discrete Contin. Dyn. Syst.* 23.3, pp. 755–764. MR: 2461825 (cit. on p. 2389).

– (2015). "Double exponential growth of the vorticity gradient for the two-dimensional Euler equation". *Proc. Amer. Math. Soc.* 143.3, pp. 1199–1210. MR: 3293735 (cit. on p. 2389).

Nicolas Depauw (1999). "Poche de tourbillon pour Euler 2D dans un ouvert à bord". *J. Math. Pures Appl. (9)* 78.3, pp. 313–351. MR: 1687165 (cit. on p. 2398).

Tam Do, Alexander Kiselev, and Xiaoqian Xu (n.d.). "Stability of Blow Up for a 1D model of Axisymmetric 3D Euler Equation". Preprint, to appear at Journal of Nonlinear Science. arXiv: 1604.07118 (cit. on p. 2394).

Alexandre Dutrifoy (2003). "On 3-D vortex patches in bounded domains". *Comm. Partial Differential Equations* 28.7-8, pp. 1237–1263. MR: 1998937 (cit. on p. 2398).

Weinan E and Chi-Wang Shu (1994). "Small-scale structures in Boussinesq convection". *Phys. Fluids* 6.1, pp. 49–58. MR: 1252833 (cit. on p. 2383).

Albert Einstein (1926). "The cause of the formation of meanders in the courses of rivers and of the so-called Baer's law". *Die Naturwissenschaften* 14.11. English translation in *Ideas and Opinions*, 1954, pp. 223–224 (cit. on p. 2386).

Tarek M. Elgindi and In-Jee Jeong (n.d.[a]). "Finite-time Singularity formation for Strong Solutions to the axi-symmetric $3D$ Euler Equations". arXiv: 1802.09936 (cit. on p. 2404).

– (n.d.[b]). "Finite-time Singularity Formation for Strong Solutions to the Boussinesq System". Preprint. arXiv: 1708.02724 (cit. on p. 2404).

– (n.d.[c]). "On the Effects of Advection and Vortex Stretching". Preprint. arXiv: 1701.04050 (cit. on p. 2393).

Joachim Escher and Boris Kolev (2014). "Right-invariant Sobolev metrics of fractional order on the diffeomorphism group of the circle". *J. Geom. Mech.* 6.3, pp. 335–372. MR: 3265830 (cit. on p. 2393).

L Euler (1755). *Principes généraux du mouvement des fluides*. Vol. 11. 4 September 1755, printed 1757, pp. 217–273 (cit. on p. 2381).

Gregory L. Eyink (2008). "Dissipative anomalies in singular Euler flows". *Phys. D* 237.14-17, pp. 1956–1968. MR: 2449784 (cit. on p. 2382).

Charles L Fefferman (2006). "Existence and smoothness of the Navier-Stokes equation". *The millennium prize problems* 57, p. 67 (cit. on p. 2383).

Francisco Gancedo (2008). "Existence for the α-patch model and the QG sharp front in Sobolev spaces". *Adv. Math.* 217.6, pp. 2569–2598. MR: 2397460 (cit. on pp. 2384, 2398).

J. D. Gibbon (2008). "The three-dimensional Euler equations: where do we stand?" *Phys. D* 237.14-17, pp. 1894–1904. MR: 2449772 (cit. on p. 2383).

Rainer Grauer and Thomas C Sideris (1991). "Numerical computation of 3D incompressible ideal fluids with swirl". *Physical review letters* 67.25, pp. 3511–3514 (cit. on p. 2383).

Isaac M. Held, Raymond T. Pierrehumbert, Stephen T. Garner, and Kyle L. Swanson (1995). "Surface quasi-geostrophic dynamics". *J. Fluid Mech.* 282, pp. 1–20. MR: 1312238 (cit. on p. 2384).

Vu Hoang, Betul Orcan-Ekmekci, Maria Radosz, and Hang Yang (n.d.). "Blowup with vorticity control for a 2D model of the Boussinesq equations". arXiv: 1608.01285 (cit. on p. 2404).

Ernst Hölder (1933). "Über die unbeschränkte Fortsetzbarkeit einer stetigen ebenen Bewegung in einer unbegrenzten inkompressiblen Flüssigkeit". *Math. Z.* 37.1, pp. 727–738. MR: 1545431 (cit. on pp. 2382, 2388).

Thomas Y. Hou and Congming Li (2005). "Global well-posedness of the viscous Boussinesq equations". *Discrete Contin. Dyn. Syst.* 12.1, pp. 1–12. MR: 2121245 (cit. on p. 2383).

Thomas Y. Hou and Ruo Li (2006). "Dynamic depletion of vortex stretching and non-blowup of the 3-D incompressible Euler equations". *J. Nonlinear Sci.* 16.6, pp. 639–664. MR: 2271429 (cit. on p. 2383).

Thomas Y. Hou and Pengfei Liu (2015). "Self-similar singularity of a 1D model for the 3D axisymmetric Euler equations". *Res. Math. Sci.* 2, Art. 5, 26. MR: 3341705 (cit. on p. 2394).

Hao Jia, Samuel Stewart, and Vladimir Sverak (n.d.). "On the De Gregorio modification of the Constantin–Lax–Majda Model". Preprint. arXiv: 1710.02737 (cit. on p. 2393).

V. I. Judovič (1963). "Non-stationary flows of an ideal incompressible fluid". *Ž. Vyčisl. Mat. i Mat. Fiz.* 3, pp. 1032–1066. MR: 0158189 (cit. on pp. 2384, 2389).

– (1974). "The loss of smoothness of the solutions of Euler equations with time". *Dinamika Splošn. Sredy* Vyp. 16 Nestacionarnye Problemy Gidrodinamiki, pp. 71–78, 121. MR: 0454419 (cit. on p. 2388).

Tosio Kato (1986). "Remarks on the Euler and Navier-Stokes equations in $\mathbf{R}^2$". In: *Nonlinear functional analysis and its applications, Part 2 (Berkeley, Calif., 1983)*. Vol. 45. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 1–7. MR: 843590 (cit. on p. 2388).

Robert M. Kerr (1993). "Evidence for a singularity of the three-dimensional, incompressible Euler equations". *Phys. Fluids A* 5.7, pp. 1725–1746. MR: 1223050 (cit. on p. 2383).

Alexander Kiselev and Fedor Nazarov (2012). "A simple energy pump for the surface quasi-geostrophic equation". In: *Nonlinear partial differential equations*. Vol. 7. Abel Symp. Springer, Heidelberg, pp. 175–179. MR: 3289364 (cit. on p. 2400).

Alexander Kiselev, Lenya Ryzhik, Yao Yao, and Andrej Zlatoš (2016). "Finite time singularity for the modified SQG patch equation". *Ann. of Math. (2)* 184.3, pp. 909–948. MR: 3549626 (cit. on pp. 2384, 2398, 2399, 2403).

Alexander Kiselev and Vladimir Šverák (2014). "Small scale creation for solutions of the incompressible two-dimensional Euler equation". *Ann. of Math. (2)* 180.3, pp. 1205–1220. MR: 3245016 (cit. on pp. 2383, 2388, 2389, 2391).

Alexander Kiselev and Changhui Tan (2018). "Finite time blow up in the hyperbolic Boussinesq system". *Adv. Math.* 325, pp. 34–55. arXiv: 1609.02468. MR: 3742585 (cit. on p. 2404).

Alexander Kiselev, Yao Yao, and Andrej Zlatoš (2017). "Local regularity for the modified SQG patch equation". *Comm. Pure Appl. Math.* 70.7, pp. 1253–1315. MR: 3666567 (cit. on pp. 2384, 2399).

Herbert Koch (2002). "Transport and instability for perfect fluids". *Math. Ann.* 323.3, pp. 491–523. MR: 1923695.

Adam Larios, Mark Petersen, Edriss S. Titi, and Beth Wingate (n.d.). "The Euler-Voigt equations and a computational investigation of the finite-time blow-up of solutions to the 3D Euler Equations". In: arXiv: 1512.07877 (cit. on p. 2383).

Guo Luo and Thomas Y. Hou (2014). "Toward the finite-time blowup of the 3D axisymmetric Euler equations: a numerical investigation". *Multiscale Model. Simul.* 12.4, pp. 1722–1776. MR: 3278833 (cit. on pp. 2383, 2385, 2394).

Andrew J. Majda (2003). *Introduction to PDEs and waves for the atmosphere and ocean*. Vol. 9. Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, pp. x+234. MR: 1965452 (cit. on p. 2384).

Andrew J. Majda and Andrea L. Bertozzi (2002). *Vorticity and incompressible flow*. Vol. 27. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, pp. xii+545. MR: 1867882 (cit. on pp. 2382, 2384, 2387, 2398).

Carlo Marchioro and Mario Pulvirenti (1994). *Mathematical theory of incompressible nonviscous fluids*. Vol. 96. Applied Mathematical Sciences. Springer-Verlag, New York, pp. xii+283. MR: 1245492 (cit. on pp. 2382, 2384, 2397, 2398).

Andrey Morgulis, Alexander Shnirelman, and Victor Yudovich (2008). "Loss of smoothness and inherent instability of 2D inviscid fluid flows". *Comm. Partial Differential Equations* 33.4-6, pp. 943–968. MR: 2424384 (cit. on p. 2388).

N. S. Nadirashvili (1991). "Wandering solutions of the two-dimensional Euler equation". *Funktsional. Anal. i Prilozhen.* 25.3, pp. 70–71. MR: 1139875 (cit. on p. 2388).

Koji Ohkitani and John D. Gibbon (2000). "Numerical study of singularity formation in a class of Euler and Navier-Stokes flows". *Phys. Fluids* 12.12, pp. 3181–3194. MR: 1796376 (cit. on p. 2383).

Hisashi Okamoto, Takashi Sakajo, and Marcus Wunsch (2008). "On a generalization of the Constantin-Lax-Majda equation". *Nonlinearity* 21.10, pp. 2447–2461. MR: 2439488 (cit. on p. 2393).

J Pedlosky (1987). *Geophysical fluid mechanics* (cit. on p. 2384).

RB Pelz and Y Gulak (1997). "Evidence for a real-time singularity in hydrodynamics from time series analysis". *Physical review letters* 79.25, p. 4998 (cit. on p. 2383).

Raymond T Pierrehumbert, Isaac M Held, and Kyle L Swanson (1994). "Spectra of local and nonlocal two-dimensional turbulence". *Chaos, Solitons & Fractals* 4.6, pp. 1111–1116 (cit. on p. 2384).

L Prandtl (1952). "Essentials of fluid dynamics, Blackie and Son, Ltd". *London and Glasgow* 452 (cit. on p. 2386).

Alain Pumir and Eric D. Siggia (1992). "Development of singular solutions to the axisymmetric Euler equations". *Phys. Fluids A* 4.7, pp. 1472–1491. MR: 1167779 (cit. on p. 2383).

José Luis Rodrigo (2005). "On the evolution of sharp fronts for the quasi-geostrophic equation". *Comm. Pure Appl. Math.* 58.6, pp. 821–866. MR: 2142632 (cit. on pp. 2384, 2398).

Alejandro Sarria and Jiahong Wu (2015). "Blowup in stagnation-point form solutions of the inviscid 2d Boussinesq equations". *J. Differential Equations* 259.8, pp. 3559–3576. MR: 3369255 (cit. on p. 2404).

M. Siegel and R. E. Caflisch (2009). "Calculation of complex singular solutions to the 3D incompressible Euler equations". *Phys. D* 238.23-24, pp. 2368–2379. MR: 2576080 (cit. on p. 2383).

Elias M. Stein (1970). *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., pp. xiv+290. MR: 0290095 (cit. on p. 2388).

W. Wolibner (1933). "Un theorème sur l'existence du mouvement plan d'un fluide parfait, homogène, incompressible, pendant un temps infiniment long". *Math. Z.* 37.1, pp. 698–726. MR: 1545430 (cit. on pp. 2382, 2383, 2388).

Marcus Wunsch (2011). "The generalized Constantin–Lax–Majda equation revisited". *Commun. Math. Sci.* 9.3, pp. 929–936. MR: 2865810 (cit. on p. 2393).

Xiaoqian Xu (2016). "Fast growth of the vorticity gradient in symmetric smooth domains for 2D incompressible ideal flow". *J. Math. Anal. Appl.* 439.2, pp. 594–607. MR: 3475939 (cit. on p. 2389).

V. I. Yudovich (1962). "The flow of a perfect, incompressible liquid through a given region". *Soviet Physics Dokl.* 7, pp. 789–791. MR: 0163529 (cit. on p. 2383).

– (2000). "On the loss of smoothness of the solutions of the Euler equations and the inherent instability of flows of an ideal fluid". *Chaos* 10.3, pp. 705–719. MR: 1791984 (cit. on p. 2388).

– (2003). "Eleven great problems of mathematical hydrodynamics". *Mosc. Math. J.* 3.2. Dedicated to Vladimir I. Arnold on the occasion of his 65th birthday, pp. 711–737, 746. MR: 2025281 (cit. on p. 2383).

Andrej Zlatoš (2015). "Exponential growth of the vorticity gradient for the Euler equation on the torus". *Adv. Math.* 268, pp. 396–403. MR: 3276599 (cit. on p. 2389).

Received 2017-11-16.

ALEXANDER KISELEV

Department of Mathematics
Duke University
Durham NC 27708
USA

and

Department of Mathematics
Rice University
Houston TX 77005
USA
kiselev@math.duke.edu
kiselev@rice.edu

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2409–2430)

QUANTITATIVE PROPAGATION OF SMALLNESS FOR SOLUTIONS OF ELLIPTIC EQUATIONS

ALEXANDER LOGUNOV AND EUGENIA MALINNIKOVA

Abstract

Let u be a solution to an elliptic equation $\operatorname{div}(A\nabla u) = 0$ with Lipschitz coefficients in $\mathbb{R}^n$. Assume $|u|$ is bounded by 1 in the ball $B = \{|x| \le 1\}$. We show that if $|u| < \varepsilon$ on a set $E \subset \frac{1}{2}B$ with positive n-dimensional Hausdorf measure, then

$$|u| \le C\varepsilon^\gamma \text{ on } \frac{1}{2}B,$$

where $C > 0, \gamma \in (0,1)$ do not depend on u and depend only on A and the measure of E. We specify the dependence on the measure of E in the form of the Remez type inequality. Similar estimate holds for sets E with Hausdorff dimension bigger than $n-1$.

For the gradients of the solutions we show that a similar propagation of smallness holds for sets of Hausdorff dimension bigger than $n-1-c$, where $c > 0$ is a small numerical constant depending on the dimension only.

1 Introduction

This paper contains several quantitative results on propagation of smallness for solutions of elliptic PDE. The results concern the logarithms of the magnitudes of the solutions and their gradients. The techniques used in this paper were recently applied to estimates of zero sets of Laplace eigenfunctions Logunov and Malinnikova [2018] and Logunov [2018a,b].

The inspiration comes from the following useful fact from complex analysis: *if f is a holomorphic function on $\mathbb{C}$, then* $\log|f|$ *is subharmonic.* For this simple and powerful fact there are no known direct analogs for real valued solutions of elliptic PDE on $\mathbb{R}^n$,

A. L. was supported in part by ERC Advanced Grant 692616, ISF Grants 1380/13, 382/15 and by a Schmidt Fellowship at the Institute for Advanced Study. E. M. was supported by Project 213638 of the Research Council of Norway.

MSC2010: primary 58G25; secondary 35P99.

except for the gradients of harmonic functions on $\mathbb{R}^2$, which can be identified with the holomorphic functions. For a harmonic function u in $\mathbb{R}^n$, $n \geq 3$, it is no longer true that $\log|\nabla u|$ is necessarily subharmonic; it was shown by E. Stein that $|\nabla u|^p$ is subharmonic when $p \geq (n-2)/(n-1)$. However some logarithmic convexity properties for harmonic functions still hold. One example is the classical three spheres theorem, which claims that for solutions u to a reasonable uniformly elliptic equation $Lu = 0$ in $\mathbb{R}^n$ (one can think that $L = \Delta$) the following inequality holds

$$\sup_B |u| \leq C(\sup_{\frac{1}{2}B} |u|)^\gamma (\sup_{2B} |u|)^{1-\gamma}, \tag{1}$$

where $B = \{x \in \mathbb{R}^n : |x| \leq 1\}$, constants $C > 0, \gamma \in (0,1)$ depend only on the elliptic operator L and do not depend on u.

The three spheres theorem holds for linear uniformly elliptic PDE of higher order under some smoothness assumptions on the coefficients (Sitnikova [1970]) as well as for some non-linear elliptic equations (Capuzzo Dolcetta [2002]).

1.1 Three spheres theorem for wild sets. Throughout this paper Ω will be a bounded domain in $\mathbb{R}^n$ and u will denote a solution of an elliptic equation in the divergence form $\operatorname{div}(A\nabla u) = 0$ in Ω with Lipschitz coefficients. We will show that in the three spheres theorem one can replace $\sup_{\frac{1}{2}B} |u|$ by the supremum over any set E with positive volume.

Let E and $\mathcal{K}$ be subsets of Ω such that the distances from E and $\mathcal{K}$ to $\partial\Omega$ are positive. We assume that E has positive n-dimensional Lebesgue measure. We aim to prove the following estimate

$$\sup_{\mathcal{K}} |u| \leq C(\sup_E |u|)^\gamma (\sup_\Omega |u|)^{1-\gamma}, \tag{2}$$

where $C > 0$ and $\gamma \in (0,1)$ are independent of u, but depend on Ω, A, the measure of E, and the distances from $\mathcal{K}$ and E to the boundary of Ω.

If $\sup_\Omega |u| = 1$ and $\sup_E |u| = \varepsilon$, then (2) can be written as

$$\sup_{\mathcal{K}} |u| \leq C\varepsilon^\gamma. \tag{3}$$

This inequality explains why the result is called propagation of smallness. Typically, we start with some set, where we know that the solution is small, and then we make a conclusion that it is also small on a bigger set.

The fact that the set E is allowed to be arbitrary wild, while the constants depend only on its measure, seems to be useful for applications, see Apraiz, Escauriaza, Wang, and Zhang [2014]. Further we will specify the dependence of constants on the measure in the form of the Remez type inequality.

1.2 Preceding results. The result that we prove is expected. We would like to mention the preceding work in this direction. In the case of analytic coefficients the estimate (2) was proved by Nadiraŝvili [1979], see also Vessella [1999]. The case of C^∞-smooth coefficients remained open till now, but there were several attempts to prove it. Estimates, weaker than (3), were obtained by Nadiraŝvili [1986] and Vessella [2000]. See also Malinnikova and Vessella [2012], where the case of solutions of elliptic equations with singular lower order coefficients is treated. In the preceding results the exponent ε^γ in the right-hand side of (3) was replaced by $\exp(-c|\log \varepsilon|^p)$ for some $p = p(n) < 1$. We push p to 1 in this paper.

1.3 Remez type inequality. In this note we prove (2) in the setting of smooth coefficients, using the new results on the behavior of the doubling index of solutions to elliptic equations presented in Logunov and Malinnikova [2018] and Logunov [2018a,b]. On the way of proving (2) we obtain an interesting inequality for solutions of elliptic equations, which reminds the classical Remez inequality for polynomials, the role of the degree is now played by the doubling index.

Let Q be the unit cube. Assume u is a solution to $\operatorname{div}(A\nabla u) = 0$ with the doubling index $N = \log \frac{\sup_{2Q}|u|}{\sup_Q |u|}$. Then

$$\sup_Q |u| \le C \sup_E |u| \left(C \frac{|Q|}{|E|} \right)^{CN} \tag{4}$$

where C depends on A only, E is any subset of Q of a positive measure.

Note that if u is a harmonic polynomial in $\mathbb{R}^n$, then one can replace N by the degree of u. The doubling index for harmonic polynomials can be estimated from above by the degree of the polynomial.

Garofalo and Lin [1986] proved almost monotonicity of doubling index for solutions of second order elliptic PDEs and applied this result to prove unique continuation properties. In particular, they showed that both $|u|^2$ and $|\nabla u|^2$ are Muckenhoupt weights with parameters that depend on the maximal doubling index. This implies (4) with some implicit power $const(N)$ in place of CN and with L^2 norm in place of sup norm.

1.4 Propagation of smallness from sets with big Hausdorff dimension. The assumption that E has positive n-dimensional Lebesgue measure can be relaxed. It is enough to assume that the dimension of E is larger than $n-1$, see Malinnikova [2004] for the details in the analytic case. We fix the Hausdorff content of E of some order $n-1+\delta$ with $\delta > 0$ and obtain inequality (2). The main Lemma 4.2 gives an upper estimate for the Hausdorff content of the set where the solution is small.

1.5 Propagation of smallness for gradients. In Section 5 we prove a result for the gradients of solutions of elliptic PDEs, which is new even for ordinary harmonic functions in $\mathbb{R}^n$, $n \geq 3$. Propagation of smallness for the gradients of solutions is better than for the solutions themselves. More precisely, the inequality remains the same

$$\sup_{\mathcal{K}} |\nabla u| \leq C (\sup_E |\nabla u|)^\gamma (\sup_\Omega |\nabla u|)^{1-\gamma}, \tag{5}$$

but now the set E is allowed to be smaller. Namely, we show that there is a constant $c = c(n) \in (0, 1)$ such that (5) is valid for sets E with Hausdorff dimension

$$\dim_{\mathcal{H}}(E) > n - 1 - c.$$

We give the precise statement in Section 5.

Precaution. We warn the reader that the paper is not self-contained: sometimes we use recent results, which are proved in other papers. Namely, we use the technique of counting doubling indices developed in Logunov and Malinnikova [2018] and Logunov [2018a,b] and in Section 5 we rely on estimates for sublevel sets of the gradients of solutions obtained in Cheeger, Naber, and Valtorta [2015].

1.6 Open questions. We propagate smallness (for gradients) from sets of Hausdorff dimension bigger than $n - 1 - c$. It would be interesting to obtain quantitative estimates for propagation of smallness from sets of Hausdorff dimension greater than $n - 2$. There are qualitative stratification results for critical sets Cheeger, Naber, and Valtorta [ibid.] and exponential estimates for the $n-2$-dimensional Hausdorff measure of the critical set Naber and Valtorta [2017] that suggest that $n - 2$ is the correct threshold.

Question 1. Is it true that the inequality (5) holds for sets E with $\dim_{\mathcal{H}}(E) > n - 2$ and the constants can be chosen to depend only on the operator A, domain Ω, the distances from E and $\mathcal{K}$ to the boundary of Ω and the Hausdorff content of E of order $n - 2 + \delta$ for any $\delta > 0$?

Such estimates would be related to a conjecture by Lin [1991] on the size of the critical sets of solutions. For the sake of simplicity we formulate Lin's conjecture for ordinary harmonic functions, we also slightly modify the definition of the frequency.

Conjecture 1 (Fang-Hua Lin). Let u be a non-zero harmonic function in the unit ball $B_1 \subset \mathbb{R}^n$, $n \geq 3$. Consider

$$N = \log \frac{\sup_{B_1} |\nabla u|}{\sup_{B_{1/2}} |\nabla u|}$$

Is it true that the Hausdorff measure

$$\mathcal{H}^{n-2}(\{\nabla u = 0\} \cap B_{1/2}) \leq C_n N^2$$

for some C_n depending only on the dimension?

Recently Naber and Valtorta [2017] proved an exponential bound C^{N^2} for the Hausdorff measure of the critical set.

An interesting topic in the propagation of smallness which we don't touch in this paper is the dependence of the constants in (2) and (5) on the distance from the set E to the boundary of Ω.

Question 2. Consider the inequality (5) with a set E of $\dim_{\mathcal{H}}(E) = n-1$ and fixed $\mathcal{K}$, A and Ω. How do the constants C and γ depend on the distance from E to the boundary of Ω?

This question is connected to the quantitative version of the Cauchy uniqueness problem, see Lin [1991] for related results when E is a relatively open subset of the boundary. The situation changes when we consider wild sets on the boundary of positive surface measure. The following question is quite famous, it dates back to at least L. Bers. The two-dimensional case is not difficult due to connections with complex analysis. The fact that the question is open in higher dimensions shows that we still don't understand well the Cauchy uniqueness problem even for ordinary harmonic functions in the dimension three or higher (which is quite embarrassing for the well-developed theory of elliptic PDEs nowadays).

Conjecture 2. Assume that u is a harmonic function in the unit ball $B_1 \subset \mathbb{R}^3$ and u is C^∞-smooth in the closed ball $\overline{B_1}$. Let $S \subset \partial B_1$ be any closed set with positive area. Is it true that $\nabla u = 0$ on S implies $\nabla u \equiv 0$?

Usually this question is asked in the form of the Cauchy uniqueness problem, where the condition $\nabla u = 0$ is replaced by the condition that the Cauchy data $(u, \frac{\partial u}{\partial n})$ are zero on S. If one takes any Lebesgue point of S, then harmonicity of u and C^∞-smoothness automatically implies that all the derivatives of u of any order are zero at this point. Since the area of Lebesgue points of S is the same as of S, one can also assume (in the question above) that all the derivatives of u vanish at the boundary subset of positive area and the question is whether the harmonic function u should be identically zero.

For the class $C^{1+\varepsilon}(\overline{B_1})$ there is a striking counterexample Bourgain and Wolff [1990], which however is not C^∞-smooth up to the boundary. The attempts to construct C^2-smooth counterexamples were not successful.

1.7 Estimates for Laplace eigenfunctions.

Let (M, g) be a C^∞ smooth closed Riemannian manifold and let Δ denote the Laplace operator on M. Consider the sequence of Laplace eigenfunctions φ_λ on M with $\Delta\varphi_\lambda + \lambda\varphi_\lambda = 0$.

We would like to make a remark that the Remez type inequality (4) for harmonic functions implies the following bound for Laplace eigenfunctions, which was conjectured in

Donnelly and Fefferman [1990]. For any subset E of M with positive volume the following holds:

$$\sup_E |\varphi_\lambda| \geq \frac{1}{C} \sup_M |\varphi_\lambda| \left(\frac{|E|}{C|M|} \right)^{C\sqrt{\lambda}}, \tag{6}$$

where $C = C(M, g) > 1$ does not depend on E and λ. Note that $\sqrt{\lambda}$ corresponds to the degree of the polynomial in Remez inequality.

Looking at the following example of spherical harmonics $u(x, y, z) = \Re(x+iy)^n$ one can see that L^2 norm of restriction of u on the unit sphere is concentrated near the equator very fast and $|u|$ is exponentially small on most of the unit sphere. This example shows that a sequence Laplace eigenfunctions can be $e^{-c\sqrt{\lambda}}$ small on a fixed open subset of the manifold.

The proof of implication (4) $\implies$ (6) is a standard trick, we give a sketch of the proof of the implication, which is not difficult.

The function $u(x,t) = \varphi_\lambda(x)e^{\sqrt{\lambda}t}$ is a harmonic function on the product manifold $M \times \mathbb{R}$. The doubling index N of φ in any geodesic ball is smaller than $C_1\sqrt{\lambda}$ (Donnelly and Fefferman [1988]). Then the doubling index for u in any geodesic ball of radius smaller than the diameter of M is also smaller than $C_2\sqrt{\lambda}$. One can apply (4) to u with N=$C_2\sqrt{\lambda}$ and get the bound (6) for φ.

It seems that for negatively curved Riemannian manifolds one can prove better versions of (6). We don't feel the curvature in our methods.

We would like to mention an outstanding recent result from the works by Bourgain and Dyatlov [n.d.] and Dyatlov and Jin [n.d.].

Theorem 1.1 (Bourgain and Dyatlov [n.d.],Dyatlov and Jin [n.d.]). *Under assumption that (M, g) is a closed Riemannian surface with constant negative curvature the following inequality holds for Laplace eigenfunctions on M. Given an open subset E of M there exists $c = c(E, M, g) > 0$ such that*

$$\int_E \varphi_\lambda^2 \geq c \int_M \varphi_\lambda^2.$$

The constant c does not depend on the eigenvalue λ. Note that the situation on closed surfaces of constant negative curvature is different from the case of the sphere.

A beautiful result Bourgain and Rudnick [2009] states that on a two dimensional torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ equipped with the standard metric the toral Laplace eigenfunctions φ_λ satisfy L^2 lower and upper restriction bounds on curves. Namely, if S is a smooth curve on T^2 with non-zero curvature and $\lambda > const(S)$, then

$$c\|\varphi_\lambda\|_{L^2(S)} \leq \|\varphi_\lambda\|_{L^2(T^2)} \leq C\|\varphi_\lambda\|_{L^2(S)}.$$

In particular that implies that on a given smooth curve, which is not geodesic, only a finite number of Laplace eigenfunctions can vanish.

A very interesting question that we don't touch here is how L^2 mass of Laplace eigenfunctions φ_λ are asymptotically distributed on the manifold as $\lambda \to \infty$. In particular, for negatively curved surfaces the quantum unique ergodicity conjecture states that asymptotically the L^2 mass of eigenfunctions is distrubed uniformly. We refer to Sarnak [2011], Lindenstrauss [2006], Šnirelman [1974], Zelditch [1987, 2010], Zworski [2012], and Dyatlov and Jin [n.d.] for the results on ergodic properties of eigenfunctions.

Acknowledgments. This work was partly done when the second author held a one year visiting position at the Department of Mathematics at Purdue University and it is a pleasure to thank the department for its hospitality.

2 Preliminaries

2.1 Hausdorff content. Remind that the Hausdorff content of a set $E \subset \mathbb{R}^n$ is

$$C^d_{\mathcal{H}}(E) = \inf\Big\{ \sum_j r_j^d \,:\, E \subset \cup_j B(x_j, r_j) \Big\},$$

and the Hausdorff dimension of E is defined as

$$\dim_{\mathcal{H}}(E) = \inf\{d \,:\, C^d_{\mathcal{H}}(E) = 0\}.$$

Clearly the Hausdorff content is sub-additive

$$C^d_{\mathcal{H}}(E_1 \cup E_2) \le C^d_{\mathcal{H}}(E_1) + C^d_{\mathcal{H}}(E_2).$$

It also satisfies the natural scaling identity, if ϕ_t is a homothety of $\mathbb{R}^n$ with coefficient t then

$$C^d_{\mathcal{H}}(\phi_t(E)) = t^d C^d_{\mathcal{H}}(E).$$

The advantage of the Hausdorff content over the corresponding Hausdorff measure is that the former is always finite on bounded sets, it is bounded from above by $\operatorname{diam}(E)^d$. The Hausdorff content of order n is equivalent to the Lebesgue measure.

2.2 Three spheres theorem for wild sets. We always assume that u is a solution of an elliptic equation in divergence form in a bounded domain $\Omega \subset \mathbb{R}^n$,

$$\operatorname{div}(A \nabla u) = 0, \tag{7}$$

where $A(x) = [a_{ij}(x)]_{1\le i,j\le n}$ is a symmetric uniformly elliptic matrix with Lipschitz entries,

$$\Lambda_1^{-1}\|\zeta\|^2 \le \langle A\zeta,\zeta\rangle \le \Lambda_1\|\zeta\|^2, \quad |a_{ij}(x)-a_{ij}(y)| \le \Lambda_2|x-y|. \tag{8}$$

Let m,δ,ρ be positive numbers. Suppose a set $E\subset\Omega$ satisfies

$$C_{\mathcal{H}}^{n-1+\delta}(E) > m, \quad \operatorname{dist}(E,\partial\Omega) > \rho.$$

Let $\mathcal{K}$ be a subset of Ω with $\operatorname{dist}(\mathcal{K},\partial\Omega) > \rho$. Our main result is the following.

Theorem 2.1. *There exist $C,\gamma>0$, depending on m,δ,ρ,A,Ω only such that*

$$\sup_{\mathcal{K}}|u| \le C(\sup_E|u|)^{\gamma}(\sup_{\Omega}|u|)^{1-\gamma}$$

for any solution u of $\operatorname{div}(A\nabla u)=0$ *in* Ω.

2.3 Doubling index. We formulate several well-known lemmas connected to the three spheres theorem (or monotonicity property of the frequency function of a solution). We refer to Han and Lin [n.d.] for an introduction to the frequency function, which is almost a synonym for the doubling index (the term "'frequency"' will not be used in this paper).

Let B be a ball in $\mathbb{R}^n$. Define the doubling index of a non-zero function u (defined in $2B$) by

$$N(u,B) = \log\frac{\sup_{2B}|u|}{\sup_B|u|}.$$

It is a non-trivial fact that the doubling index of solutions to an elliptic second order PDE in divergence form is almost monotonic in the following sense:

$$N(tB) \le N(B)(1+c) + C \tag{9}$$

for any positive $t\le 1/2$. Here as usual tB denotes a ball of radius t times the radius of B with the same center as B, the constants $c,C>0$ depend on A, but are independent of u. Almost monotonicity of the doubling index implies the three spheres theorem. The three spheres theorem related to the almost monotonicity property of the doubling index was proved in the work Landis [1963]. Garofalo and Lin [1986] proved a sharper version of the monotonicity property. In particular, the results of Garofalo and Lin [ibid.] imply that if the elliptic operator is a small perturabtion of the Laplace operator, then c in (9) can be chosen to be small (C is still big, but it is less important). We refer the reader to Mangoubi [2013] and Logunov [2018a] for further discussion.

For a cube Q in $\mathbb{R}^n$ let $s(Q)$ denote its side length and let tQ be the cube with the same center as Q and such that $s(tQ) = ts(Q)$. Suppose that $(20n)Q \subset \Omega$. We define the doubling index of a function u in the cube Q by

$$N(u, Q) = \sup_{x \in Q, r \leq s(Q)} \log \frac{\sup_{B(x,10nr)} |u|}{\sup_{B(x,r)} |u|}. \tag{10}$$

This is a kind of maximal version of the doubling index, which is convenient in the sense that if a cube q is a subset of Q, then $N(u, q) \leq N(u, Q)$. The definition implies the following estimate. Let q be a subcube of Q and $K = \frac{s(Q)}{s(q)} \geq 2$. Then

$$\sup_q |u| \geq cK^{-CN} \sup_Q |u|, \tag{11}$$

where $N = N(u, Q)$ and c and C depend on n only.

3 Auxiliary lemmas

3.1 Estimates of the zero set. The doubling index is useful for estimates of the zero set of solutions of elliptic equations. We will need the following known result.

Lemma 3.1 (Hardt and Simon [1989]). *Let u be a solution to* $\operatorname{div}(A\nabla u) = 0$ *in* $\Omega \supset 20nQ$. *For any $N > 0$ there exists C_N, which is independent of u, but depends on A, N and Ω, such that if $N(u, Q) \leq N$, then*

$$\mathcal{H}^{n-1}(\{u = 0\} \cap Q) \leq C_N s(Q)^{n-1}. \tag{12}$$

We will use only finiteness of C_N and will apply it for N smaller than some numerical constant.

Remark 3.2. One can ask what is the optimal upper bound. The harmonic counterpart of the Yau conjecture suggests that there is a linear estimate:

$$\mathcal{H}^{n-1}(\{u = 0\} \cap Q) \overset{?}{\leq} C_{\Omega,A} N.$$

The conjecture is open, but known in the case of analytic coefficients due to results by Donnelly and Fefferman [1988]. In the setting of smooth coefficients an exponential bound (CN^{CN}) was proved in Hardt and Simon [1989], a recent result in Logunov [2018a] provides a polynomial upper bound CN^{α}, $\alpha > 1$ depends on the dimension.

The measure of the zero set can be also estimated from below. We assume that u and Q are as in Lemma 3.1.

Lemma 3.3. *Let q be a subcube of Q and suppose that u has a zero in q. Then*

$$\mathcal{H}^{n-1}(\{u=0\}\cap 2q) \geq c_N s^{n-1}(q), \tag{13}$$

where c_N depends on A, Ω and $N = N(u, Q)$.

Remark 3.4. The following much stronger version of this estimate is proved in Logunov [2018b],

$$\mathcal{H}^{n-1}(\{u=0\}\cap 2q) \geq c s^{n-1}(q)$$

where c depends on A, Ω only. We will use only the weak inequality (13) above, which is not difficult (see for example Logunov and Malinnikova [2018]).

3.2 Estimate for sub-level sets. The following lemma gives an estimate for the size of the set where a solution to an elliptic PDE is small in terms of the doubling index. We note that the lemma below is qualitative, but not quantitative (in a sense that there is no control of constants in terms of N). The lemma will be further refined to a quantitative version (Lemma 4.2).

Lemma 3.5. *Let $\delta \in (0,1]$, $N > 0$. Assume that u satisfies* $\operatorname{div}(A\nabla u) = 0$ *in* $(20n)Q$, $\sup_Q |u| = 1$ *and* $N(u, Q) \leq N$. *Let*

$$E_a = \{x \in \frac{1}{2}Q : |u(x)| < e^{-a}\}.$$

Then

$$C_{\mathcal{H}}^{n-1+\delta}(E_a) \leq M e^{-\beta a} s(Q)^{n-1+\delta},$$

for some $\beta = \beta(N, \delta, A, \Omega)$ and $M = M(N, \delta, A, \Omega)$.

Proof. By $c, C, \kappa, c_1, C_1 \ldots$ we will denote positive constants that depend on δ, A, and Ω only, while constants c_N, C_N additionally depend on N.

Clearly, it is enough to prove the statement for $a >> N$ and $a >> 1$. For small a the inequality holds if we choose M large enough to satisfy the inequality. Without loss of generality we may assume that $N \geq 2$.

Let $K = [e^{\kappa a/N}]$ where $\kappa > 0$ is a sufficiently small constant to be specified later. Partition $\frac{1}{2}Q$ into K^n equal subcubes q_i. We will assume that $K > 4$, then $4q_i \subset Q$. We will estimate the number of cubes q_i that intersect E_a.

Let q_i be a cube with $q_i \cap E_a \neq \varnothing$. So $\inf_{q_i} |u| \leq e^{-a}$.

Assume first that u does not change sign in $2q_i$. Then by the Harnack inequality

$$\sup_{q_i} |u| \leq c_1 \inf_{q_i} |u| \leq c_1 e^{-a}.$$

On the other hand by (11) we have

$$\sup_{q_i} |u| \geq c_2 K^{-C_1 N} \geq \frac{c_2}{2} e^{-C_1 \kappa a}.$$

Now, we specify $\kappa = \frac{1}{2C_1}$. Then the two inequalities above cannot coexist for large a.

Hence if q_i intersects E_a, then u changes sign in $2q_i$. Denote by S the set of cubes q_i such that u changes sign in $2q_i$. Note that

$$C_{\mathcal{H}}^{n-1+\delta}(E_a) \leq C_2 |S| s(Q)^{n-1+\delta} K^{-n+1-\delta}. \tag{14}$$

Now, we will estimate $|S|$ using the bounds for the size of the zero set of u. Note that u has a zero in each $2q_i$ for $q_i \in S$. Recall that $4q_i \subset Q$ and each point in Q may be covered only by a finite number of $4q_i$, depending only on the dimension. By Lemma 3.3

$$\mathcal{H}^{n-1}(\{u=0\} \cap Q) \geq c_3 \sum_S \mathcal{H}^{n-1}(\{u=0\} \cap 4q_i) \geq c_4 c_N |S| s(Q)^{n-1} K^{-n+1}.$$

On the other hand, by Lemma 3.1

$$\mathcal{H}^{n-1}(\{u=0\} \cap Q) \leq C_N s(Q)^{n-1}.$$

We therefore have

$$|S| \leq \frac{C_N}{c_4 c_N} K^{n-1}.$$

Thus by (14)

$$C_{\mathcal{H}}^{n-1+\delta}(E_a) \leq C_3 \frac{C_N}{c_N} K^{-\delta} s(Q)^{n-1+\delta} \leq C_4 \frac{C_N}{c_N} e^{-\kappa\delta a/N} s(Q)^{n-1+\delta},$$

which is the required estimate with $\beta = \kappa\delta/N$ and $M = C_4 C_N c_N^{-1}$. □

Remark 3.6. Note that the almost monotonicity of the doubling index implies that for any subcube q of Q with $4\sqrt{n}s(q) < s(Q)$ one has $N(u, 2q) \leq CN(u, Q)$. Then partitioning Q into a finite number of small cubes q and applying Lemma 3.5 to cubes $2q$, we obtain the following estimate

$$C_{\mathcal{H}}^{n-1+\delta}(\{x \in Q : |u(x)| < e^{-a}\}) \leq M_1 e^{-\beta a} s(Q)^{n-1+\delta}.$$

Remark 3.7. In Logunov [2018b,a] it was shown that one can choose c_N independent of N and $C_N = CN^\alpha$, where α depends only on the dimension. Hence for $N \geq 1$,

$$C_{\mathcal{H}}^{n-1+\delta}(\{|u| < e^{-a}\} \cap \frac{1}{2}Q) \leq CN^\alpha e^{-c\delta a/N} s(Q)^{n-1+\delta}.$$

The optimal estimates for c_N and C_N will appear to be not necessary for the purposes of this paper. In Lemma 4.2 we will prove a better bound for $C_{\mathcal{H}}^{n-1+\delta}(E_a)$ without using the uniform lower bound for c_N or polynomial bound for C_N.

3.3 Main tool. The following lemma will be severely exploited in the proof of main results. See Section "Number of cubes with big doubling index" in Logunov [2018b] for the proof of the lemma formulated below. We note that the definition of the doubling index in Logunov [ibid.] is slightly different (but the proof remains the same).

Lemma A. *Let u be a solution to $\operatorname{div}(A\nabla u) = 0$ in Ω. There exist positive constants s_0, N_0, B_0 that depend on A, Ω only such that if Q is a cube with $s(Q) < s_0$, $(20n)Q \subset \Omega$, and Q is divided into B^n equal subcubes with $B > B_0$, then the number of subcubes q with $N(u, q) \geq \max(\frac{1}{2}N(u, Q), N_0)$ is less than B^{n-1-c}, where c depends on the dimension n only.*

Remark 3.8. If we are interested in sets of positive Lebesgue measure only, it would be enough to apply this result with a weaker bound on the number of subcubes with large doubling index, namely B^{n-c}, see the combinatorial lemma in Logunov and Malinnikova [2018], which is simpler.

4 Proof of the Main result

4.1 Reformulations of Theorem 2.1. Clearly Theorem 2.1 is a local result. We formulate an equivalent local version.

Proposition 4.1. *Let Ω be a bounded domain in $\mathbb{R}^n$, A satisfy (8) and δ and m be positive. There exist $C, \gamma > 0$, depending on A, Ω, m and δ such that the following holds. Suppose that u is a solution to $\operatorname{div}(A\nabla u) = 0$ in $\Omega \supset (10n^2)Q$ and let $E \subset \frac{1}{20n}Q$ satisfy $C_{\mathcal{H}}^{n-1+\delta}(E) \geq m s(Q)^{n-1+\delta}$, then*

$$\sup_Q |u| \leq C(\sup_E |u|)^\gamma (\sup_{(10n^2)Q} |u|)^{1-\gamma}.$$

The constants $20n$ and $10n^2$ are for technical purposes only. One can replace them by the constant 2 and the lemma above will remain true.

One can use the standard argument to deduce Theorem 2.1 from Proposition 4.1. We give only a sketch without details. First, find a suitable cube Q with $20n^2 Q \subset\subset \Omega$ and $C_{\mathcal{H}}^{n-1+\delta}(\frac{1}{20n}Q \cap E) > 0$. Second, apply Proposition 4.1. It shows that we can propagate smallness from E onto the cube Q. Third, with the help of the three spheres theorem the standard Harnack chain argument allows to propagate smallness from Q onto the whole $\mathcal{K} \subset\subset \Omega$.

It remains to prove Proposition 4.1, which will follow from the next lemma. All the main ideas of the paper are used in the proof of the lemma, reduction of the proposition to the lemma will be given below and is not difficult.

Lemma 4.2. *Suppose that* $\operatorname{div}(A\nabla u) = 0$ *in* $(20n)Q$ *and* $\sup_Q |u| = 1$. *Let* $N = N(u, Q) \geq 1$. *Set as above*

$$E_a = \{x \in Q : |u(x)| < e^{-a}\}.$$

Then

$$C_{\mathcal{H}}^{n-1+\delta}(E_a) < Ce^{-\beta a/N} s(Q)^{n-1+\delta},$$

for some $C, \beta > 0$ *that depend on* A, δ *only.*

Remark 4.3. This lemma with $\delta = 1$ can be written as a version of Remez inequality (see Remes [1936]) for solutions of $\operatorname{div}(A\nabla u) = 0$ and the role of the degree of a polynomial is played by the doubling index:

$$\sup_Q |u| \leq C \sup_E |u| \left(C\frac{|Q|}{|E|} \right)^{CN}$$

where C depends on A only, E is a subset of Q of a positive measure and $N = N(u, Q)$ is defined by (10). Note that one can also replace the maximal version of the doubling index $N(u, Q)$ by $\log \frac{\sup_{2Q} |u|}{\sup_Q |u|}$ and the statement will remain true. The standard reduction, which we omit, uses the monotonicity property of the doubling index.

4.2 Lemmas 3.5 and 4.2 imply Proposition 4.1.

Consider two cases.

First case: $N = N(u, \frac{1}{10n}Q) \leq 1$. Here Lemma 3.5 is applicable for $\frac{1}{10n}Q$ and since $C_{\mathcal{H}}^{n-1+\delta}(E) > m$ we have

$$\sup_E |u| \geq c_m \sup_{\frac{1}{10n}Q} |u|.$$

And by the three spheres (squares) theorem we know

$$\sup_Q |u| \leq C(\sup_{\frac{1}{10n}Q} |u|)^{\gamma} (\sup_{(10n^2)Q} |u|)^{1-\gamma}.$$

Second case: $N = N(u, \frac{1}{10n}Q) \geq 1$. Assume that $C_{\mathcal{H}}^{n-1+\delta}(E) = ms(Q)^{n-1+\delta} > 0$, $|u| < \varepsilon$ on E and $\sup_{\frac{1}{10n}Q} |u| = 1$. We apply Lemma 4.2 in the cube $\frac{1}{10n}Q$ with $a = |\log \varepsilon|$. Then $E \subset E_a$ and the lemma implies that

$$m < C\varepsilon^{\beta/N}$$

and therefore

$$N \geq \gamma |\log \varepsilon|,$$

where $\gamma = \gamma(C, m, \beta)$.

It is time to use the definition of the doubling index, see Section 2.3. There exists a ball $B_r(x)$, $x \in \frac{1}{10n}Q$, $r \leq \frac{1}{10n}s(Q)$ such that

$$\log \frac{\sup\limits_{B_{10nr}(x)} |u|}{\sup\limits_{B_r(x)} |u|} \geq N - 1/100.$$

Note that $B_{10nr}(x) \subset B_{\sqrt{n}s(Q)}(x)$ and $B_{\sqrt{n}s(Q)}(x)$ also contains Q. Then the monotonicity of the doubling index (9) and the assumption $N \geq 1$ implies

$$\log \frac{\sup\limits_{(10n^2)Q} |u|}{\sup\limits_{Q} |u|} \geq \log \frac{\sup\limits_{B_{10n\sqrt{n}s(Q)}(x)} |u|}{\sup\limits_{B_{\sqrt{n}s(Q)}(x)} |u|} \geq c_1 \log \frac{\sup\limits_{B_{10nr}(x)} |u|}{\sup\limits_{B_r(x)} |u|} \geq c_2 N \geq c_2 \gamma |\log \epsilon|$$

Thus Proposition 4.1 follows. It remains to prove Lemma 4.2.

4.3 Proof of Lemma 4.2. Now, the ellipticity and Lipschitz constants (see (8)) $\Lambda_1 \geq 1$ and $\Lambda_2 > 0$ are fixed parameters and Q_0 is the unit square in $\mathbb{R}^n$. Numbers $N > 1$ and $a > 0$ are variables. Let

$$m(u,a) = C_{\mathcal{H}}^{n-1+\delta}\{x \in Q_0 : |u(x)| < e^{-a} \sup_{Q_0} |u|\},$$

and

$$M(N,a) = \sup_{*} m(u,a),$$

where the supremum is taken over all elliptic operators $\operatorname{div}(A\nabla \cdot)$ and functions u satisfying the following conditions in $20nQ_0$:

(i) $A(x) = [a_{ij}(x)]_{1 \leq i,j \leq n}$ is a symmetric uniformly elliptic matrix with Lipschitz entries satisfying (8),

(ii) u is a solution to $\operatorname{div}(A\nabla u) = 0$ in $20nQ_0$,

(iii) $N(u, Q_0) \leq N$.

Our aim is to show that

$$M(N,a) \leq Ce^{-\beta a/N}. \tag{15}$$

The constant $\beta > 0$ will be chosen later and will not depend on N.

We can always assume that

$$a/N \gg 1$$

by making the constant C sufficiently large. By Lemma 3.5 we can also assume that N is sufficiently large, in particular $N/2 \geq N_0$, where N_0 is the constant from Lemma A.

The proof contains several steps. First, with the help of Lemma A we prove a recursive inequality for $M(N,a)$. Then we show how this inequality implies the exponential bound (15) by a double induction argument on a, N.

Recursive inequality. We show that

$$M(N,a) \leq B^{1-\delta} M(N/2, a - C_1 N \log B) + B^{-\delta-c} M(N, a - C_1 N \log B). \tag{16}$$

The constant C_1 will be specified later; we choose $B = B_0 + 1$ and c from Lemma A.

Fix a solution u to the elliptic equation $\operatorname{div}(A\nabla u) = 0$ with $N(u, Q_0) \leq N$. Divide Q_0 into B^n subcubes q. Lemma A claims that we can partition cubes q into two groups: a group of good cubes with $N(u,q) \leq N/2$ and a group of bad cubes with $N/2 \leq N(u,q) \leq N$ such that the number of all bad cubes is smaller than B^{n-1-c} (and the number of all good cubes is smaller than the total number of cubes B^n). We have

$$m(u,a) \leq \sum_q C_{\mathcal{H}}^{n-1+\delta}(\{x \in q : |u(x)| < e^{-a} \sup_{Q_0} |u|\}).$$

By Equation (11) we see that

$$\sup_q |u| \geq c_1 B^{-C_1 N} \sup_{Q_0} |u|. \tag{17}$$

Since N is sufficiently large, we can forget about c_1 above by increasing C_1. We continue to estimate $m(u,a)$:

$$\begin{aligned} m(u,a) &\leq \sum_q C_{\mathcal{H}}^{n-1+\delta}(\{x \in q : |u(x)| < e^{-a} B^{C_1 N} \sup_q |u|\}) \\ &= \sum_{\text{good } q} + \sum_{\text{bad } q} C_{\mathcal{H}}^{n-1+\delta}(\{x \in q : |u(x)| < e^{-\tilde{a}} \sup_q |u|\}) \end{aligned} \tag{18}$$

where

$$\tilde{a} = a - C_1 N \log B.$$

Now, we estimate each sum individually. If q is a good cube, then

$$C_{\mathcal{H}}^{n-1+\delta}(\{x \in q : |u(x)| < e^{-\tilde{a}} \sup_q |u|\}) \leq B^{-(n-1+\delta)} M(N/2, \tilde{a})$$

Above we used the scaling property of $C_{\mathcal{H}}^{n-1+\delta}$ and the fact that the restriction of u to a cube q corresponds to a solution of another elliptic PDE in the unit cube, the new equation can be written in the divergence form with some coefficient matrix which satisfies the same estimate (8).

Since the total number of good cubes is smaller than B^n

$$\sum_{\text{good } q} \leq B^{1-\delta} M(N/2, \tilde{a})$$

We know that the number of bad cubes q is smaller than B^{n-1-c}. Hence

$$\sum_{\text{bad } q} \leq B^{n-1-c} B^{-(n-1+\delta)} M(N, \tilde{a}) = B^{-c-\delta} M(N, \tilde{a}).$$

Adding the inequalities for bad and good cubes and taking the supremum over u, we obtain the recursive inequality (16) for $M(N, a)$.

Recursive inequality implies exponential bound. We will now prove that

$$M(N, a) \leq C e^{-\beta a/N}. \tag{19}$$

by a double induction on N and a. Without loss of generality we may assume $N = 2^l$, where l is an integer number. Suppose that we know (19) for $N = 2^{l-1}$ and all $a > 0$ and now we wish to establish it for $N = 2^l$. By Lemma 3.5 and Remark 3.6 we may assume l is sufficiently large. So we can say that Lemma 3.5 gives the basis for the induction. For a fixed l we argue by induction on a with step $C_1 N \log B$. Recall that B is a sufficiently large number for which Lemma A holds. We will assume that $a >> N$, namely $a > C_0 N \log B$, where $C_0 > 0$ will be chosen later. For $a < C_0 N \log B$ the inequality is true if C is large enough.

By the induction assumption we have

$$M(N, a - C_1 N \log B) \leq C e^{-\beta a/N + C_1 \beta \log B}$$

and

$$M(N/2, a - C_1 N \log B) \leq C e^{-2\beta a/N + 2C_1 \beta \log B}.$$

Finally, we use the recursive inequality (16) and get

$$M(N, a) \leq C B^{1-\delta} e^{-2\beta a/N + 2C_1 \beta \log B} + C B^{-\delta-c} e^{-\beta a/N + C_1 \beta \log B}.$$

Our goal is to obtain the following inequality

$$B^{1-\delta} e^{-2\beta a/N + 2C_1 \beta \log B} + B^{-\delta-c} e^{-\beta a/N + C_1 \beta \log B} \leq e^{-\beta a/N}$$

for $a/N > C_0 \log B$. Dividing by $e^{-\beta a/N}$ we reduce it to

$$B^{1-\delta+2C_1\beta}e^{-\beta a/N} + B^{-\delta-c+C_1\beta} \leq 1.$$

Now, recall that $a/N > C_0 \log B$ and the above inequality follows from

$$B^{1-\delta+2C_1\beta-C_0\beta} + B^{-\delta-c+C_1\beta} \leq 1.$$

The last inequality can be achieved with the proper choice of the parameters: $B > 2, \delta, c, C_1 > 0$ are fixed, we choose β to be small enough so that the second term is less than $1-\varepsilon$ and then choose large C_0 to make the first term smaller than ε. Thus the inequality above holds for all sufficiently large a/N. As we mentioned above, for small a/N the inequality (19) is true if we choose C to be large.

Remark 4.4. One can notice that the induction step is working for negative δ such that $-c < \delta$. However the induction basis step (Lemma 3.5) is not true for negative δ. For instance, zeroes of harmonic functions in $\mathbb{R}^n$ are sets of dimension $n-1$. But the induction basis step appears to be true for the gradients of solutions, which have better unique continuation properties than the solutions.

5 Propagation of smallness for the gradients of solutions

5.1 Formulation of the result. As above we assume that u is a solution of an elliptic equation (7) in divergence form in a bounded domain $\Omega \subset \mathbb{R}^n$ and the coefficients satisfy (8).

Theorem 5.1. *There exists a constant $c \in (0,1)$ that depends only on the dimension n such that the following holds. Let m, δ, ρ be positive numbers and suppose sets $E, \mathcal{K} \subset \Omega$ satisfy*

$$C_{\mathcal{H}}^{n-1-c+\delta}(E) > m, \quad \operatorname{dist}(E, \partial\Omega) > \rho, \quad \operatorname{dist}(\mathcal{K}, \partial\Omega) > \rho.$$

Then there exist $C, \gamma > 0$, depending on $m, \delta, \rho, \Lambda_1, \Lambda_2, \Omega$ only (and independent of u) such that

$$\sup_{\mathcal{K}} |\nabla u| \leq C(\sup_E |\nabla u|)^\gamma (\sup_\Omega |\nabla u|)^{1-\gamma}.$$

5.2 Modifications of the proof. We shall use the notion of doubling index for $|\nabla u|$. Let $B = B(x_0, r)$ be a ball in $\mathbb{R}^n$. Define

$$N(\nabla u, B) = \log \frac{\sup_{2B} |\nabla u|}{\sup_B |\nabla u|}.$$

Assume $r \le 1$. The doubling index is almost monotonic:

$$N(\nabla u, tB) \le N(\nabla u, B)(1+c) + C \tag{20}$$

for $t \le 1/2$. The constants $c, C > 0$ depend on Λ_1, Λ_2 (the ellipticity and Lipschitz constants) and the dimension n. The monotonicity of the doubling index for $|\nabla u|$ follows from the three spheres theorem for the function $|u(\cdot) - u(x_0)|$ and standard elliptic estimates. A similar modification appeared in Cheeger, Naber, and Valtorta [2015], see also Garofalo and Lin [1986]. We also need a modified doubling index for a cube Q:

$$N(\nabla u, Q) = \sup_{x \in Q, r \le s(Q)} \log \frac{\sup_{B(x,10nr)} |\nabla u|}{\sup_{B(x,r)} |\nabla u|}.$$

The proof of Theorem 5.1 is parallel to the proof of Theorem 2.1. We need to establish analogs of Lemma A (induction step), Lemma 3.5 (basis of induction), and Lemma 4.2 (estimate of the Hausdorff content), where $|u|$ should be replaced by $|\nabla u|$. We formulate such statements below.

Lemma B. *There exist positive constants s_0, N_0, B_0 that depend on Λ_1, Λ_2 and the dimension n only such that if Q is a cube with side $s(Q) < s_0$ and Q is divided into B^n equal subcubes with $B > B_0$, then the number of subcubes q with $N(\nabla u, q) \ge \max(\frac{1}{2} N(\nabla u, Q), N_0)$ is less than B^{n-1-c}, where $c \in (0,1)$ depends on the dimension n only.*

Lemma 5.2. *Let Q_0 be the unit cube in $\mathbb{R}^n$. Suppose that* $\operatorname{div}(A\nabla u) = 0$ *in $(20n)Q_0$,* $\sup_{Q_0} |\nabla u| = 1$*, and $N(\nabla u, Q_0) \le N_0$, then for*

$$E_a = \{x \in Q_0 : |\nabla u(x)| < e^{-a}\}$$

we have

$$C_{\mathcal{H}}^{n-2+\delta}(E_a) < Ce^{-\beta a}$$

for some β, C depending on N_0, Λ_1, Λ_2, δ.

Lemma 5.3. *Let Q_0 be the unit cube in $\mathbb{R}^n$. Suppose that* $\operatorname{div}(A\nabla u) = 0$ *in $(20n)Q_0$ and* $\sup_{Q_0} |\nabla u| = 1$*. Let a number $N = N(\nabla u, Q_0) \ge 1$. Set*

$$E_a = \{x \in Q_0 : |\nabla u(x)| < e^{-a}\}.$$

There exists $c \in (0,1)$ that depends only on the dimension n such that

$$C_{\mathcal{H}}^{n-1-c+\delta}(E_a) < Ce^{-\beta a/N},$$

for some $C, \beta > 0$ that depend on $\Lambda_1, \Lambda_2, \delta, n$ only.

Only the proof of Lemma 5.2 requires modifications, the other changes are minor.

5.3 Outline of changes. The reduction of Theorem 5.1 to Lemma 5.3 is not difficult and remains the same as in Section 4. To prove Lemma 5.3 one has to replace the used Lemma 3.5 by its analog for $|\nabla u|$ (Lemma 5.2), which we prove below. The proof is based on new results from Cheeger, Naber, and Valtorta [2015].

The proof of Lemma B repeats the proof of Lemma A (Logunov [2018b]). There are two main ingredients in the proof: simplex lemma and hyperplane lemma from Logunov [2018a]. We don't formulate those lemmas here, see Logunov [ibid.]. There are no changes in the proof of hyperplane lemma, except that one has to subtract a constant from the function.

To prove the simplex lemma we need a sharper version of the monotonicity property of the doubling index as it was in the proof of the original simplex lemma. Namely, one has to make c in inequality (20) a sufficiently small constant, depending only on the dimension. One has to make a linear change of coordinates such that $A(0)$ turns into the identity matrix and Λ_1 is close to 1 in a small neighborhood of the origin. After that one obtains a sharper version of the three spheres theorem for $u - u(0)$ as it is done in Logunov [ibid.]. Then one should use standard elliptic estimates to provide smallness of c in (20).

To prove Lemma 5.3 one has to use the same induction argument as in Lemma 4.2. The induction step remains the same, but one has to work with the doubling index for $|\nabla u|$ in place of $|u|$ and use Lemma B in place of Lemma A. Concerning the basis of induction, which is Lemma 5.2, a different argument is needed, and we will use a result from Cheeger, Naber, and Valtorta [2015], which estimates the size of the neighborhood of the effective critical set. That would give us an analog of Lemma 3.5 for $|\nabla u|$, but now the dimension of the set E will be allowed to be smaller than $n-1$, but bigger than $n-2$. Unfortunately, the induction step works only for dimensions bigger $n-1-c$ only and that is the main obstacle for improvement towards $n-2$.

5.4 Proof of Lemma 5.2. The lemma is a corollary from Theorem 1.17 (estimate of the effective critical set) from Cheeger, Naber, and Valtorta [ibid.]. We warn the reader that we formulate it below in our own notation and don't bring the proof of Theorem 1.17.

Reformulation of Theorem 1.17 from Cheeger, Naber, and Valtorta [ibid.]. Let u be as in Lemma 5.2. For any $\delta > 0$ there exist positive constants C and c depending on $n, \Lambda_1, \Lambda_2, \delta, N_0$ such that the following holds for all integer K. Partition the unit cube Q_0 into K^n sub-cubes q with side length $1/K$. We call q bad if

$$\inf_q |\nabla u| < c \sup_{2q} |\nabla u|.$$

Then the number of bad cubes q is not greater than $CK^{n-2+\delta}$.

Now, we are ready to finish the proof of Lemma 5.2. We divide the unit cube Q_0 into K^n sub-cubes q with side length $1/K$, the integer K will be chosen later.

The monotonicity of the doubling index implies

$$\sup_q |\nabla u| \geq c_1 K^{-C_1 N_0 - C_1} \sup_{Q_0} |\nabla u| = c_1 K^{-C_1 N_0 - C_1}.$$

If q is not bad, then

$$\inf_q |\nabla u| \geq c_2 K^{-C_1 N_0 - C_1}.$$

Given $a > 0$ we want to estimate the Hausdorff content $C_{\mathcal{H}}^{n-2+2\delta}$ of

$$E_a = \{x \in Q_0 : |\nabla u(x)| < e^{-a}\}.$$

We may assume $a > 1$. Now, we specify the choice of K. The K is smallest integer number greater than 2 such that

$$e^{-a} > c_2 K^{-C_1 N_0 - C_1}.$$

So $\log K$ is comparable to a. And the set E_a is contained in the union of bad cubes of size $1/K$. The number of bad cubes is not greater than $CK^{n-2+\delta}$. We therefore have

$$C_{\mathcal{H}}^{n-2+2\delta}(E_a) \leq C_2 K^{\delta} \leq C_3 e^{-c_3 a}.$$

Replacing 2δ by δ we finish the proof.

References

John Apraiz, Luis Escauriaza, Gengsheng Wang, and Can Zhang (2014). "Observability inequalities and measurable sets". *J. Eur. Math. Soc. (JEMS)* 16.11, pp. 2433–2475. MR: 3283402 (cit. on p. 2410).

Jean Bourgain and Semyon Dyatlov (n.d.). "Spectral gaps without the pressure condition". arXiv: 1612.09040 (cit. on p. 2414).

Jean Bourgain and Zeév Rudnick (2009). "Restriction of toral eigenfunctions to hypersurfaces". *C. R. Math. Acad. Sci. Paris* 347.21-22, pp. 1249–1253. MR: 2561032 (cit. on p. 2414).

Jean Bourgain and Thomas Wolff (1990). "A remark on gradients of harmonic functions in dimension ≥ 3". *Colloq. Math.* 60/61.1, pp. 253–260. MR: 1096375 (cit. on p. 2413).

Italo Capuzzo Dolcetta (2002). "Some properties of solutions of fully nonlinear partial differential inequalities". In: *Nonlinear problems in mathematical physics and related topics, II*. Vol. 2. Int. Math. Ser. (N. Y.) Kluwer/Plenum, New York, pp. 75–88. MR: 1971991 (cit. on p. 2410).

Jeff Cheeger, Aaron Naber, and Daniele Valtorta (2015). "Critical sets of elliptic equations". *Comm. Pure Appl. Math.* 68.2, pp. 173–209. MR: 3298662 (cit. on pp. 2412, 2426, 2427).

Harold Donnelly and Charles Fefferman (1988). "Nodal sets of eigenfunctions on Riemannian manifolds". *Invent. Math.* 93.1, pp. 161–183. MR: 943927 (cit. on pp. 2414, 2417).

– (1990). "Growth and geometry of eigenfunctions of the Laplacian". In: *Analysis and partial differential equations*. Vol. 122. Lecture Notes in Pure and Appl. Math. Dekker, New York, pp. 635–655. MR: 1044811 (cit. on p. 2414).

Semyon Dyatlov and Long Jin (n.d.). "Semiclassical measures on hyperbolic surfaces have full support". arXiv: 1705.05019 (cit. on pp. 2414, 2415).

Nicola Garofalo and Fang-Hua Lin (1986). "Monotonicity properties of variational integrals, A_p weights and unique continuation". *Indiana Univ. Math. J.* 35.2, pp. 245–268. MR: 833393 (cit. on pp. 2411, 2416, 2426).

Qing Han and Fang-Hua Lin (n.d.). "Nodal sets of solutions of elliptic differential equations" (cit. on p. 2416).

Robert Hardt and Leon Simon (1989). "Nodal sets for solutions of elliptic equations". *J. Differential Geom.* 30.2, pp. 505–522. MR: 1010169 (cit. on p. 2417).

Evgenii M. Landis (1963). "Some questions in the qualitative theory of second-order elliptic equations (case of several independent variables)". *Uspehi Mat. Nauk* 18.1 (109), pp. 3–62. MR: 0150437 (cit. on p. 2416).

Fang-Hua Lin (1991). "Nodal sets of solutions of elliptic and parabolic equations". *Comm. Pure Appl. Math.* 44.3, pp. 287–308. MR: 1090434 (cit. on pp. 2412, 2413).

Elon Lindenstrauss (2006). "Invariant measures and arithmetic quantum unique ergodicity". *Ann. of Math. (2)* 163.1, pp. 165–219. MR: 2195133 (cit. on p. 2415).

Alexander Logunov (2018a). "Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure". *Ann. of Math. (2)* 187.1, pp. 221–239. arXiv: 1605.02587. MR: 3739231 (cit. on pp. 2409, 2411, 2412, 2416, 2417, 2419, 2427).

– (2018b). "Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture". *Ann. of Math. (2)* 187.1, pp. 241–262. arXiv: 1605.02589. MR: 3739232 (cit. on pp. 2409, 2411, 2412, 2418–2420, 2427).

Alexander Logunov and Eugenia Malinnikova (2018). "Nodal Sets of Laplace Eigenfunctions: Estimates of the Hausdorff Measure in Dimensions Two and Three". In: *50 Years with Hardy Spaces. Operator Theory: Advances and Applications*. Ed. by Baranov A., Kisliakov S., and Nikolski N. Vol. 261, pp. 333–344. arXiv: 1605.02595 (cit. on pp. 2409, 2411, 2412, 2418, 2420).

Eugenia Malinnikova (2004). "Propagation of smallness for solutions of generalized Cauchy-Riemann systems". *Proc. Edinb. Math. Soc. (2)* 47.1, pp. 191–204. MR: 2064745 (cit. on p. 2411).

Eugenia Malinnikova and Sergio Vessella (2012). "Quantitative uniqueness for elliptic equations with singular lower order terms". *Math. Ann.* 353.4, pp. 1157–1181. MR: 2944026 (cit. on p. 2411).

Dan Mangoubi (2013). "The effect of curvature on convexity properties of harmonic functions and eigenfunctions". *J. Lond. Math. Soc. (2)* 87.3, pp. 645–662. MR: 3073669 (cit. on p. 2416).

Aaron Naber and Daniel Valtorta (2017). "Estimates on the Critical Sets of Solutions to Elliptic PDEs". *Comm. Pure Appl. Math.* 70, pp. 1835–1897. MR: 3688031 (cit. on pp. 2412, 2413).

Nikolai S. Nadirašvili (1979). "Estimation of the solutions of elliptic equations with analytic coefficients which are bounded on some set". *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* 2, pp. 42–46, 102. MR: 531646 (cit. on p. 2411).

– (1986). "Uniqueness and stability of continuation from a set to the domain of solution of an elliptic equation". *Mat. Zametki* 40.2, pp. 218–225, 287. MR: 864285 (cit. on p. 2411).

Evgenii Ja. Remes (1936). "Sur une propriété extrémale des polynômes de Tchebychef". *Commun. Inst. Sci. Math. Mech., Univ. Kharkoff* 13, pp. 93–95 (cit. on p. 2421).

Peter Sarnak (2011). "Recent progress on the quantum unique ergodicity conjecture". *Bull. Amer. Math. Soc. (N.S.)* 48.2, pp. 211–228. MR: 2774090 (cit. on p. 2415).

Elena G. Sitnikova (1970). "A theorem on three balls for an elliptic equation of higher order". *Mat Sbornik (N. S.)* 82.124, pp. 213–219. MR: 0265758 (cit. on p. 2410).

Alexander I. Šnirelman (1974). "Ergodic properties of eigenfunctions". *Uspehi Mat. Nauk* 29.6(180), pp. 181–182. MR: 0402834 (cit. on p. 2415).

Sergio Vessella (1999). "A continuous dependence result in the analytic continuation problem". *Forum Math.* 11.6, pp. 695–703. MR: 1724631 (cit. on p. 2411).

– (2000). "Quantitative continuation from a measurable set of solutions of elliptic equations". *Proc. Roy. Soc. Edinburgh Sect. A* 130.4, pp. 909–923. MR: 1776684 (cit. on p. 2411).

Steven Zelditch (1987). "Uniform distribution of eigenfunctions on compact hyperbolic surfaces". *Duke Math. J.* 55.4, pp. 919–941. MR: 916129 (cit. on p. 2415).

– (2010). "Recent developments in mathematical quantum chaos". In: *Current developments in mathematics, 2009*. Int. Press, Somerville, MA, pp. 115–204. MR: 2757360 (cit. on p. 2415).

Maciej Zworski (2012). *Semiclassical analysis*. Vol. 138. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xii+431. MR: 2952218 (cit. on p. 2415).

Received 2017-11-28.

ALEXANDER LOGUNOV

School of Mathematical Sciences
Tel Aviv University
Tel Aviv 69978
Israel

and

Chebyshev Laboratory
St. Petersburg State University 14th Line V.O., 29B
Saint Petersburg, 199178
Russia

and

Institute for Advanced Study Princeton, NJ, 08540
USA
log239@yandex.ru

Eugenia Malinnikova
Department of Mathematics
Purdue University
150 N. University Street
West Lafayette, IN 47907-2067

and

Department of Mathematical Sciences
Norwegian University of Science and Technology
7491, Trondheim
Norway
eugenia.malinnikova@ntnu.no

ICM 2018 RIO DE JANEIRO

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2431–2456)

WELL-POSEDNESS, GLOBAL EXISTENCE AND DECAY ESTIMATES FOR THE HEAT EQUATION WITH GENERAL POWER-EXPONENTIAL NONLINEARITIES

Mohamed Majdoub and Slim Tayachi

Abstract

In this paper we consider the problem: $\partial_t u - \Delta u = f(u),\ u(0) = u_0 \in \exp L^p(\mathbb{R}^N)$, where $p > 1$ and $f : \mathbb{R} \to \mathbb{R}$ having an exponential growth at infinity with $f(0) = 0$. We prove local well-posedness in $\exp L^p_0(\mathbb{R}^N)$ for $f(u) \sim \mathrm{e}^{|u|^q},\ 0 < q \le p,\ |u| \to \infty$. However, if for some $\lambda > 0$, $\liminf_{s\to\infty}\left(f(s)\,\mathrm{e}^{-\lambda s^p}\right) > 0$, then non-existence occurs in $\exp L^p(\mathbb{R}^N)$. Under smallness condition on the initial data and for exponential nonlinearity f such that $|f(u)| \sim |u|^m$ as $u \to 0$, $\frac{N(m-1)}{2} \ge p$, we show that the solution is global. In particular, $p - 1 > 0$ sufficiently small is allowed. Moreover, we obtain decay estimates in Lebesgue spaces for large time which depend on m.

1 Introduction

In this paper we study the Cauchy problem:

$$\text{(1-1)} \qquad \begin{cases} \partial_t u - \Delta u = f(u), \\ u(0) = u_0 \in \exp L^p(\mathbb{R}^N), \end{cases}$$

where $p > 1$ and $f : \mathbb{R} \to \mathbb{R}$ having an exponential growth at infinity with $f(0) = 0$.

As is a standard practice, we study (1-1) via the associated integral equation:

$$\text{(1-2)} \qquad u(t) = \mathrm{e}^{t\Delta} u_0 + \int_0^t \mathrm{e}^{(t-s)\Delta}\, f(u(s))\, ds,$$

where $\mathrm{e}^{t\Delta}$ is the linear heat semi-group. The Cauchy problem (1-1) has been extensively studied in the scale of Lebesgue spaces, especially for polynomial type nonlinearities. It is

MSC2010: primary 35K58; secondary 35K15, 46E30, 35A01.
Keywords: Nonlinear heat equation, Existence, Non-existence, Orlicz spaces.

known that in this case one can always find a Lebesgue space L^q, $q < \infty$ for which (1-1) is locally well-posed. See for instance Brezis and Cazenave [1996], Haraux and Weissler [1982], and Weissler [1979, 1980].

By analogy with the Lebesgue spaces, which are well-adapted to the heat equations with power nonlinearities (Weissler [1981]), we are motivated to consider the Orlicz spaces, in order to study heat equations with power-exponential nonlinearities. Such spaces were introduced by Birnbaum and Orlicz [1931] as a natural generalization of the classical Lebesgue spaces L^q, $1 < q < \infty$. For this generalization the function x^q entering in the definition of L^q space is replaced by a more general convex function: in particular $e^{x^q} - 1$.

For the particular case where $f(u) \sim e^{|u|^2}$, u large, well-posedness results are proved in the Orlicz space $\exp L^2(\mathbb{R}^N)$. See Ibrahim, Jrad, Majdoub, and Saanouni [2014], Ioku [2011], Ioku, Ruf, and Terraneo [2015], and Ruf and Terraneo [2002]. It is also proved that if $f(u) \sim e^{|u|^s}$, $s > 2$, u large then the existence is no longer guaranteed and in fact there is nonexistence in the Orlicz space $\exp L^2(\mathbb{R}^N)$. See Ioku, Ruf, and Terraneo [2015]. Global existence and decay estimates are also established for the nonlinear heat equation with $f(u) \sim e^{|u|^2}$, u large. See Ioku [2011], Majdoub, Otsmane, and Tayachi [2018], and Furioli, Kawakami, Ruf, and Terraneo [2017].

Here we consider the general case $f(u) \sim e^{|u|^q}$, $q > 1$, u large. For such exponential nonlinearities, the most adaptable space is the so-called Orlicz space $\exp L^p(\mathbb{R}^N)$, $p \geq q > 1$. We aim to study local well-posedness and look for the maximum power of the nonlinearity in terms of the existence of solutions in these spaces. We also study the global existence for small initial data and determine the decay estimates for large time. For the global existence, we aim to allow f to behave like $|u|^{m-1}u$ near the origin, with $m > 1 + 2/N$. That is to reach the Fujita critical exponent $1 + 2/N$.

The Orlicz space $\exp L^p(\mathbb{R}^N)$ is defined as follows

$$\exp L^p(\mathbb{R}^N) = \left\{ u \in L^1_{loc}(\mathbb{R}^N);\ \int_{\mathbb{R}^N} \left(e^{\frac{|u(x)|^p}{\lambda^p}} - 1 \right) dx < \infty, \ \text{ for some } \lambda > 0 \right\},$$

endowed with the Luxembourg norm

$$\|u\|_{\exp L^p(\mathbb{R}^N)} := \inf\left\{ \lambda > 0;\ \ \int_{\mathbb{R}^N} \left(e^{\frac{|u(x)|^p}{\lambda^p}} - 1 \right) dx \leq 1 \right\}.$$

Since the space of smooth compactly supported functions $C_0^\infty(\mathbb{R}^N)$ is not dense in the Orlicz space $\exp L^p(\mathbb{R}^N)$ (see Ioku, Ruf, and Terraneo [2015] and Ioku [2011]), we use the space $\exp L^p_0(\mathbb{R}^N)$ which is the closure of $C_0^\infty(\mathbb{R}^N)$ with respect to the Luxemburg norm $\|\cdot\|_{\exp L^p(\mathbb{R}^N)}$. It is known that Ioku, Ruf, and Terraneo [2015]

(1-3)
$$\exp L^p_0(\mathbb{R}^N) = \left\{ u \in L^1_{loc}(\mathbb{R}^N);\ \int_{\mathbb{R}^N} \left(\mathrm{e}^{\alpha|u(x)|^p} - 1\right) dx < \infty,\ \text{ for every } \alpha > 0 \right\}.$$

It is easy to show that the linear heat semi-group $\mathrm{e}^{t\Delta}$ is continuous at $t = 0$ in $\exp L^p_0(\mathbb{R}^N)$. However, this is not the case in $\exp L^p(\mathbb{R}^N)$.

In the sequel, we adopt the following definitions of weak, weak-mild and classical solutions to Cauchy problem (1-1).

Definition 1.1 (Weak solution)**.** *Let $u_0 \in \exp L^p_0(\mathbb{R}^N)$ and $T > 0$. We say that the function $u \in C([0, T];\ \exp L^p_0(\mathbb{R}^N))$ is a weak solution of (1-1) if u verifies (1-1) in the sense of distribution and $u(t) \to u_0$ in the weak* topology as $t \searrow 0$.*

Definition 1.2 (Weak-mild solution)**.** *We say that $u \in L^\infty(0, T;\ \exp L^p(\mathbb{R}^N))$ is a weak-mild solution of the Cauchy problem (1-1) if u satisfies the associated integral equation (1-2) in $\exp L^p(\mathbb{R}^N)$ for almost all $t \in (0, T)$ and $u(t) \to u_0$ in the weak* topology as $t \searrow 0$.*

Definition 1.3 ($\exp L^p$−classical solution)**.** *Let $u_0 \in \exp L^p(\mathbb{R}^N)$ and $T > 0$. A function $u \in C((0, T]; \exp L^p(\mathbb{R}^N)) \cap L^\infty_{loc}(0, T; L^\infty(\mathbb{R}^N))$ is said to be $\exp L^p$−classical solution of (1-1) if $u \in C^{1,2}((0, T) \times \mathbb{R}^N)$, verifies (1-1) in the classical sense and $u(t) \to u_0$ in the weak* topology as $t \searrow 0$.*

We are first interested in the local well-posedness. Since $C^\infty_0(\mathbb{R}^N)$ is dense in $\exp L^p_0(\mathbb{R}^N)$, we are able to prove local existence and uniqueness to (1-1) for initial data in $\exp L^p_0(\mathbb{R}^N)$. We assume that the nonlinearity f satisfies

$$\text{(1-4)}\qquad f(0) = 0, \qquad |f(u) - f(v)| \le C|u - v|(\mathrm{e}^{\lambda|u|^p} + \mathrm{e}^{\lambda|v|^p}),\ \forall\ u, v \in \mathbb{R},$$

for some constants $C > 0$, $p > 1$ and $\lambda > 0$. Our first main result reads as follows.

Theorem 1.4 (Local well-posedness)**.** *Suppose that f satisfies (1-4). Given any $u_0 \in \exp L^p_0(\mathbb{R}^N)$ with $p > 1$, there exist a time $T = T(u_0) > 0$ and a unique weak solution $u \in C([0, T];\ \exp L^p_0(\mathbb{R}^N))$ to (1-1).*

We stress that the density of $C^\infty_0(\mathbb{R}^N)$ in $\exp L^p_0(\mathbb{R}^N)$ is crucial in the above Theorem. In fact we have obtained the following non-existence result in $\exp L^p(\mathbb{R}^N)$.

Theorem 1.5 (Non-existence)**.** *Let $p > 1$, $\alpha > 0$ and*

$$\text{(1-5)}\qquad \Phi_\alpha(x) = \begin{cases} \alpha\Big(-\log|x|\Big)^{\frac{1}{p}}, & |x| < 1, \\ 0, & |x| \ge 1. \end{cases}$$

Assume that $f:\mathbb{R}\to\mathbb{R}$ *is continuous, positive on* $[0,\infty)$ *and satisfies*

$$\liminf_{s\to\infty}\Big(f(s)\,\mathrm{e}^{-\lambda s^p}\Big)>0,\quad \lambda>0. \tag{1-6}$$

Then $\Phi_\alpha\in\exp L^p(\mathbb{R}^N)\backslash\exp L^p_0(\mathbb{R}^N)$ *and there exists* $\alpha_0>0$ *such that for every* $\alpha\geq\alpha_0$ *and* $T>0$ *the Cauchy problem* (1-1) *with* $u_0=\Phi_\alpha$ *has no nonnegative* $\exp L^p$*−classical solution in* $[0,T]$.

The results of Theorems 1.4-1.5 are known for $p=2$ in Ioku, Ruf, and Terraneo [2015].

Our next interest is the global existence and the decay estimate. It depends on the behavior of the nonlinearity $f(u)$ near $u=0$. The following behavior near 0 will be allowed

$$|f(u)|\sim|u|^m,$$

where $\frac{N(m-1)}{2}\geq p$. More precisely, we suppose that the nonlinearity f satisfies

$$f(0)=0,\ |f(u)-f(v)|\leq C\,|u-v|\left(|u|^{m-1}\mathrm{e}^{\lambda|u|^p}+|v|^{m-1}\mathrm{e}^{\lambda|v|^p}\right),\quad\forall u,v\in\mathbb{R}, \tag{1-7}$$

where $\frac{N(m-1)}{2}\geq p>1$, $C>0$, and $\lambda>0$ are constants. Our aim is to obtain global existence to the Cauchy problem (1-1) for small initial data in $\exp L^p(\mathbb{R}^N)$. We have obtained the following.

Theorem 1.6 (Global existence). *Let* $N\geq1$, $p>1$, *such that* $N(p-1)/2>p$. *Assume that* $m\geq p$ *(hence* $N(m-1)/2>p$*) and the nonlinearity* f *satisfies* (1-7). *Then, there exists a positive constant* $\varepsilon>0$ *such that every initial data* $u_0\in\exp L^p(\mathbb{R}^N)$ *with* $\|u_0\|_{\exp L^p(\mathbb{R}^N)}\leqslant\varepsilon$, *there exists a weak-mild solution* $u\in L^\infty(0,\infty;\exp L^p(\mathbb{R}^N))$ *of the Cauchy problem* (1-1) *satisfying*

$$\lim_{t\longrightarrow0}\|u(t)-\mathrm{e}^{t\Delta}u_0\|_{\exp L^p(\mathbb{R}^N)}=0. \tag{1-8}$$

Moreover, if $m>3/2$ *then there exists a constant* $C>0$ *such that,*

$$\|u(t)\|_a\leq C\,t^{-\sigma},\quad\forall\,t>0, \tag{1-9}$$

where

$$\frac{N(m-1)}{2}<a<\frac{N(m-1)}{2}\,\frac{1}{(2-m)_+},\ a>N/2,\quad\textit{and}\quad\sigma=\frac{1}{m-1}-\frac{N}{2a}>0\,.$$

Remarks 1.7.

(i) The case $N(p-1)/2 \leq p$ will be investigated in a forthcoming paper.

(ii) Note that in the proof of the decay estimates, we require $a > N/2$ which is compatible with the other assumptions only if we impose the additional condition $m > 3/2$.

(iii) If only we want to prove global existence, we change the space of contraction that is we omit the Lebesgue part and we do not need such a supplementary condition on m.

Hereafter, $\|\cdot\|_r$ denotes the norm in the Lebesgue space $L^r(\mathbb{R}^N),\ 1 \leq r \leq \infty$. We mention that the assumption for the nonlinearity covers the cases

$$f(u) = \pm|u|^{m-1}u\mathrm{e}^{|u|^p}, \quad m \geq 1 + \frac{2p}{N}.$$

The global existence part of Theorem 1.6 is known for $p = 2$ (see Ioku [2011]). The estimate (1-9) was obtained in Ioku [ibid.] for $p = 2$ and $m = 1 + \frac{4}{N}$. This is improved in Majdoub, Otsmane, and Tayachi [2018] for $p = 2$ and any $m \geq 1 + \frac{4}{N}$. The fact that estimate (1-9) depends on the smallest power of the nonlinearity $f(u)$ is known in Snoussi, Tayachi, and Weissler [2001] but only for nonlinearities having polynomial growth.

Using similar arguments as in Weissler [1980], we can show the following lower estimate of the blow-up rate.

Theorem 1.8 (Blow-up rate). *Assume that the nonlinearity f satisfies (1-4) with $\lambda > 0$. Let $u_0 \in L^p(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$ and $u \in C([0, T_{\max}); L^p(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N))$ be the maximal solution of (1-1). If $T_{\max} < \infty$, then there exist two positive constants $C_1,\ C_2$ such that*

$$\lambda \|u(t)\|^p_{L^p(\mathbb{R}^N)\cap L^\infty(\mathbb{R}^N)} \geq C_1 \big| \log(T_{\max} - t)\big| + C_2, \quad 0 \leq t < T_{\max}.$$

See Souplet and Tayachi [2016] and references therein for similar blow-up estimates for parabolic problems with exponential nonlinearities.

The rest of this paper is organized as follows. In the next section, we collect some basic facts and useful tools about Orlicz spaces. Section 3 is devoted to some crucial estimates on the linear heat semi-group. The sketches of the proofs of Theorems 1.4 and 1.8 are done in Section 4. Section 5 is devoted to Theorem 1.5 about nonexistence. Finally, in Section 6 we give the proof of Theorem 1.6. In all this paper, C will be a positive constant which may have different values at different places. Also, $L^r(\mathbb{R}^N)$, $\exp L^r(\mathbb{R}^N)$, $\exp L^r_0(\mathbb{R}^N)$ will be written respectively L^r, $\exp L^r$ and $\exp L^r_0$.

2 Orlicz spaces: basic facts and useful tools

Let us recall the definition of the so-called Orlicz spaces on $\mathbb{R}^N$ and some related basic facts. For a complete presentation and more details, we refer the reader to Adams and Fournier [2003], Rao and Ren [2002], and Trudinger [1967].

Definition 2.1.
Let $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ be a convex increasing function such that

$$\phi(0) = 0 = \lim_{s\to 0^+} \phi(s), \qquad \lim_{s\to\infty} \phi(s) = \infty.$$

We say that a function $u \in L^1_{loc}(\mathbb{R}^N)$ belongs to $L^\phi(\mathbb{R}^N)$ if there exists $\lambda > 0$ such that

$$\int_{\mathbb{R}^N} \phi\left(\frac{|u(x)|}{\lambda}\right) dx < \infty.$$

We denote then

$$\|u\|_{L^\phi} = \inf\left\{\lambda > 0, \quad \int_{\mathbb{R}^N} \phi\left(\frac{|u(x)|}{\lambda}\right) dx \le 1\right\}. \tag{2-1}$$

It is known that $\left(L^\phi(\mathbb{R}^N), \|\cdot\|_{L^\phi}\right)$ is a Banach space. Note that, if $\phi(s) = s^p$, $1 \le p < \infty$, then L^ϕ is nothing else than the Lebesgue space L^p. Moreover, for $u \in L^\phi$ with $K := \|u\|_{L^\phi} > 0$, we have

$$\left\{\lambda > 0, \quad \int_{\mathbb{R}^N} \phi\left(\frac{|u(x)|}{\lambda}\right) dx \le 1\right\} = [K, \infty[.$$

In particular

$$\int_{\mathbb{R}^N} \phi\left(\frac{|u(x)|}{\|u\|_{L^\phi}}\right) dx \le 1. \tag{2-2}$$

We also recall the following well known properties.

Proposition 2.2. *We have*

(i) $L^1 \cap L^\infty \subset L^\phi(\mathbb{R}^N) \subset L^1 + L^\infty$.

(ii) Lower semi-continuity:

$$u_n \to u \quad a.e. \quad \Longrightarrow \quad \|u\|_{L^\phi} \le \liminf \|u_n\|_{L^\phi}.$$

(iii) Monotonicity:

$$|u| \le |v| \quad a.e. \quad \Longrightarrow \quad \|u\|_{L^\phi} \le \|v\|_{L^\phi}.$$

(iv) *Strong Fatou property:*

$$0 \le u_n \nearrow u \quad a.e. \quad \Longrightarrow \quad \|u_n\|_{L^\phi} \nearrow \|u\|_{L^\phi}.$$

(v) *Strong and modular convergence:*

$$u_n \to u \;\; in \;\; L^\phi \;\; \Longrightarrow \;\; \int_{\mathbb{R}^N} \phi(u_n - u)dx \to 0.$$

Denote by

$$L_0^\phi(\mathbb{R}^N) = \left\{u \in L^1_{loc}(\mathbb{R}^N), \quad \int_{\mathbb{R}^N} \phi\left(\frac{|u(x)|}{\lambda}\right) dx < \infty, \;\forall\;\; \lambda > 0\right\}.$$

It can be shown (see for example Ioku, Ruf, and Terraneo [2015]) that

$$L_0^\phi(\mathbb{R}^N) = \overline{C_0^\infty(\mathbb{R}^N)}^{L^\phi} = \text{the colsure of } C_0^\infty(\mathbb{R}^N) \text{ in } L^\phi(\mathbb{R}^N).$$

Clearly $L_0^\phi(\mathbb{R}^N) = L^\phi(\mathbb{R}^N)$ for $\phi(s) = s^p,\; p \ge 1$, but this is not the case for any ϕ (see Ioku, Ruf, and Terraneo [ibid.]). When $\phi(s) = \mathrm{e}^{s^p} - 1$, we denote the space $L^\phi(\mathbb{R}^N)$ by $\exp L^p$ and $L_0^\phi(\mathbb{R}^N)$ by $\exp L_0^p$.

The following Lemma summarize the relationship between Orlicz and Lebesgue spaces.

Lemma 2.3. *We have*

(i) $\exp L_0^p \subsetneq \exp L^p,\;\; p \ge 1.$

(ii) $\exp L_0^p \not\hookrightarrow L^\infty$, *hence* $\exp L^p \not\hookrightarrow L^\infty, \quad p \ge 1.$

(iii) $\exp L^p \not\hookrightarrow L^r$, *for all* $1 \le r < p, \quad p > 1.$

(iv) $L^q \cap L^\infty \hookrightarrow \exp L_0^p$, *for all* $1 \le q \le p$. *More precisely*

$$\|u\|_{\exp L^p} \le \frac{1}{(\log 2)^{\frac{1}{p}}}\Big(\|u\|_q + \|u\|_\infty\Big). \tag{2-3}$$

Proof of Lemma 2.3. (i) Let u be the function defined by

$$u(x) = \Big(-\log|x|\Big)^{1/p} \quad \text{if} \quad |x| \le 1,$$
$$u(x) = 0 \quad \text{if} \quad |x| > 1.$$

For $\alpha > 0$, we have

$$\int_{\mathbb{R}^N} \left(\mathrm{e}^{\frac{|u(x)|^p}{\alpha^p}} - 1\right) dx < \infty \Longleftrightarrow \alpha > N^{-1/p}.$$

Therefore $u \in \exp L^p$ and $u \notin \exp L_0^p$.

(ii) Let u be the function defined by

$$u(x) = \Big(\log(1-\log|x|)\Big)^{1/p} \quad \text{if} \quad |x| \le 1,$$
$$u(x) = 0 \qquad \text{if} \qquad |x| > 1.$$

Clearly $u \notin L^\infty$. Moreover, for any $\alpha > 0$, we have

$$\int_{\mathbb{R}^N} \Big(\mathrm{e}^{\frac{|u(x)|^p}{\alpha^p}} - 1\Big)\, dx = |\mathbb{S}^{N-1}| \int_0^1 r^{N-1}\Big((1-\log r)^{\frac{1}{\alpha^p}} - 1\Big)\, dr < \infty,$$

where $|\mathbb{S}^{N-1}|$ is the measure of the unit sphere $\mathbb{S}^{N-1}$ in $\mathbb{R}^N$. The second assertion follows since $\exp L^p_0 \hookrightarrow \exp L^p$.

(iii) Let u be the function defined by

$$u(x) = |x|^{-\frac{N}{r}} \quad \text{if} \quad |x| \ge 1,$$
$$u(x) = 0 \qquad \text{if} \qquad |x| < 1.$$

Then $u \in \exp L^p_0$ but $u \notin L^r$. Indeed, it is clear that $u \notin L^r$, and for $\alpha > 0$, we have

$$\int_{\mathbb{R}^N} \Big(\mathrm{e}^{\frac{|u(x)|^p}{\alpha^p}} - 1\Big)\, dx = \frac{|\mathbb{S}^{N-1}|}{Nr} \sum_{k=1}^{\infty} \frac{1}{(pk-r)k!\alpha^{pk}} < \infty.$$

(iv) Let $u \in L^q \cap L^\infty$ and let $\alpha > 0$. Using the interpolation inequality

$$\|u\|_r \le \|u\|_q^{q/r}\, \|u\|_\infty^{1-q/r} \le \|u\|_q + \|u\|_\infty, \quad q \le r \le \infty,$$

we obtain

$$\begin{aligned}\int_{\mathbb{R}^N} \Big(\mathrm{e}^{\frac{|u(x)|^p}{\alpha^p}} - 1\Big)\, dx &= \sum_{k=1}^{\infty} \frac{1}{k!\alpha^{pk}} \|u\|_{L^{pk}}^{pk} \\ &\le \sum_{k=1}^{\infty} \frac{1}{k!\alpha^{pk}} \big(\|u\|_q + \|u\|_\infty\big)^{pk} \\ &= \mathrm{e}^{\frac{(\|u\|_q + \|u\|_\infty)^2}{\alpha^p}} - 1.\end{aligned}$$

This clearly implies (2-3).

□

We have the embedding: $\exp L^p \hookrightarrow L^q$ for every $1 < p \le q$. More precisely:

Lemma 2.4. *For every $1 \leq p \leq q < \infty$, we have*

$$\|u\|_q \leqslant \left(\Gamma\left(\frac{q}{p}+1\right)\right)^{\frac{1}{q}} \|u\|_{\exp L^p}, \tag{2-4}$$

where $\Gamma(x) := \int_0^\infty \tau^{x-1} \mathrm{e}^{-\tau}\, d\tau,\ x > 0.$

The proof of the previous lemma is similar to that in Ruf and Terraneo [2002]. For reader's convenience, we give it here.

Proof of Lemma 2.4. Let $K = \|u\|_{\exp L^p} > 0$. Using the inequality

$$\frac{|x|^{pr}}{\Gamma(r+1)} \leq \mathrm{e}^{|x|^p} - 1,\ r \geq 1,\ x \in \mathbb{R},$$

we have

$$\int_{\mathbb{R}^N} \frac{(|u(x)|/K)^{pr}}{\Gamma(r+1)} dx \leq \int_{\mathbb{R}^N} \left(\mathrm{e}^{(|u(x)|/K)^p} - 1\right) dx \leq 1.$$

This leads to

$$\|u\|_{pr} \leq (\Gamma(r+1))^{\frac{1}{pr}} K.$$

The result follows by taking $r = \frac{q}{p} \geq 1$. □

Remark 2.5. *For $\phi(s) = \mathrm{e}^s - 1 - s$, on can prove the following inequality*

$$\|u\|_q \leq C(q)\|u\|_{L^\phi},\ 2 \leq q < \infty,$$

for some constant $C(q) > 0$ depending only on q.

We recall that the following properties of the functions Γ and $\mathfrak{B}$ given by

$$\mathfrak{B}(x, y) = \int_0^1 \tau^{1-x}(1-\tau)^{1-y} d\tau, \quad x,\ y > 0.$$

We have

$$\mathfrak{B}(x, y) = \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)},\ \forall\ x,\ y > 0, \tag{2-5}$$

$$\Gamma(x) \geq C > 0,\ \forall\ x > 0, \tag{2-6}$$

$$\Gamma(x+1) \sim \left(\frac{x}{\mathrm{e}}\right)^x \sqrt{2\pi x}, \text{ as } x \to \infty, \tag{2-7}$$

and

$$\Gamma(x+1) \le C x^{x+\frac{1}{2}}, \ \forall\ x \ge 1. \tag{2-8}$$

The following Lemmas will be useful in the proof of the global existence.

Lemma 2.6. *Let $\lambda > 0$, $1 \le p,\ q < \infty$ and $K > 0$ such that $\lambda q\, K^p \le 1$. Assume that*

$$\|u\|_{\exp L^p} \le K\,.$$

Then

$$\|\mathrm{e}^{\lambda |u|^p} - 1\|_q \le (\lambda q\, K^p)^{\frac{1}{q}}\ .$$

Proof of Lemma 2.6. Write

$$\begin{aligned}
\int_{\mathbb{R}^N} \left(\mathrm{e}^{\lambda |u|^p} - 1\right)^q\, dx &\le \int_{\mathbb{R}^N} \left(\mathrm{e}^{\lambda\, q\, |u|^p} - 1\right)\, dx \\
&\le \int_{\mathbb{R}^N} \left(\mathrm{e}^{\lambda\, q\, K^p \frac{|u|^p}{\|u\|^p_{\exp L^p}}} - 1\right)\, dx \\
&\le \lambda\, q\, K^p \int_{\mathbb{R}^N} \left(\mathrm{e}^{\frac{|u|^p}{\|u\|^p_{\exp L^p}}} - 1\right)\, dx \le \lambda\, q\, K^p,
\end{aligned}$$

where we have used the fact that $\mathrm{e}^{\theta s} - 1 \le \theta\,(\mathrm{e}^s - 1)$, $0 \le \theta \le 1$, $s \ge 0$ and (2-2). ☐

Lemma 2.7. *Let $m \ge p > 1$, $a > \frac{N(m-1)}{2}$, $a > \frac{N}{2}$. Define*

$$\sigma = \frac{1}{m-1} - \frac{N}{2a} > 0.$$

Assume that

$$N > \frac{2p}{p-1}, \tag{2-9}$$

and

$$a < \frac{N(m-1)}{2}\,\frac{1}{(2-m)_+}\,. \tag{2-10}$$

Then, there exist $r,\ \ q,\ \ (\theta_k)_{k=0}^{\infty}\ ,(\rho_k)_{k=0}^{\infty}$ such that

$$1 \le r \le a\,. \tag{2-11}$$

$$q \geq 1 \quad and \quad \frac{1}{r} = \frac{1}{a} + \frac{1}{q}. \tag{2-12}$$

$$0 < \theta_k < 1 \quad and \quad \frac{1}{q(pk+m-1)} = \frac{\theta_k}{a} + \frac{1-\theta_k}{\rho_k}. \tag{2-13}$$

$$p \leq \rho_k < \infty. \tag{2-14}$$

$$\frac{N}{2}\left(\frac{1}{r} - \frac{1}{a}\right) < 1. \tag{2-15}$$

$$\sigma\Big[\theta_k(pk+m-1)+1\Big] < 1. \tag{2-16}$$

$$1 - \frac{N}{2}\left(\frac{1}{r} - \frac{1}{a}\right) - \sigma\theta_k(pk+m-1) = 0. \tag{2-17}$$

Moreover,

$$\theta_k \longrightarrow 0 \quad as \quad k \longrightarrow \infty. \tag{2-18}$$

$$\rho_k \longrightarrow \infty \quad as \quad k \longrightarrow \infty. \tag{2-19}$$

$$\frac{(pk+m-1)(1-\theta_k)}{p\rho_k}(1+\rho_k) \leq k, \quad \forall\ k \geq 1. \tag{2-20}$$

Remark 2.8. *The assumption (2-10) together with $a > \frac{N}{2}$ implies that $m > \frac{3}{2}$.*

Proof of Lemma 2.7. Note that the assumption (2-10) implies that $\sigma < 1$. It follows that, for all integer $k \geq 0$ one can choose θ_k such that

$$0 < \theta_k < \frac{1}{pk+m-1}\min\left(m-1, \frac{1-\sigma}{\sigma}\right). \tag{2-21}$$

Next, we choose ρ_k such that

$$\frac{1-\theta_k}{\rho_k} = \frac{2}{N(pk+m-1)} - \frac{2\theta_k}{N(m-1)}. \tag{2-22}$$

Finally, we choose q such that

$$\frac{1}{q(pk+m-1)} = \frac{\theta_k}{a} + \frac{1-\theta_k}{\rho_k}. \tag{2-23}$$

This leads to all remainder parameters. □

We state the following proposition which is needed for the local well-posedness in the space $\exp L_0^p$.

Proposition 2.9. *Let $1 \le p < \infty$ and $u \in C([0,T]; \exp L^p)$. Then for every $\alpha > 0$ there holds*

$$\left(\mathrm{e}^{\alpha|u|^p} - 1\right) \in C([0,T]; L^r),\ 1 \le r < \infty.$$

Proof of Proposition 2.9. Although the proof is similar to that given in Majdoub, Otsmane, and Tayachi [2018], we give it here for completeness. Using the inequality

$$|\mathrm{e}^x - \mathrm{e}^y|^r \le |\mathrm{e}^{rx} - \mathrm{e}^{ry}|, \quad x,\ y \in \mathbb{R},$$

it suffices to consider only the case $r = 1$. Note that the proof for $p = 2$ was done in Ibrahim, Jrad, Majdoub, and Saanouni [2014]. The case $p = 1$ follows by the inequality

$$\left|\mathrm{e}^{|x|-|y|} - 1\right| \le \mathrm{e}^{|x-y|} - 1, \quad x,\ y \in \mathbb{R},$$

and property (v) in Proposition 2.2. The general case follows from the following lemmas.

Lemma 2.10. *Assume that*

$$v_n \to v \quad in \quad \exp L^p.$$

Then, for any $\alpha > 0$, we have

$$\mathrm{e}^{\alpha|v_n - v|^p} - 1 \to 0 \quad in \quad L^1.$$

Proof of Lemma 2.10. It suffices to consider the case $v = 0$ and $\alpha = 1$. For given $0 < \varepsilon \le 1$, there exists $N \ge 1$ such that $\|v_n\|_{\exp L^p} \le \varepsilon$ for all $n \ge N$. By definition of the norm $\|\cdot\|_{\exp L^p}$, there exists $0 < \lambda = \lambda_n < \varepsilon$ such that

$$\int_{\mathbb{R}^N} \left(\mathrm{e}^{|\frac{v_n}{\lambda}|^p} - 1\right) dx \le 1, \quad \forall \quad n \ge N.$$

By convexity argument, we deduce that

$$\begin{aligned}\int_{\mathbb{R}^N}\left(\mathrm{e}^{|v_n|^p}-1\right)dx &= \int_{\mathbb{R}^N}\left(\mathrm{e}^{\lambda^p|\frac{v_n}{\lambda}|^p}-1\right)dx\\ &\le \int_{\mathbb{R}^N}\left(\mathrm{e}^{|\varepsilon\frac{v_n}{\lambda}|^p}-1\right)dx\\ &\le \varepsilon\int_{\mathbb{R}^N}\left(\mathrm{e}^{|\frac{v_n}{\lambda}|^p}-1\right)dx\\ &\le \varepsilon.\end{aligned}$$

□

Lemma 2.11. *Let* $1<p<\infty$ *and* $v\in \exp L^p$. *Assume that*

$$w_n\to 0 \quad \text{in} \quad \exp L^p.$$

Then, for any $\alpha>0$, *we have*

$$\mathrm{e}^{\alpha|w_n||v|^{p-1}}-1\to 0 \quad \text{in} \quad L^1.$$

Proof of Lemma 2.11. Write

$$\begin{aligned}\left\|\mathrm{e}^{\alpha|w_n||v|^{p-1}}-1\right\|_{L^1} &= \sum_{k=1}^{\infty}\frac{\alpha^k}{k!}\int |w_n|^k|v|^{k(p-1)}\,dx\\ &\le \sum_{k=1}^{\infty}\frac{\alpha^k}{k!}\|w_n\|^k_{L^{kp}}\,\|v\|^{k(p-1)}_{L^{kp}}\end{aligned}$$

where we have used Hölder's inequality with

$$\frac{1}{k}=\frac{1}{kp}+\frac{1}{k\frac{p}{p-1}}.$$

Hence, using (2-4), we deduce that

$$\begin{aligned}\left\|\mathrm{e}^{\alpha|w_n||v|^{p-1}}-1\right\|_{L^1} &\le \sum_{k=1}^{\infty}\frac{\alpha^k}{k!}(k!)^{1/p}(k!)^{1-1/p}\,\|w_n\|^k_{\exp L^p}\,\|v\|^{k(p-1)}_{\exp L^p}\\ &\le \sum_{k=1}^{\infty}\left(\alpha\|w_n\|_{\exp L^p}\|v\|^{p-1}_{\exp L^p}\right)^k\\ &\le \frac{\alpha\,\|w_n\|_{\exp L^p}\,\|v\|^{p-1}_{\exp L^p}}{1-\alpha\,\|w_n\|_{\exp L^p}\,\|v\|^{p-1}_{\exp L^p}}\longrightarrow 0.\end{aligned}$$

□

Lemma 2.12. *Let $1 < p < \infty$. Assume that*

$$v_n \to v \quad in \quad \exp L^p.$$

Then,

$$\mathrm{e}^{|v_n|^p} - \mathrm{e}^{|v|^p} \to 0 \quad in \quad L^1.$$

Proof of Lemma 2.12. Set $w_n = v_n - v$, then

$$\mathrm{e}^{|v_n|^p} - \mathrm{e}^{|v|^p} = \left(\mathrm{e}^{|v|^p} - 1\right)\left(\mathrm{e}^{|w_n+v|^p-|v|^p} - 1\right) + \left(\mathrm{e}^{|w_n+v|^p-|v|^p} - 1\right).$$

Using the following elementary inequality

$$\exists \quad \alpha > 0 \quad \text{such that} \quad \Big| |a+b|^p - |b|^p \Big| \leq \alpha \left(|a|^p + |a||b|^{p-1}\right), \quad \forall \quad a, b \in \mathbb{R},$$

it follows that

$$\left\| \mathrm{e}^{|w_n+v|^p-|v|^p} - 1 \right\|_{L^1} \leq \left\| \mathrm{e}^{\alpha|w_n|^p+\alpha|w_n||v|^{p-1}} - 1 \right\|_{L^1}.$$

Let us write

$$\mathrm{e}^{\alpha|w_n|^p+\alpha|w_n||v|^{p-1}} - 1 = \mathbf{I}_n + \mathbf{J}_n + \mathbf{K}_n,$$

where

$$\begin{aligned} \mathbf{I}_n &= \left(\mathrm{e}^{\alpha|w_n|^p} - 1\right)\left(\mathrm{e}^{\alpha|w_n||v|^{p-1}} - 1\right) \\ \mathbf{J}_n &= \left(\mathrm{e}^{\alpha|w_n|^p} - 1\right) \\ \mathbf{K}_n &= \left(\mathrm{e}^{\alpha|w_n||v|^{p-1}} - 1\right) \end{aligned}$$

By Lemma 2.11 and since $w_n \to 0$ in $\exp L^p$, $v \in \exp L^p$, we deduce that

$$\begin{aligned} \mathbf{I}_n &\longrightarrow 0 \quad \text{in} \quad L^1, \\ \mathbf{J}_n &\longrightarrow 0 \quad \text{in} \quad L^1, \\ \mathbf{K}_n &\longrightarrow 0 \quad \text{in} \quad L^1. \end{aligned}$$

The proof of Lemma 2.12 is complete. □

Combining Lemmas 2.10-2.11-2.12, we easily deduce the desired result; that is

$$\mathrm{e}^{\alpha|u|^p} - 1 \in C([0,T]; L^1),$$

whenever $u \in C([0,T]; \exp L^p)$. This finishes the proof of Proposition 2.9. □

A straightforward consequence is:

Corollary 2.13. *Let* $1 \leq p < \infty$ *and* $u \in C([0,T]; \exp L^p)$. *Assume that* f *satisfies* (1-4). *Then for every* $p \leq r < \infty$ *there holds*

$$f(u) \in C\left([0,T]; L^r\right).$$

Proof. Fix $p \leq r < \infty$, $0 \leq t \leq T$ and let $(t_n) \subset [0,T]$ such that $t_n \to t$. Using Hölder's inequality, we obtain

$$\begin{aligned}
\|f(u(t_n)) - f(u(t))\|_r \leq{}& 2C\,\|u(t_n) - u(t)\|_r + \\
& C\,\||u(t_n) - u(t)|(\mathrm{e}^{\lambda\,|u(t_n)|^p} - 1 + \mathrm{e}^{\lambda\,|u(t)|^p} - 1)\|_r \\
& \leq 2C\,\|u(t_n) - u(t)\|_r + C\,\|u(t_n) - u(t)\|_{2r}\times \\
& \left(\|\mathrm{e}^{\lambda|u(t_n)|^p} - 1\|_{2r} + \|\mathrm{e}^{\lambda|u(t)|^p} - 1\|_{2r}\right) \\
\leq C\,\|u(t_n) - u(t)\|_{\exp L^p} & \left(1 + \|\mathrm{e}^{\lambda|u(t_n)|^p} - 1\|_{2r} + \|\mathrm{e}^{\lambda|u(t)|^p} - 1\|_{2r}\right),
\end{aligned}$$

where we have used Lemma 2.4 in the last inequality. From Proposition 2.9 we know that $\|\mathrm{e}^{\lambda|u(t_n)|^p} - 1\|_{2r} \to \|\mathrm{e}^{\lambda|u(t)|^p} - 1\|_{2r}$ as $n \to \infty$. It follows that $\|f(u(t_n)) - f(u(t))\|_r \to 0$ which is the desired conclusion. $\square$

3 Linear estimates

In this section we establish some results needed for the proofs of the main theorems. We first recall some basic estimates for the linear heat semigroup $\mathrm{e}^{t\Delta}$. The solution of the linear heat equation

$$\begin{cases} \partial_t u = \Delta u, \ t > 0, \ x \in \mathbb{R}^N, \\ \qquad u(0,x) = u_0(x), \end{cases}$$

can be written as a convolution:

$$u(t,x) = \big(G_t \star u_0\big)(x) := \big(\mathrm{e}^{t\Delta} u_0\big)(x),$$

where

$$G_t(x) := G(t,x) = \frac{\mathrm{e}^{-\frac{|x|^2}{4t}}}{(4\pi t)^{\frac{N}{2}}}, \ t > 0, \ x \in \mathbb{R}^N,$$

is the heat kernel. We will frequently use the $L^r - L^\rho$ estimate as stated in the Proposition below.

Proposition 3.1. *For all* $1 \leq r \leq \rho \leq \infty$, *we have*

$$\|\mathrm{e}^{t\Delta}\varphi\|_\rho \leqslant t^{-\frac{N}{2}(\frac{1}{r}-\frac{1}{\rho})}\|\varphi\|_r, \qquad \forall\, t>0,\ \forall\,\varphi\in L^r. \tag{3-1}$$

The following Proposition is a generalization of Ioku [2011, Lemma 2.2, p. 1176].

Proposition 3.2. *Let* $1 \leqslant q \leqslant p,\ 1 \leqslant r \leqslant \infty$. *Then the following estimates hold:*

(i) $\|\mathrm{e}^{t\Delta}\varphi\|_{\exp L^p} \leqslant \|\varphi\|_{\exp L^p},\ \forall\, t>0,\ \forall\,\varphi\in \exp L^p.$

(ii) $\|\mathrm{e}^{t\Delta}\varphi\|_{\exp L^p} \leqslant t^{-\frac{N}{2q}}\left(\log(t^{-\frac{N}{2}}+1)\right)^{-\frac{1}{p}}\|\varphi\|_q,\ \forall\, t>0,\ \forall\,\varphi\in L^q.$

(iii) $\|\mathrm{e}^{t\Delta}\varphi\|_{\exp L^p} \leqslant \frac{1}{(\log 2)^{\frac{1}{p}}}\left[t^{-\frac{N}{2r}}\|\varphi\|_r+\|\varphi\|_q\right],\ \forall\, t>0, \forall\,\varphi\in L^r\cap L^q.$

Proof of Proposition 3.2. We begin by proving (i). For any $\alpha>0$, expanding the exponential function leads to

$$\int_{\mathbb{R}^N}\left(\exp\left|\frac{\mathrm{e}^{t\Delta}\varphi}{\alpha}\right|^p-1\right)dx=\sum_{k=1}^{\infty}\frac{\|\mathrm{e}^{t\Delta}\varphi\|_{pk}^{pk}}{k!\alpha^{pk}}.$$

Then by the $L^{pk}-L^{pk}$ estimate of the heat semi-group (3-1), we obtain

$$\begin{aligned}\int_{\mathbb{R}^N}\left(\exp\left|\frac{\mathrm{e}^{t\Delta}\varphi}{\alpha}\right|^p-1\right)dx &\leqslant \sum_{k=1}^{\infty}\frac{\|\varphi\|_{pk}^{pk}}{k!\alpha^{pk}}\\ &=\int_{\mathbb{R}^N}\left(\exp\left|\frac{\varphi}{\alpha}\right|^p-1\right)dx.\end{aligned}$$

Therefore we obtain

$$\begin{aligned}\|\mathrm{e}^{t\Delta}\varphi\|_{\exp L^p}&=\inf\left\{\alpha>0,\ \int_{\mathbb{R}^N}\left(\exp\left|\frac{\mathrm{e}^{t\Delta}\varphi}{\alpha}\right|^p-1\right)dx\leq 1\right\}\\ &\leqslant \inf\left\{\alpha>0,\ \int_{\mathbb{R}^N}\left(\exp\left|\frac{\varphi}{\alpha}\right|^p-1\right)dx\leq 1\right\}\\ &=\|\varphi\|_{\exp L^p}.\end{aligned}$$

This proves (i).

We now turn to the proof of (ii). Using (3-1) with $q \leq p$, we have

$$\begin{aligned}\int_{\mathbb{R}^N}\left(\exp\left|\frac{e^{t\Delta}\varphi}{\alpha}\right|^p - 1\right)dx &= \sum_{k=1}^{\infty}\frac{\|e^{t\Delta}\varphi\|_{pk}^{pk}}{k!\alpha^{pk}}\\ &\leqslant \sum_{k=1}^{\infty}\frac{t^{-\frac{N}{2}(\frac{1}{q}-\frac{1}{pk})pk}\|\varphi\|_q^{pk}}{k!\alpha^{pk}}\\ &= t^{\frac{N}{2}}\left(\exp\left(\frac{t^{-\frac{N}{2q}}\|\varphi\|_q}{\alpha}\right)^p - 1\right).\end{aligned}$$

It follows that

$$\|e^{t\Delta}\varphi\|_{\exp L^p} \leq t^{-\frac{N}{2q}}\left(\log(t^{-\frac{N}{2}}+1)\right)^{-\frac{1}{p}}\|\varphi\|_q.$$

This proves (ii).

We now prove (iii). By the embedding $L^q \cap L^\infty \hookrightarrow \exp L^p$ (2-3), we have

$$\|e^{t\Delta}\varphi\|_{\exp L^p} \leqslant \frac{1}{(\log 2)^{1/p}}\left[\|e^{t\Delta}\varphi\|_\infty + \|e^{t\Delta}\varphi\|_q\right].$$

Using the $L^r - L^\infty$ estimate (3-1), we get

$$\|e^{t\Delta}\varphi\|_{\exp L^p} \leqslant \frac{1}{(\log 2)^{1/p}}\left[t^{-\frac{N}{2r}}\|\varphi\|_r + \|\varphi\|_q\right].$$

This proves (iii). The proof of the proposition is now complete. □

As a consequence we have the following, the proof of which can be done as in Majdoub, Otsmane, and Tayachi [2018].

Corollary 3.3. *Let $p > 1$, $N > \frac{2p}{p-1}$, $r > \frac{N}{2}$. Then, for every $g \in L^1 \cap L^r$, we have*

$$\|e^{t\Delta}g\|_{\exp L^p} \leq \kappa(t)\,\|g\|_{L^1\cap L^r},\ \forall\ t > 0,$$

where $\kappa \in L^1(0,\infty)$ is given by

$$\kappa(t) = \frac{1}{(\log 2)^{\frac{1}{p}}}\min\left\{t^{-\frac{N}{2r}}+1,\ t^{-\frac{N}{2}}\left(\log(t^{-\frac{N}{2}}+1)\right)^{-\frac{1}{p}}\right\}.$$

Here we use $\|g\|_{L^1\cap L^q} = \|g\|_1 + \|g\|_q$.

Proof of Corollary 3.3. We have, by Proposition 3.2 (ii) with $q = 1$,

$$\|e^{t\Delta} g\|_{\exp L^p} \leq t^{-\frac{N}{2}} \Big(\log(t^{-\frac{N}{2}} + 1)\Big)^{-\frac{1}{p}} \|g\|_1. \tag{3-2}$$

Using Proposition 3.2 (iii) with $q = 1$, we get

$$\|e^{t\Delta} g\|_{\exp L^p} \leq \frac{1}{(\log 2)^{\frac{1}{p}}} \Big(t^{-\frac{N}{2r}} + 1\Big)\Big[\|g\|_r + \|g\|_1\Big]. \tag{3-3}$$

Combining the inequalities (3-2) and (3-3), we obtain

$$\|e^{t\Delta} g\|_{\exp L^p} \leq \kappa(t) \Big(\|g\|_1 + \|g\|_r\Big). \tag{3-4}$$

By the assumption $N > \frac{2p}{p-1}$, $r > \frac{N}{2}$, we can see that $\kappa \in L^1(0, \infty)$. □

We will also need the following result for the proofs.

Proposition 3.4. *If* $u_0 \in \exp L^p_0$ *then* $e^{t\Delta} u_0 \in C([0, \infty); \exp L^p_0)$.

It is known that $e^{t\Delta}$ is a C^0−semigroup on L^p. By Proposition 3.4, it is also a C^0−semigroup on $\exp L^p_0$. This is not the case on $\exp L^p$. We have the following result.

Proposition 3.5. *There exist* $u_0 \in \exp L^p$ *and a constant* $C > 0$ *such that*

$$\|e^{t\Delta} u_0 - u_0\|_{\exp L^p} \geqslant C, \ \forall\, t > 0. \tag{3-5}$$

The proof of the previous proposition uses the notion of rearrangement of functions and can be done as in Majdoub, Otsmane, and Tayachi [2018].

4 Local well-posedness

In this section we prove the existence and the uniqueness of solution to (1-1) in $C([0, T]; \exp L^p_0)$ for some $T > 0$, namely Theorem 1.4. Throughout this section we assume that the nonlinearity $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(0) = 0$ and

$$|f(u) - f(v)| \leq C|u - v| \Big(e^{\lambda |u|^p} + e^{\lambda |v|^p}\Big), \quad \forall\, u,\ v \in \mathbb{R} \tag{4-1}$$

for some constants $C > 0$, $\lambda > 0$ $p \geq 1$. We emphasize that, thanks to Corollary 2.13, the Cauchy problem (1-1) admits the equivalent integral formulation (1-2). This is formulated as follows.

Proposition 4.1. *Let $T > 0$ and u_0 be in $\exp L^p_0$. If u belongs to $C([0,T];\exp L^p_0)$, then u is a weak solution of (1-1) if and only if $u(t)$ satisfies the integral equation (1-2) for any $t \in (0,T)$.*

Now we are ready to prove Theorem 1.4. The idea is to split the initial data $u_0 \in \exp L^p_0$ into a small part in $\exp L^p$ and a smooth one. This will be done using the density of $C^\infty_0(\mathbb{R}^N)$ in $\exp L^p_0$. First we solve the initial value problem with smooth initial data to obtain a local and bounded solution v. Then we consider the perturbed equation satisfied by $w := u - v$ and with small initial data. Now we come to the details. For $\varepsilon > 0$ to be chosen later, we write $u_0 = v_0 + w_0$, where $v_0 \in C^\infty_0(\mathbb{R}^N)$ and $\|w_0\|_{\exp L^p} \le \varepsilon$. Then, we consider the two Cauchy problems:

$$(\mathcal{P}_1) \qquad \begin{cases} \partial_t v - \Delta v = f(v), & t > 0,\ x \in \mathbb{R}^N, \\ v(0) = v_0, \end{cases}$$

and

$$(\mathcal{P}_2) \qquad \begin{cases} \partial_t w - \Delta w = f(w+v) - f(v), & t > 0,\ x \in \mathbb{R}^N, \\ w(0) = w_0. \end{cases}$$

We first, prove the following existence result concerning $(\mathcal{P}_1)$.

Proposition 4.2. *Let $v_0 \in L^p \cap L^\infty$. Then there exist a time $T > 0$ and a solution $v \in C([0,T], \exp L^p_0) \cap L^\infty(0,T;L^\infty)$ to $(\mathcal{P}_1)$.*

Proof of Proposition 4.2. We use a fixed point argument. We introduce, for any $M > 0$, and positive time T the following complete metric space

$$\mathcal{Y}(M,T) := \Big\{ v \in C([0,T];\exp L^p_0) \cap L^\infty(0,T;L^\infty); \quad \|v\|_T \le M \Big\},$$

where $\|v\|_T := \|v\|_{L^\infty(0,T;L^p)} + \|v\|_{L^\infty(0,T;L^\infty)}$, and $\|v_0\|_{L^p\cap L^\infty} = \|v_0\|_p + \|v_0\|_\infty$. Set

$$\Phi(v)(t) := \mathrm{e}^{t\Delta} v_0 + \int_0^t \mathrm{e}^{(t-s)\Delta} f(v(s))\, ds.$$

We will prove that, for suitable $M > 0$ and $T > 0$, Φ is a contraction map from $\mathcal{Y}(M,T)$ into itself.

First, since $v_0 \in L^p \cap L^\infty$, then by Lemma 2.3 (iv), $v_0 \in \exp L^p_0$ and by Proposition 3.4, $\mathrm{e}^{t\Delta} v_0 \in C([0,T];\exp L^p_0)$. Obviously $\mathrm{e}^{t\Delta} v_0 \in L^\infty(0,T;L^\infty)$. Second by (1-4), $f(v) \in L^1(0,T;\exp L^p_0)$ whenever $v \in C([0,T];\exp L^p_0) \cap L^\infty(0,T;L^\infty)$. Then, by Proposition 3.4, we conclude that $\Phi(v) \in C([0,T];\exp L^p_0) \cap L^\infty(0,T;L^\infty)$.

Now, for every $v_1,\ v_2 \in \mathcal{Y}(M,T)$, we have thanks to (4-1),

$$\begin{aligned}
\|\Phi(v_1)-\Phi(v_2)\|_{L^\infty(0,T;L^q)} &\le C\int_0^T \|f(v_1(s))-f(v_2(s))\|_q\,ds\\
&\le T\|f(v_1)-f(v_2)\|_{L^\infty(0,T;L^q)}\\
&\le CT\left(\mathrm{e}^{\lambda\|v_1\|^p_{L^\infty_t(L^\infty_x)}}+\mathrm{e}^{\lambda\|v_2\|^p_{L^\infty_t(L^\infty_x)}}\right)\|v_1-v_2\|_{L^\infty(0,T;L^q)}
\end{aligned}$$

where $q=p$ or $q=\infty$. Then, it follows that

$$\begin{aligned}
\|\Phi(v_1)-\Phi(v_2)\|_T &\le 2C\ T\,\mathrm{e}^{\lambda M^p}\|v_1-v_2\|_T\\
&\le 2C\ T\,\mathrm{e}^{\lambda M^p}\|v_1-v_2\|_T.
\end{aligned}\tag{4-2}$$

Similarly we have

$$\begin{aligned}
\|\Phi(v)\|_T &\le \|v_0\|_{L^p\cap L^\infty}+C\ \ T\,\mathrm{e}^{\lambda M^p}\ \|v\|_T\\
&\le \|v_0\|_{L^p\cap L^\infty}+2CM\mathrm{e}^{\lambda M^p}T.
\end{aligned}\tag{4-3}$$

From (4-2) and (4-3) we conclude that for $T>0$ and $M>\|v_0\|_{L^p\cap L^\infty}$ such that

$$2C\mathrm{e}^{\lambda M^p}T<1,\ \ \|v_0\|_{L^p\cap L^\infty}+2CM\mathrm{e}^{\lambda M^p}T\le M,$$

Φ is a contraction map on $\mathcal{Y}(M,T)$. In particular, one can take $M>\|v_0\|_{L^p\cap L^\infty}$ and $T<\frac{M-\|v_0\|_{L^p\cap L^\infty}}{2MC\mathrm{e}^{\lambda M^p}}$. This finishes the proof of Proposition 4.2. □

Following similar arguments as in Majdoub, Otsmane, and Tayachi [2018] and using Propositions 4.1-4.2, we end the proof of Theorem 1.4.

The solution constructed by the above Proposition can be extended to a maximal solution by well known argument. Moreover, if $T_{\max}<\infty$, then $\lim\limits_{t\to T_{\max}}\|u(t)\|_{L^p\cap L^\infty}=\infty$. Let us now give the proof of the lower blow-up estimates.

Proof of Theorem 1.8. Let $u_0\in L^p\cap L^\infty$ and $u\in C([0,T_{\max}),\exp L^p_0)$ be the maximal solution of (1-1) given by Theorem 1.4 (or Proposition 4.2). To prove the lower blow-up estimates we use an argument introduced by Weissler in Weissler [1981, Section 4 and Remark (6)2]. See also Mueller and Weissler [1985, Proposition 5.3, p. 901]. Assume that $T_{\max}<\infty$. Then $\lim\limits_{t\to T_{\max}}\|u(t)\|_{L^p\cap L^\infty}=\infty$. Consider u the solution starting at $u(t)$ for some $t\in[0,T_{\max})$. If for some M

$$\|u(t)\|_{L^p\cap L^\infty}+2CM\mathrm{e}^{\lambda M^p}(T-t)\le M,$$

then $T < T_{\max}$. Therefore, for any $M > 0$,

$$\|u(t)\|_{L^p\cap L^\infty} + 2CM\mathrm{e}^{\lambda M^p}(T_{\max}-t) > M.$$

Choosing $M = 2\|u(t)\|_{L^p\cap L^\infty}$ it follows that

$$4C\,\|u(t)\|_{L^p\cap L^\infty}\mathrm{e}^{2^p\lambda\|u(t)\|^p_{L^p\cap L^\infty}}(T_{\max}-t) > \|u(t)\|_{L^p\cap L^\infty}.$$

That is

$$\mathrm{e}^{2^p\lambda\|u(t)\|^p_{L^p\cap L^\infty}} \geq C(T_{\max}-t)^{-1},$$

for some positive constant C. Hence,

$$2^p\lambda\|u(t)\|^p_{L^p\cap L^\infty} \geq -\log(T_{\max}-t)+C.$$

Then

$$\lambda\|u(t)\|^p_{L^p\cap L^\infty} \geq -C_1\log(T_{\max}-t)+C_2$$

for some positive constants $C_1,\ C_2$. This completes the proof of Theorem 1.8. □

We obtain the following concerning problem $(\mathcal{P}_2)$.

Proposition 4.3. *Let $T > 0$ and $v \in L^\infty(0,T;L^\infty)$ given by Proposition 4.2. Let $w_0 \in \exp L^p_0$. Then for $\|w_0\|_{\exp L^p} \leq \varepsilon$, with $\varepsilon > 0$ small enough, there exist a time $\widetilde{T} = \widetilde{T}(w_0,\varepsilon,v) > 0$ and a solution $w \in C([0,\widetilde{T}],\exp L^p_0)$ to problem $(\mathcal{P}_2)$.*

The proof of Proposition 4.3 uses the following lemma.

Lemma 4.4. *Let $v \in L^\infty$ and $w_1,\ w_2 \in \exp L^p$ with $\|w_1\|_{\exp L^p},\ \|w_2\|_{\exp L^p} \leq M$ for some constant $M > 0$. Let $p \leq q < \infty$, and assume that $2^p\lambda q M^p \leq 1$ where λ is given by (4-1). Then there exists a constant $C > 0$ such that*

$$\Big\|f(w_1+v)-f(w_2+v)\Big\|_q \leq C\,\mathrm{e}^{2^{p-1}\lambda\|v\|^p_\infty}\Big\|w_1-w_2\Big\|_{\exp L^p}.$$

Proof of the Lemma 4.4. By the assumption (4-1) on f, we have

$$
\begin{aligned}
\Big\| f(w_1+v) - f(w_2+v)\Big\|_q \leq \\
&\leq C\Big\| |w_1-w_2|\left(\mathrm{e}^{2^{p-1}\lambda|w_1|^p+2^{p-1}\lambda|v|^p}+\mathrm{e}^{2^{p-1}\lambda|w_2|^p+2^{p-1}\lambda|v|^p}\right)\Big\|_q \\
&\leq \mathrm{e}^{2^{p-1}\lambda\|v\|_\infty^p}\left(2C\Big\|w_1-w_2\Big\|_q + C\Big\||w_1-w_2|\left(\mathrm{e}^{2^{p-1}\lambda|w_1|^p}-1\right)\Big\|_q\right) \\
&\qquad + C\mathrm{e}^{2^{p-1}\lambda\|v\|_\infty^p}\Big\||w_1-w_2|\left(\mathrm{e}^{2^{p-1}\lambda|w_2|^p}-1\right)\Big\|_q \\
&\leq \mathrm{e}^{2^{p-1}\lambda\|v\|_\infty^p}\left(2C\Big\|w_1-w_2\Big\|_q + C\Big\|w_1-w_2\Big\|_{2q}\Big\|\mathrm{e}^{2^{p-1}\lambda|w_1|^p}-1\Big\|_{2q}\right) \\
&\qquad + C\mathrm{e}^{2^{p-1}\lambda\|v\|_\infty^p}\Big\|w_1-w_2\Big\|_{2q}\Big\|\mathrm{e}^{2^{p-1}\lambda|w_2|^p}-1\Big\|_{2q} \\
&\leq C\,\mathrm{e}^{2^{p-1}\lambda\|v\|_\infty^p}\Big\|w_1-w_2\Big\|_{\exp L^p},
\end{aligned}
$$

where we have used Hölder inequality, Lemma 2.4, Lemma 2.6 and the fact that $(a+b)^p \leq 2^{p-1}(a^p+b^p)$, for every $a,b \geq 0$ and any $p \geq 1$. This finishes the proof of Lemma 4.4. □

5 Non-existence

The following lemma is the key of the proof of Theorem 1.5.

Lemma 5.1. *Let $p>1$, $\alpha>0$. Let Φ_α be given by (1-5) and f, $\lambda>0$ be as in (1-6). Then, there exists $\alpha_0>0$ such that for any $\alpha\geq\alpha_0$, $\varepsilon>0$ and $r>0$, we have*

$$\int_0^\varepsilon \int_{|x|<r} \exp\left(\lambda\left(\mathrm{e}^{t\Delta}\Phi_\alpha\right)^p\right)\,dx\,dt = \infty\,.$$

Proof of Lemma 5.1. Let $B(a,\rho)$ denotes the open ball centered at $a\in\mathbb{R}^N$ and with radius $\rho>0$. Fix ε, $r>0$. For $\rho=\min\left(r,\frac{1}{4}\right)$, we have $B(3x,|x|)\subset B(0,1)$ for any $|x|<\rho$. Therefore, for any $|x|<\rho$, it holds

$$
\begin{aligned}
\left(\mathrm{e}^{t\Delta}\Phi_\alpha\right)(x) &= \frac{1}{(4\pi t)^{N/2}}\int_{|x|<1}\mathrm{e}^{-\frac{|x-y|^2}{4t}}\,\Phi_\alpha(y)\,dy \\
&\geq \frac{\alpha}{(4\pi t)^{N/2}}\int_{|y-3x|<|x|}\mathrm{e}^{-\frac{|x-y|^2}{4t}}\left(-\log|y|\right)^{\frac{1}{p}}\,dy \\
&\geq C\alpha\left(\frac{|x|^2}{t}\right)^{N/2}\mathrm{e}^{-\frac{9}{4}\frac{|x|^2}{t}}\left(-\log 4|x|\right)^{1/p}.
\end{aligned}
$$

Let $\eta = \min\left(\varepsilon, \rho^2\right)$. Then, for any $0 < t < \eta$, we have $B(0, \sqrt{t}) \subset B(0, \rho)$. Hence

$$\begin{aligned} \int_0^\varepsilon \int_{|x|<r} \exp\left(\lambda\left(\mathrm{e}^{t\Delta}\Phi_\alpha\right)^p\right) dx\,dt &\geq \int_0^\eta \int_{|x|<\rho} \exp\left(\lambda\left(\mathrm{e}^{t\Delta}\Phi_\alpha\right)^p\right) dx\,dt \\ &\geq \int_0^\eta \int_{\frac{\sqrt{t}}{2}<|x|<\sqrt{t}} \exp\left(-C\lambda\alpha^p \log(4|x|)\right) dx\,dt \\ &\geq C_\alpha \int_0^\eta t^{\frac{N}{2}-\frac{C\lambda\alpha^p}{2}}\,dt = \infty, \end{aligned}$$

for $\alpha \geq \alpha_0 := \left(\frac{N+2}{C\lambda}\right)^{1/p}$. This finishes the proof of Lemma 5.1. $\square$

The proof of Theorem 1.5 follows similar arguments as in Ioku, Ruf, and Terraneo [2015] and uses the previous Lemma.

6 Global Existence

This section is devoted to the proof of Theorem 1.6. The proof uses a fixed point argument on the associated integral equation

$$u(t) = \mathrm{e}^{t\Delta}u_0 + \int_0^t \mathrm{e}^{(t-s)\Delta}(f(u))(s)ds, \tag{6-1}$$

where $\|u_0\|_{\exp L^p} \leq \varepsilon$, with small $\varepsilon > 0$ to be fixed later. The nonlinearity f satisfies $f(0) = 0$ and

$$|f(u) - f(v)| \leq C\,|u - v|\left(|u|^{m-1}\mathrm{e}^{\lambda|u|^p} + |v|^{m-1}\mathrm{e}^{\lambda|v|^p}\right), \tag{6-2}$$

for some constants $C > 0$ and $\lambda > 0$, $p \geq 1$ and m is larger than $1 + \frac{2p}{N}$. From (6-2), we obviously deduce that

$$|f(u) - f(v)| \leq C|u - v| \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\left(|u|^{pk+m-1} + |v|^{pk+m-1}\right). \tag{6-3}$$

We will perform a fixed point argument on a suitable metric space. For $M > 0$ we introduce the space

$$Y_M := \left\{u \in L^\infty(0, \infty, \exp L^p);\ \sup_{t>0} t^\sigma \|u(t)\|_a + \|u\|_{L^\infty(0,\infty;\exp L^p)} \leq M\right\},$$

where $a > \frac{N(m-1)}{2} \geq p$ and

$$\sigma = \frac{1}{m-1} - \frac{N}{2a} = \frac{N}{2}\left(\frac{2}{N(m-1)} - \frac{1}{a}\right) > 0.$$

Endowed with the metric $d(u,v) = \sup_{t>0}\left(t^{\sigma}\|u(t)-v(t)\|_r\right)$, Y_M is a complete metric space. This follows by Proposition 2.2.

For $u \in Y_M$, we define $\Phi(u)$ by

$$\Phi(u)(t) := \mathrm{e}^{t\Delta}u_0 + \int_0^t \mathrm{e}^{(t-s)\Delta}(f(u(s)))ds. \tag{6-4}$$

By Proposition 3.2 (i), Proposition 3.1 and Lemma 2.4, we have

$$\|\mathrm{e}^{t\Delta}u_0\|_{\exp L^p} \leq \|u_0\|_{\exp L^p},$$

and

$$\begin{aligned} t^{\sigma}\|\mathrm{e}^{t\Delta}u_0\|_a &\leq t^{\sigma} t^{-\frac{N}{2}\left(\frac{2}{N(m-1)}-\frac{1}{a}\right)}\|u_0\|_{\frac{N(m-1)}{2}} \\ &= \|u_0\|_{\frac{N(m-1)}{2}} \leq C\|u_0\|_{\exp L^p}, \end{aligned}$$

where we have used $1 \leq p \leq \frac{N(m-1)}{2} < a$.

Let $u \in Y_M$. Using Proposition 3.2 and Corollary 3.3, we get for $q > \frac{N}{2}$,

$$\begin{aligned} \|\Phi(u)(t)\|_{\exp L^p} &\leq \|\mathrm{e}^{t\Delta}u_0\|_{\exp L^p} + \int_0^t \left\|\mathrm{e}^{(t-s)\Delta}(f(u(s)))\right\|_{\exp L^p} ds \\ &\leq \|\mathrm{e}^{t\Delta}u_0\|_{\exp L^p} + \int_0^t \kappa(t-s)\Big(\|f(u(s))\|_{L^1\cap L^q}\Big)\, ds \\ &\leq \|\mathrm{e}^{t\Delta}u_0\|_{\exp L^p} + \|f(u)\|_{L^{\infty}(0,\infty;(L^1\cap L^q))}\int_0^{\infty}\kappa(s)\, ds \\ &\leq \|\mathrm{e}^{t\Delta}u_0\|_{\exp L^p} + C\|f(u)\|_{L^{\infty}(0,\infty;(L^1\cap L^q))}. \end{aligned}$$

Hence by Part (i) of Proposition 3.2, we get

$$\|\Phi(u)\|_{L^{\infty}(0,\infty;\exp L^p)} \leq \|u_0\|_{\exp L^p} + C\|f(u)\|_{L^{\infty}(0,\infty;\, L^1\cap L^q)}.$$

It remains to estimate the nonlinearity $f(u)$ in L^r for $r = 1,\, q$. To this end, let us remark that

$$|f(u)| \leq C|u|^m\left(\mathrm{e}^{\lambda|u|^p} - 1\right) + C|u|^m. \tag{6-5}$$

By Hölder's inequality and Lemma 2.4, we have for $1 \le r \le q$ and since $m \ge p$,

$$\begin{aligned}\|f(u)\|_r &\le C\|u\|_{mr}^m + C\||u|^m(\mathrm{e}^{\lambda|u|^p}-1)\|_r \\ &\le C\|u\|_{mr}^m + C\|u\|_{2mr}^m\|\mathrm{e}^{\lambda|u|^p}-1\|_{2r} \\ &\le C\|u\|_{\exp L^p}^m\Big(\|\mathrm{e}^{\lambda|u|^p}-1\|_{2r}+1\Big).\end{aligned} \tag{6-6}$$

According to Lemma 2.6, and the fact that $u \in Y_M$, we have for $2q\lambda M^p \le 1$,

$$\|f(u)\|_{L^\infty(0,\infty;L^r)} \le CM^m. \tag{6-7}$$

Finally, we obtain

$$\begin{aligned}\|\Phi(u)\|_{L^\infty(0,\infty,\exp L^p)} &\le \|u_0\|_{\exp L^p} + CM^m \\ &\le \varepsilon + CM^m.\end{aligned}$$

Let $u,\ v$ be two elements of Y_M. By using (6-3) and Proposition 3.1, we obtain

$$\begin{aligned}t^\sigma\|\Phi(u)(t)-\Phi(v)(t)\|_a &\le t^\sigma\int_0^t \Big\|\mathrm{e}^{(t-s)\Delta}(f(u(s))-f(v(s)))\Big\|_a\,ds \\ &\le t^\sigma\int_0^t (t-s)^{-\frac{N}{2}(\frac{1}{r}-\frac{1}{a})}\,\|f(u(s))-f(v(s))\|_r\,ds \\ \le C\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}t^\sigma&\int_0^t (t-s)^{-\frac{N}{2}(\frac{1}{r}-\frac{1}{a})}\|(u-v)(|u|^{pk+m-1}+|v|^{pk+m-1})\|_r ds,\end{aligned}$$

where $1 \le r \le a$. We use the Hölder inequality with $\frac{1}{r}=\frac{1}{a}+\frac{1}{q}$ to obtain

$$\begin{aligned}t^\sigma\|\Phi(u)(t)-\Phi(v)(t)\|_a &\le C\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}t^\sigma\int_0^t (t-s)^{-\frac{N}{2}(\frac{1}{r}-\frac{1}{a})}\|u-v\|_a\times \\ &\qquad \||u|^{pk+m-1}+|v|^{pk+m-1}\|_q ds, \\ &\le C\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}t^\sigma\int_0^t (t-s)^{-\frac{N}{2}(\frac{1}{r}-\frac{1}{a})}\|u-v\|_a\times \\ &\qquad \Big(\|u\|_{q(pk+m-1)}^{pk+m-1}+\|v\|_{q(pk+m-1)}^{pk+m-1}\Big)\,ds.\end{aligned}$$

Using interpolation inequality where $\frac{1}{q(pk+m-1)} = \frac{\theta}{a} + \frac{1-\theta}{\rho}$, $p \le \rho < \infty$, we find that

$$t^{\sigma} \left\| \int_0^t \mathrm{e}^{(t-s)\Delta} \left(f(u) - f(v) \right) ds \right\|_a \le C \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} t^{\sigma} \int_0^t (t-s)^{-\frac{N}{2}\left(\frac{1}{r}-\frac{1}{a}\right)} \|u-v\|_a$$
$$\times \left(\|u\|_a^{(pk+m-1)\theta} \|u\|_{\rho}^{(pk+m-1)(1-\theta)} + \|v\|_a^{(pk+m-1)\theta} \|v\|_{\rho}^{(pk+m-1)(1-\theta)} \right) ds.$$

By Lemma 2.4, we obtain

$$t^{\sigma} \left\| \int_0^t \mathrm{e}^{(t-s)\Delta} \left(f(u) - f(v) \right) ds \right\|_a$$
$$\le C \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} t^{\sigma} \int_0^t (t-s)^{-\frac{N}{2}\left(\frac{1}{r}-\frac{1}{a}\right)} \|u-v\|_a \Gamma\left(\frac{\rho}{p}+1\right)^{\frac{(pk+m-1)(1-\theta)}{\rho}}$$
$$\times \left(\|u\|_a^{(pk+m-1)\theta} \|u\|_{\exp L^p}^{(pk+m-1)(1-\theta)} + \|v\|_a^{(pk+m-1)\theta} \|v\|_{\exp L^p}^{(pk+m-1)(1-\theta)} \right) ds. \tag{6-8}$$

Applying the fact that $u,\ v \in \ Y_M$ in (6-8), we see that

$$t^{\sigma} \left\| \int_0^t \mathrm{e}^{(t-s)\Delta} \left(f(u) - f(v) \right) ds \right\|_a$$
$$\le C d(u,v) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \Gamma\left(\frac{\rho}{p}+1\right)^{\frac{(pk+m-1)(1-\theta)}{\rho}} M^{pk+m-1}$$
$$\times t^{\sigma} \left(\int_0^t (t-s)^{-\frac{N}{2}\left(\frac{1}{r}-\frac{1}{a}\right)} s^{-\sigma(1+(pk+m-1)\theta)}\, ds \right)$$
$$\le C d(u,v) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \Gamma\left(\frac{\rho}{p}+1\right)^{\frac{(pk+m-1)(1-\theta)}{\rho}} M^{pk+m-1}$$
$$\times \mathcal{B}\left(1 - \frac{N}{2}\left(\frac{1}{r}-\frac{1}{a}\right), 1-\sigma\big(1+(pk+m-1)\theta\big)\right), \tag{6-9}$$

where the parameters $a,\ q,\ r,\ \theta = \theta_k,\ \rho = \rho_k$ are given by Lemma 2.7. For these parameters, using (2-5) and (2-6), we obtain that

$$\mathcal{B}\left(1 - \frac{N}{2}\left(\frac{1}{r}-\frac{1}{a}\right), 1-\sigma\big(1+(pk+m-1)\theta\big)\right) \le C. \tag{6-10}$$

Moreover, using (2-18)-(2-19)-(2-20) together with (2-8) and (2-7) gives

$$\Gamma\left(\frac{\rho_k}{p}+1\right)^{\frac{(pk+m-1)(1-\theta_k)}{\rho_k}} \leq C^k k!. \tag{6-11}$$

Combining (6-9), (6-10) and (6-11) we get

$$t^\sigma \left\| \int_0^t \mathrm{e}^{(t-s)\Delta} (f(u)-f(v))\, ds \right\|_a \leq C d(u,v) \sum_{k=0}^{\infty} (C\lambda)^k M^{pk+m-1}.$$

Hence, we get for M small,

$$t^\sigma \left\| \int_0^t \mathrm{e}^{(t-s)\Delta} (f(u)-f(v))\, ds \right\|_a \leq C M^{m-1} d(u,v).$$

The above estimates show that $\Phi : Y_M \to Y_M$ is a contraction mapping. By Banach's fixed point theorem, we thus obtain the existence of a unique u in Y_M with $\Phi(u) = u$. By (6-4), u solves the integral equation (6-1) with f satisfying (6-2). The estimate (1-9) follows from $u \in Y_M$. This terminates the proof of the existence of a global solution to (6-1) for $N > \frac{2p}{p-1}$.

We will now prove the statement (1-8). For $q \geq \frac{N}{2}$ and $q \geq p$, we have

$$\begin{aligned}
&\|u(t) - \mathrm{e}^{t\Delta} u_0\|_{\exp L^p} \\
&\quad \leq \int_0^t \|\mathrm{e}^{(t-s)\Delta} f(u(s))\|_{\exp L^p}\, ds \\
&\quad \leq C \int_0^t \|\mathrm{e}^{(t-s)\Delta} f(u(s))\|_p ds + C \int_0^t \|\mathrm{e}^{(t-s)\Delta} f(u(s))\|_\infty\, ds \\
&\quad \leq C \int_0^t \|f(u(s))\|_p ds + C \int_0^t (t-s)^{-\frac{N}{2q}} \|f(u(s))\|_q\, ds.
\end{aligned} \tag{6-12}$$

Now, let us estimate $\|f(u(t))\|_r$ for $r = p, q$. We have

$$|f(u)| \leq C|u|^m \mathrm{e}^{\lambda|u|^p}.$$

Therefore, we obtain

$$\|f(u)\|_r \leq C \||u|^m (\mathrm{e}^{\lambda|u|^p} - 1 + 1)\|_r.$$

By using Hölder inequality and Lemma 2.4, we obtain

$$\begin{aligned}
\|f(u)\|_r &\leq C\|u\|_{2mr}^m \|\mathrm{e}^{\lambda|u|^p} - 1\|_{2r} + \|u\|_{mr}^m \\
&\leq C\|u\|_{\exp L^p}^m \left(\|\mathrm{e}^{\lambda|u|^p} - 1\|_{2r} + 1\right).
\end{aligned}$$

Using Lemma 2.6 we conclude that

$$\|f(u)\|_r \leq C\|u\|^m_{\exp L^p}\left((2\lambda r M^p)^{\frac{1}{2r}}+1\right) \leq C\|u\|^m_{\exp L^p}. \tag{6-13}$$

Substituting (6-13) in (6-12), we have

$$\begin{aligned}\|u(t)-\mathrm{e}^{t\Delta}u_0\|_{\exp L^p} &\leq C\int_0^t \left[\|u\|^m_{\exp L^p}+(t-s)^{-\frac{N}{2q}}\|u\|^m_{\exp L^p}\right]ds\\ &\leq Ct\|u\|^m_{L^\infty(0,\infty;\,\exp L^p)}+Ct^{1-\frac{N}{2q}}\|u\|^m_{L^\infty(0,\infty;\,\exp L^p)}\\ &\leq C_1 t + C_2 t^{1-\frac{N}{2q}},\end{aligned}$$

where C_1, C_2 are finite positive constants. This gives

$$\lim_{t\longrightarrow 0}\|u(t)-\mathrm{e}^{t\Delta}u_0\|_{\exp L^p}=0,$$

and proves statement (1-8).

Finally the fact that $u(t)\to u_0$ as $t\to 0$ in the weak* topology can be done as in Ioku [2011]. So we omit the proof here.

References

Robert A. Adams and John J. F. Fournier (2003). *Sobolev spaces*. Second. Vol. 140. Pure and Applied Mathematics (Amsterdam). Elsevier/Academic Press, Amsterdam, pp. xiv+305. MR: 2424078 (cit. on p. 2436).

Z. W. Birnbaum and W. Orlicz (1931). "Über die Verallgemeinerung des Begriffes der zueinander Konjugierten Potenzen". *Studia Mathematica* 3, pp. 1–67 (cit. on p. 2432).

Haïm Brezis and Thierry Cazenave (1996). "A nonlinear heat equation with singular initial data". *J. Anal. Math.* 68, pp. 277–304. MR: 1403259 (cit. on p. 2432).

Thierry Cazenave and Fred B. Weissler (1998). "Asymptotically self-similar global solutions of the nonlinear Schrödinger and heat equations". *Math. Z.* 228.1, pp. 83–120. MR: 1617975.

Giulia Furioli, Tatsuki Kawakami, Bernhard Ruf, and Elide Terraneo (2017). "Asymptotic behavior and decay estimates of the solutions for a nonlinear parabolic equation with exponential nonlinearity". *J. Differential Equations* 262.1, pp. 145–180. MR: 3567484 (cit. on p. 2432).

Alain Haraux and Fred B. Weissler (1982). "Nonuniqueness for a semilinear initial value problem". *Indiana Univ. Math. J.* 31.2, pp. 167–189. MR: 648169 (cit. on p. 2432).

Slim Ibrahim, Rym Jrad, Mohamed Majdoub, and Tarek Saanouni (2014). "Local well posedness of a 2D semilinear heat equation". *Bull. Belg. Math. Soc. Simon Stevin* 21.3, pp. 535–551. MR: 3250777 (cit. on pp. 2432, 2442).

Slim Ibrahim, Mohamed Majdoub, and Nader Masmoudi (2011). "Well- and ill-posedness issues for energy supercritical waves". *Anal. PDE* 4.2, pp. 341–367. MR: 2859857.

Norisuke Ioku (2011). "The Cauchy problem for heat equations with exponential nonlinearity". *J. Differential Equations* 251.4-5, pp. 1172–1194. MR: 2812586 (cit. on pp. 2432, 2435, 2446, 2458).

Norisuke Ioku, Bernhard Ruf, and Elide Terraneo (2015). "Existence, non-existence, and uniqueness for a heat equation with exponential nonlinearity in $\mathbb{R}^2$". *Math. Phys. Anal. Geom.* 18.1, Art. 29, 19. MR: 3414493 (cit. on pp. 2432–2434, 2437, 2453).

Carlos E. Kenig, Gustavo Ponce, and Luis Vega (2000). "Global well-posedness for semi-linear wave equations". *Comm. Partial Differential Equations* 25.9-10, pp. 1741–1752. MR: 1778778.

M. Majdoub, S. Otsmane, and S. Tayachi (2018). "Local well-posedness and global existence for the biharmonic heat equation with exponential nonlinearity". *Advances in Differential Equations* 23, pp. 489–522 (cit. on pp. 2432, 2435, 2442, 2447, 2448, 2450).

Carl E. Mueller and Fred B. Weissler (1985). "Single point blow-up for a general semilinear heat equation". *Indiana Univ. Math. J.* 34.4, pp. 881–913. MR: 808833 (cit. on p. 2450).

M. M. Rao and Z. D. Ren (2002). *Applications of Orlicz spaces*. Vol. 250. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, pp. xii+464. MR: 1890178 (cit. on p. 2436).

Bernhard Ruf and Elide Terraneo (2002). "The Cauchy problem for a semilinear heat equation with singular initial data". In: *Evolution equations, semigroups and functional analysis (Milano, 2000)*. Vol. 50. Progr. Nonlinear Differential Equations Appl. Birkhäuser, Basel, pp. 295–309. MR: 1944169 (cit. on pp. 2432, 2439).

Seifeddine Snoussi, Slim Tayachi, and Fred B. Weissler (2001). "Asymptotically self-similar global solutions of a general semilinear heat equation". *Math. Ann.* 321.1, pp. 131–155. MR: 1857372 (cit. on p. 2435).

Philippe Souplet and Slim Tayachi (2016). "Single-point blow-up for parabolic systems with exponential nonlinearities and unequal diffusivities". *Nonlinear Anal.* 138, pp. 428–447. MR: 3485155 (cit. on p. 2435).

Neil S. Trudinger (1967). "On imbeddings into Orlicz spaces and some applications". *J. Math. Mech.* 17, pp. 473–483. MR: 0216286 (cit. on p. 2436).

Fred B. Weissler (1979). "Semilinear evolution equations in Banach spaces". *J. Funct. Anal.* 32.3, pp. 277–296. MR: 538855 (cit. on p. 2432).

– (1980). "Local existence and nonexistence for semilinear parabolic equations in L^p". *Indiana Univ. Math. J.* 29.1, pp. 79–102. MR: 554819 (cit. on pp. 2432, 2435).

– (1981). "Existence and nonexistence of global solutions for a semilinear heat equation". *Israel J. Math.* 38.1-2, pp. 29–40. MR: 599472 (cit. on pp. 2432, 2450).

Received 2017-11-30.

Mohamed Majdoub
Department of Mathematics
College of Science
Imam Abdulrahman Bin Faisal University
P.O. Box 1982, Dammam
Saudi Arabia
mohamed.majdoub@fst.rnu.tn
mmajdoub@iau.edu.sa

Slim Tayachi
slim.tayachi@fst.rnu.tn
slimtayachi@gmail.com
Slim Tayachi
Université de Tunis El Manar
Faculté des Sciences de Tunis
Département de Mathématiques
Laboratoire équations aux dérivées partielles (LR03ES04)
2092 Tunis
Tunisie

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2457–2484)

INTERACTION OF SOLITONS FROM THE PDE POINT OF VIEW

YVAN MARTEL

Abstract

We review recent results concerning the interactions of solitary waves for several universal nonlinear dispersive or wave equations. Though using quite different techniques, these results are partly inspired by classical papers based on the inverse scattering theory for integrable models.

1 Introduction

Pioneering numerical experiments of Fermi, Pasta, and Ulam [1955] in 1955, and of Zabusky and Kruskal [1965] in 1965, revealed unexpected phenomena related to the interactions of nonlinear waves[1]. Shortly thereafter, the inverse scattering theory and its generalizations, developed by many influential mathematicians such as Ablowitz, Kaup, Newell, and Segur [1974], Gardner, C. S. Greene, Kruskal, and Miura [1967], Gardner, J. M. Greene, Kruskal, and Miura [1974], Lax [1968], Miura [1976], Miura, Gardner, and Kruskal [1968], and Zakharov and Shabat [1971], provided a rigorous ground and a unified approach to these observations. It led very rapidly to an accurate and deep understanding of remarkable properties of several universal nonlinear models, referred to as *completely integrable*, such as for example, the Korteweg-de Vries equation, the one dimensional cubic Schrödinger equation and the sine-Gordon equation. It has created a very active and inspiring field of research since then[2]. Among the most notable achievements of this theory, we mention

MSC2010: primary 35B40; secondary 37K40, 35Q51, 35Q53, 35Q55, 35L71.
Keywords: soliton, multi-soliton, asymptotic behavior of solutions, elastic and non-elastic collisions, nonlinear Schrödinger equation, generalized Korteweg-de Vries equation, semilinear wave equation.

[1] We refer to Chapter 8 of Dauxois and Peyrard [2010] for details on this discovery and on the relation between the model considered in Fermi, Pasta, and Ulam [1955] and the KdV equation. It is quite rightly suggested in Dauxois and Peyrard [2010] to recognize the work of M. Tsingou, contributor to the numerical computations of Fermi, Pasta, and Ulam [1955].

[2] See for example Lamb [1980] and Chapter 7 of Dauxois and Peyrard [2010] for synthetic presentations of the inverse scattering transform.

(i) the existence of *infinitely many conservation laws*;

(ii) the *purely elastic nature of the collision* of any number of solitary waves, which means that the interacting solitary waves recover their exact shape and velocity after a collision. Solitary waves enjoying such remarkable property were called *solitons*[3];

(iii) the *decomposition into solitons*, saying that from any solution should emerge in large time a sum of nonlinear states, such as solitons, plus a dispersive part.

These rigorous mathematical facts are known to be physically relevant in numerous contexts, though sometimes under less extreme forms. For example, in several practical applications or for more elaborate nonlinear models, the collision of nonlinear waves is not purely elastic and some loss of energy takes place during collisions[4]. This reveals that the inverse scattering theory is restricted to models with specific algebraic structure and despite many extensions to nearly integrable systems (see *e.g.* Kivshar and Malomed [1989]), it cannot be applied to general nonlinear models.

In view of the beautiful achievements of the integrability theory but also of its inevitable limitations, it appeared necessary to investigate similar questions for general nonlinear models with solitary waves using tools from the theory of partial differential equations. In these notes, we review some results on interactions of solitary waves obtained for models that are not close to any known integrable equation, such as the generalized Korteweg-de Vries equation, the nonlinear Schrödinger equation in any space dimension, the ϕ^4 equation and the nonlinear wave equation.

Mainly in the 80s, the solitary wave theory, proving existence, uniqueness, symmetry and stability properties of nonlinear waves, was successfully developed using the elliptic theory, ODE analysis and general variational arguments, at least for *ground states* (see Section 3). More recently, *asymptotic stability* results appeared (see Section 4). Then, energy type arguments extending the elliptic theory have allowed to consider several solitary waves in *weak interactions*, *i.e.* cases where the soliton dynamics is only slightly perturbed by the interactions. Pushing the perturbative analysis one step forward, some examples of *strong interactions* have also been exhibited; the solitons are still distant, but their dynamics is substantially modified by the interactions (see Section 5). Next, we review the few recent cases where a version of the *soliton resolution conjecture* was proved for non-integrable wave models in Section 6. Finally, we discuss in Section 7 some situations where collisions were proved to be *inelastic*.

This review points out that despite some impressive and surprizing recent progress, notably on the soliton resolution conjecture, most of the questions raised above on the interaction of solitary waves remain open for general nonlinear models.

[3]This term is now commonly used for solitary waves even in the non-integrable context.

[4]We refer to Craig, Guyenne, Hammack, Henderson, and Sulem [2006] for a discussion on this topic.

2 Integrable equations

In this section, we briefly highlight some results from the inverse scattering theory that inspired mathematical research much beyond their range of applicability.

2.1 KdV solitons and multi-solitons. For the Korteweg-de Vries equation[5]

$$\partial_t u + \partial_x(\partial_x^2 u + u^2) = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{R}, \tag{1}$$

the inverse scattering transform led to a very striking property which is the existence of exact multi-soliton solutions (see *e.g.* Hirota [1971], Miura [1976], and Whitham [1974]).

Let $Q(x) = \frac{3}{2}\cosh^{-2}(\frac{x}{2})$ be the unique positive even solution of $Q'' + Q^2 = Q$, and for $c > 0$, let $Q_c(x) = cQ(\sqrt{c}x)$. Then, for any $c > 0$, $\sigma \in \mathbb{R}$, the function defined by $u(t,x) = Q_c(x - ct - \sigma)$ is a solution of (1), called *soliton*, traveling with speed c.

Solutions containing an arbitrary number of such solitons (called *multi-solitons*) have been obtained by the inverse scattering theory.

Theorem 1 (Multi-solitons for KdV, Hirota [1971] and Miura [1976])**.** *Let $K \in \mathbb{N}$, $K \geq 2$. Let $0 < c_K < \cdots < c_1$ and $\sigma_1^-, \ldots, \sigma_K^- \in \mathbb{R}$. There exist $\sigma_1^+, \ldots, \sigma_K^+ \in \mathbb{R}$ and an explicit solution u of* (1) *such that*

$$\lim_{t\to\pm\infty} \left\| u(t) - \sum_{k=1}^{K} Q_{c_k}\left(\cdot - c_k t - \sigma_k^{\pm}\right)\right\|_{H^1} = 0.$$

The most remarkable fact is that all the solitons recover exactly the same sizes and speeds after the collision. Moreover, the values of σ_k^+ are explicit. It it interesting to recall that the multi-soliton behavior, even in the simple case of two solitons, differs qualitatively according to the relative sizes of the solitons. We refer to Lemma 2.3 in Lax [1968] for a definition of the three *Lax categories* of two-solitons and to Zabusky and Kruskal [1965] for a previous formal discussion. In particular, if their sizes are close (*i.e.* $c_1 \sim c_2$), the two solitons never cross, but rather repulse each other at a large distance (this is category (c) in Lax [1968]). See Sections 7.1 to 7.3.

2.2 Decomposition into solitons for KdV. The multi-soliton behavior is fundamental for general solutions of the KdV equation as shown by the following decomposition result.

Theorem 2 (Decomposition into solitons, Eckhaus and Schuur [1983] and Schuur [1986])**.** *Let u_0 be a C^4 function such that for any $j \in \{0, ..., 4\}$, for all $x \in \mathbb{R}$, $\left|(\partial^j u_0/\partial x^j)(x)\right| \lesssim$*

[5] We refer to Chapter 1 of Dauxois and Peyrard [2010] for historical facts on this equation and its applications to Physics.

$\langle x\rangle^{-10}$. *Let u be the solution of* (1) *corresponding to u_0. Then, there exist $K \in \mathbb{N}$, $\sigma_1, \dots, \sigma_K \in \mathbb{R}$ and $c_1 > \cdots > c_K > 0$ such that, for all $x > 0$,*

$$\lim_{t\to+\infty} \left\{ u(t,x) - \sum_{k=1}^{K} Q_{c_k}(x - c_k t - \sigma_k) \right\} = 0.$$

This result has a rich history, see Ablowitz, Kaup, Newell, and Segur [1974], Cohen [1979], Dauxois and Peyrard [2010], Dodd, Eilbeck, Gibbon, and Morris [1982], Eckhaus and Schuur [1983], Kruskal [1974], Lax [1968], Schuur [1986], and Zabusky and Kruskal [1965] and the references therein. Note that if some space decay is necessary to apply the inverse scattering transform, the decay assumption on the initial data in the above result is not optimal. Note also that the asymptotic behavior of the solution is described for $x > 0$ (see results in Schuur [1986] for slight improvement). For the region $x < 0$, see Deift, Venakides, and Zhou [1994], Eckhaus and Schuur [1983], and Schuur [1986] and references therein.

Last, we mention that the modified KdV equation (i.e. the KdV equation with a cubic nonlinearity) is also an integrable model that enjoys most of the properties of the KdV equation, like the infinitely many conservation laws and the existence of pure multi-soliton solutions (see *e.g.* Miura [1976]). Actually, it even has a richer family of exceptional solutions: breather solutions (see Alejo and Muñoz [2013], Lamb [1980], and Wadati [1973]) and dipole solitons, *i.e.* special multi-solitons where solitons are distant like $C \log t$ (see Gorshkov and Ostrovsky [1981], Karpman and Solovev [1981], and Wadati and Ohkuma [1982]). This complicates any possible soliton resolution conjecture on this equation (see Schuur [1986]).

2.3 One dimensional cubic NLS.

The 1D cubic nonlinear Schrödinger equation

$$i\partial_t u + \partial_x^2 u + |u|^2 u = 0, \quad (t,x) \in \mathbb{R}\times\mathbb{R}, \tag{2}$$

is also an integrable equation, widely studied for its numerous physical applications and remarkable mathematical properties. See *e.g.* Dauxois and Peyrard [2010], Deift and Trubowitz [1979], Deift and Zhou [1993], Dodd, Eilbeck, Gibbon, and Morris [1982], Faddeev and Takhtajan [2007], Novokšenov [1980], Olmedilla [1987], Yang [2010], Zabusky and Kruskal [1965], Zakharov and Manakov [1976], and Zakharov and Shabat [1971].

Here, we denote $Q(x) = \sqrt{2}\cosh^{-1}(x)$ the unique positive even solution of $Q'' + Q^3 = Q$, and for any $c > 0$, $Q_\lambda(x) = \sqrt{c}Q(\sqrt{c}x)$. Then, for any $c > 0$, $\beta \in \mathbb{R}$, $\sigma \in \mathbb{R}$ and $\gamma \in \mathbb{R}$,

$$u(t,x) = Q_c(x - \beta t - \sigma)e^{i\Gamma(t,x)}, \quad \Gamma(t,x) = \frac{1}{2}(\beta\cdot x) - \frac{1}{4}|\beta|^2 t + ct + \gamma,$$

is a solitary wave of (2), moving on the line $x = \sigma + \beta t$ and also called *soliton*.

As the KdV equation, the 1D cubic NLS admits explicit multi-solitons. However, the possible behaviors of multi-solitons is richer for NLS. In addition to multi-solitons distant like Ct, which is the generic situation, the equation also admits multi-solitons where the distance between some solitons is $C \log t$ (see Olmedilla [1987] and Zakharov and Shabat [1971]; this requires solitons of exactly the same size, like for mKdV) and solutions where some solitons are staying at a finite distance from each other for all time (see Yang [2010] and Zakharov and Shabat [1971]). As for mKdV, the presence of such multi-solitons complicates any general decomposition result but does not prevent it. For such questions, we refer to the recent work Borghese, Jenkins, and McLaughlin [2016] and its references.

2.4 The sine-Gordon equation. The sine-Gordon equation

$$\partial_t^2 u - \partial_x^2 u + \sin u = 0, \quad (t, x) \in \mathbb{R} \times \mathbb{R},$$

was also widely studied as a physically relevant and completely integrable model (see *e.g.* Dauxois and Peyrard [2010], Dodd, Eilbeck, Gibbon, and Morris [1982], and Lamb [1980]). This equation has an explicit kink solution $S(x) = 4 \arctan(e^x)$. It also has other exceptional solutions, like time-periodic *wobbling kinks* (see Cuenda, Quintero, and Sánchez [2011] and Segur [1983]), and *breathers* (see Lamb [1980]).

2.5 Other integrable models and nearly integrable models. For the derivative NLS equation, we refer to Jenkins, J. Liu, Perry, and Sulem [2017] and its references. For the KP-I equation, see Lamb [1980]. For integrable models set on the torus, see Kuksin [2000] and references therein.

Several *nearly integrable equations* have also been studied in the context of the theory of inverse scattering. We refer to Dauxois and Peyrard [2010], Deift and Zhou [2002], Kivshar and Malomed [1989], and Yang [2010] and to the references therein.

2.6 Formal works and numerical simulations. Note that shortly after the development of the inverse scattering and the discovery of explicit multi-solitons, other approaches appeared, like in Ei and Ohta [1994], Gorshkov and Ostrovsky [1981], and Karpman and Solovev [1981], to investigate possible multi-soliton behaviors for integrable or non-integrable models. Such papers focus on the modulation equations of the parameters of the solitons and lack the analysis of the error terms, but they aim at justifying formally multi-solitons behaviors beyond any integrability property or proximity to integrable equations. In particular, as for the rigorous results presented in Section 5 below, they are *asymptotic results*, restricted to cases where the distances between the various solitons are large enough.

Theoretical and numerical works have been developed in parallel. As mentioned in the Introduction, the subject started with two fundamental numerical experiments presented in Fermi, Pasta, and Ulam [1955] and Zabusky and Kruskal [1965]. Since then, there has been an intense activity on studying solitary waves interactions from the numerical point of view. We refer to Craig, Guyenne, Hammack, Henderson, and Sulem [2006] which compares KdV multi-solitons, the water wave problem from the numerical point of view and real experiments on waves generated in water tanks. For Klein-Gordon equations, we refer to Ablowitz, Kruskal, and Ladik [1979]. We also refer to Bona, Pritchard, and Scott [1980], Dauxois and Peyrard [2010], Hammack, Henderson, Guyenne, and Yi [2004], Li and Sattinger [1999], Shih [1980], and Yang [2010] and references therein. One of the main questions studied by numerical experiments is the elastic versus inelastic character of the collisions of nonlinear waves.

3 Nonlinear models with solitary waves

In these notes, we consider four typical nonlinear models and work with the notion of solution in the energy space.

3.1 The generalized Korteweg-de Vries equation. Consider the following 1D model, for any integer $p \geq 2$,

$$\partial_t u + \partial_x(\partial_x^2 u + u^p) = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{R}. \tag{3}$$

As seen before, the case $p = 2$ corresponds to the KdV equation and $p = 3$ to the mKdV equation, which are both completely integrable.

The mass and energy

$$\int u^2(t), \quad \int \left(\frac{1}{2}u_x^2(t) - \frac{1}{p+1}u^{p+1}(t)\right)$$

are formally conserved for solutions of (3). We refer to Kenig, Ponce, and Vega [1993] for the local well-posedness of the Cauchy problem in the energy space H^1 (see also Kato [1983]). For $1 < p < 5$, all solutions in H^1 are global and bounded, and the problem is called *sub-critical*. For $p = 5$, the problem is mass critical (blow up solution do exist, see Martel, Merle, and Raphaël [2014] and references therein) and $p > 5$ correspond to the super-critical case. The notion of criticality corresponds to the scaling invariance of equation (3): indeed, if $u(t,x)$ is solution then for any $c > 0$, $u_c(t,x) = c^{\frac{1}{p-1}} u(c^{\frac{3}{2}}t, c^{\frac{1}{2}}x)$ is also solution and $\|u_c(t)\|_{L^2} = c^{\frac{1}{p-1}-\frac{1}{4}}\|u(t)\|_{L^2}$.

Let Q be the unique (up to sign change if p is odd) non-trivial even solution of $Q'' + Q^p = Q$ on $\mathbb{R}$, explicitely given by

$$Q(x) = \left(\frac{p+1}{2}\right)^{\frac{1}{p-1}} \cosh^{-\frac{2}{p-1}}\left(\frac{p-1}{2}x\right).$$

For $c > 0$, let $Q_c(x) = c^{\frac{1}{p-1}} Q(c^{\frac{1}{2}}x)$. Note that these formulas for Q and Q_c generalize the previous ones given for $p = 2$ and $p = 3$. As before, solitary waves (also called solitons by abuse of terminology) are solutions of (3) of the form $u(t,x) = Q_c(x-ct-\sigma)$, for any $c > 0$ and $\sigma \in \mathbb{R}$.

The *orbital stability of solitons* with respect to small perturbations in the energy space H^1 is known in the sub-critical case.

Theorem 3 (Stability of the soliton for sub-critical gKdV Benjamin [1972], Bona [1975], Cazenave and Lions [1982], and Weinstein [1985, 1986])**.** *Let $1 < p < 5$. For all $\epsilon > 0$, there exists $\delta > 0$, such that if $\|u_0 - Q\|_{H^1} \le \delta$, then the solution u of* (3) *with initial data u_0 satisfies, for all $t \in \mathbb{R}$, $\|u(t, . + \sigma(t)) - Q\|_{H^1} \le \epsilon$, for some function σ.*

In contrast, solitons are unstable in the critical and super-critical case $p \ge 5$. Note that the instability phenomenon is quite different in the critical case (linear stability holds and the nonlinear instability is related to the scaling parameter) and in the super-critical case (linear exponential instability). See Bona, Souganidis, and Strauss [1987], Cazenave [2003], Grillakis, Shatah, and Strauss [1987], Martel and Merle [2000], and Pego and Weinstein [1992].

3.2 The nonlinear Schrödinger equation.

Recall the nonlinear NLS equation

$$i\partial_t u - \Delta u - |u|^{p-1}u = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^d. \tag{4}$$

We consider the case $p > 1$ for $d = 1, 2$, and $1 < p < \frac{d+2}{d-2}$ for $d \ge 3$. For $d = 1$ and $p = 3$, the model is completely integrable, as seen before. Note that, similarly as for the gKdV equation, $p = 1 + \frac{4}{d}$ corresponds to L^2 criticality, while for $d \ge 3$, $p = \frac{d+2}{d-2}$ corresponds to $\dot{H}^1$ criticality.

The mass, energy and momentum

$$\int |u(t)|^2, \quad \int \left(\frac{1}{2}|\nabla u(t)|^2 - \frac{1}{p+1}|u(t)|^{p+1}\right), \quad \Im\left(\int \nabla u(t)\bar{u}(t)\right)$$

are formally conserved for solutions of (4). We refer to Cazenave [2003], Ginibre and Velo [1979], and Tao [2006] for the local well-posedness of the Cauchy problem in the energy space H^1.

We denote by Q the unique positive radially symmetric H^1 solution of $\Delta Q + |Q|^{p-1}Q = Q$ on $\mathbb{R}^d$ (the function Q is called the *ground state*; see existence and uniqueness results in Berestycki and Lions [1983], Cazenave [2003], Kwong [1989], and Tao [2006]). For $c > 0$, let $Q_c(x) = c^{\frac{1}{p-1}} Q(c^{\frac{1}{2}} x)$. Note that this is a further generalization of the notation for gKdV, for any space dimension $d \geq 1$. For $d \geq 2$, ground states are no longer explicit, but their properties are well-understood (see references above). Then, for any $c > 0$, $\beta \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^d$ and $\gamma \in \mathbb{R}$, the function u defined by

$$u(t,x) = Q_c(x - \beta t - \sigma) e^{i\Gamma(t,x)} \quad \text{where} \quad \Gamma(t,x) = \frac{1}{2}(\beta \cdot x) - \frac{1}{4}|\beta|^2 t + ct + \gamma,$$

is a traveling wave of (4), with speed β.

The stability and instability properties of solitary waves of NLS are similar: stability in the L^2 sub-critical case, and instability in the critical and super-critical cases. We refer to Cazenave [2003], Cazenave and Lions [1982], Grillakis [1990], Grillakis, Shatah, and Strauss [1987], and Weinstein [1985] for details.

3.3 The ϕ^4 equation. We consider the ϕ^4 model (see *e.g.* Dauxois and Peyrard [2010] and Manton and Sutcliffe [2004])

$$\partial_t^2 \phi - \partial_x^2 \phi - \phi + \phi^3 = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{R}. \tag{5}$$

Recall that the energy

$$E(\phi, \partial_t \phi) = \int \frac{1}{2}|\partial_t \phi|^2 + \frac{1}{2}|\partial_x \phi|^2 + \frac{1}{4}\left(1 - |\phi|^2\right)^2$$

is formally conserved along the flow. The *kink*, defined by $H(x) = \tanh\left(x/\sqrt{2}\right)$ is the unique (up to sign change), bounded, odd solution of the equation $-H'' = H - H^3$ on $\mathbb{R}$. We recall that the orbital stability of the kink with respect to small perturbations in the energy space has been proved in Henry, Perez, and Wreszinski [1982] using mainly the energy conservation. This model is analogue to the sine-Gordon equation, but it is not completely integrable and breathers solutions or woobling kinks are not expected to exist.

3.4 The energy critical nonlinear wave equation. For space dimensions $d \geq 3$, we consider the following nonlinear wave equation,

$$\partial_t^2 u = \Delta u + |u|^{\frac{4}{d-2}} u, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^d. \tag{6}$$

We denote

$$E(u,v) = \int \left(\frac{1}{2}|\nabla u|^2 + \frac{1}{2}v^2 - \frac{1}{6}|u|^6\right)$$

so that the energy of a solution $(u, \partial_t u)$ of (6), defined by $E(u, \partial_t u)$, is formally conserved by the flow. Concerning the Cauchy problem in $\dot{H}^1 \times L^2$ for the energy critical wave equation, we refer to Kenig and Merle [2008] and the references given therein. As before, the notion of criticality is related to the scaling invariance: if $u(t, x)$ is a solution, then for any $\lambda > 0$,

$$u_\lambda(t, x) = \frac{1}{\lambda^{\frac{d-2}{2}}} u\left(\frac{t}{\lambda}, \frac{x}{\lambda}\right) \quad \text{is also solution and} \quad \|\nabla u_\lambda\|_{L^2} = \|\nabla u\|_{L^2}.$$

Here, solitary waves are stationary solutions $W \in \dot{H}^1$ satisfying $\Delta W + |W|^{\frac{4}{d-2}} W = 0$, and traveling waves obtained as Lorentz transforms of such solutions. For $\boldsymbol{\ell} \in \mathbb{R}^d$, $|\boldsymbol{\ell}| < 1$, we denote

$$W_{\boldsymbol{\ell}}(x) = W\left(\left(\frac{1}{\sqrt{1 - |\boldsymbol{\ell}|^2}} - 1\right) \frac{\boldsymbol{\ell}(\boldsymbol{\ell} \cdot x)}{|\boldsymbol{\ell}|^2} + x\right),$$

so that $u(t, x) = W_{\boldsymbol{\ell}}(x - \boldsymbol{\ell} t)$ is solution of (6). As for the NLS equation, we consider only *ground states* solitary waves, *i.e.* solutions of the above elliptic equation explicitely given by

$$W(x) = \left(1 + \frac{|x|^2}{d(d-2)}\right)^{-\frac{d-2}{2}}.$$

As solutions of the evolution equation (6), they are unstable with respect to perturbation of the initial data with one direction of exponential instability (see Duyckaerts and Merle [2008] and Grillakis [1990]).

4 Asymptotic stability

We recall briefly some results of asymptotic stability of solitons.

4.1 Asymptotic stability for gKdV solitons.

Theorem 4 (Asymptotic stability of the gKdV soliton in H^1, Martel and Merle [2001]**).** *Let $p = 2, 3, 4$. For any $\beta > 0$, there exists $\delta = \delta(\beta) > 0$ such that the following is true. Let $u_0 \in H^1$ be such that $\|u_0 - Q\|_{H^1} \le \delta$. Then, the global solution u of* (3) *with initial data u_0 satisfies*

$$\lim_{t \to +\infty} \|u(t) - Q_{c^+}(\cdot - \sigma(t))\|_{H^1(x > \beta t)} = 0,$$

for some $c^+ > 0$ with $|c^+ - 1| \lesssim \delta$ and some C^1 function σ such that $\lim_{+\infty} \sigma' = c^+$.

We refer to Pego and Weinstein [1994] for the first result of asymptotic stability of gKdV solitons. Theorem 1 claims strong convergence in H^1 in the region $x > \beta t$. Strong convergence in $H^1(\mathbb{R})$ is never true since it would imply by stability that u is exactly a soliton. The region where convergence is obtained in Theorem 1 is sharp since one can construct a solution which behaves asymptotically as $t \to +\infty$ as the sum $Q(x-t) + Q_c(x-ct)$, where $0 < c \neq 1$ is arbitrary (see Martel [2005], Miura [1976], and Wadati and Toda [1972]). In particular, choosing $c \ll 1$, the H^1 norm of $Q_c(x-ct)$ is small, and this soliton travels on the line $x = ct$. This explains the necessity of a positive β in the convergence result. We also refer to the survey Tao [2009]. For $p = 4$, the result has been completed in Kenig and Martel [2009] and Tao [2007] showing that the residue scatters in a Besov space close to the critical Sobolev space $\dot{H}^{-1/6}$. For $p = 3$, we refer to Germain, Pusateri, and Rousset [2016] for a full asymptotic stability statement.

4.2 Asymptotic stability for NLS equations. In the context of the nonlinear Schrödinger equation, pioneering results on asymptotic stability of traveling waves are Buslaev and Perelman [1992, 1995] and Soffer and Weinstein [1990, 1992]. These papers initiated the method of separating modes and using dispersive estimates with potential, under assumptions on the spectrum of the linearized operator.

This question has then been extensively studied, for the NLS equation with or without potential and for various nonlinearities, see *e.g.* Buslaev and Sulem [2003], Cuccagna [2014], Gustafson, Nakanishi, and Tsai [2004], Nakanishi and Schlag [2011], Rodnianski, Schlag, and Soffer [2005], Rodnianski, Schlag, and Soffer [2003], Schlag [2006], and Schlag [2007, 2009] as typical papers, and their references. Most of these works require specific assumptions, like spectral assumptions or suitable dispersive estimates for equations with unknown potential, a suitable Fermi Golden Rule or flatness conditions on the nonlinearities at 0. It follows that no result of asymptotic stability is fully proved for any pure power NLS equation without potential with stable solitons, except for the integrable cubic 1D NLS treated in Cuccagna and Pelinovsky [2014].

In larger dimensions, or higher order nonlinearities, the solitons are unstable. The notion of *conditional asymptotic stability* and the construction of *center stable manifolds* then become relevant. For the focusing 3D cubic NLS equation (which is an $\dot{H}^{\frac{1}{2}}$ critical equation with exponentially unstable solitons) the theory has been especially well-developed, at least in the radial case, in Beceanu [2008, 2012], Costin, Huang, and Schlag [2012], Nakanishi and Schlag [2011], Schlag [2006], and Schlag [2009]. In particular, spectral assumptions implying the desired dispersive estimates for the linearized equation around the soliton have been checked, first numerically and then rigorously by computer assisted proof (see Costin, Huang, and Schlag [2012] and references therein).

4.3 Asymptotic stability of the ϕ^4 kink. The asymptotic stability of the kink H by the ϕ^4 flow (5) is known in the case of odd perturbations in the energy space. Note that for odd initial data, the corresponding solution of (5) is odd. Rewrite $\phi = H + u$. Then, one has

$$\partial_t^2 u - \partial_x^2 u - u + 3H^2 u + 3Hu^2 + u^3 = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{R}. \tag{7}$$

Theorem 5 (Asymptotic stability of the kink by odd perturbations, Kowalczyk, Martel, and Muñoz [2017]). *There exists $\delta > 0$ such that for any odd $(u_0, u_1) \in H^1 \times L^2$ with $\|(u_0,u_1)\|_{H^1\times L^2} \le \delta$, the solution $(u, \partial_t u)$ of (7) with initial data (u_0,u_1) satisfies, for any bounded interval $I \subset \mathbb{R}$,*

$$\lim_{t\to\pm\infty} \|(u,\partial_t u)(t)\|_{H^1(I)\times L^2(I)} = 0.$$

As for gKdV, if a solution u of (7) satisfies $\lim_{t\to+\infty} \|(u,\partial_t u)(t)\|_{H^1\times L^2} = 0$, then by orbital stability Henry, Perez, and Wreszinski [1982], $u(t) \equiv 0$, for all $t \in \mathbb{R}$. Thus the local result is in some sense optimal.

For previous related results, we refer to Kopylova and Komech [2011a,b] where the asymptotic stability of the kink is studied for the 1D equation $\partial_t^2 u - \partial_x^2 u + F(u) = 0$, under specific assumptions on F (not including the ϕ^4 model) and to Cuccagna [2008], where the stability and asymptotic stability of the one dimensional kink for the ϕ^4 model, subject to *localized three dimensional* perturbations is studied. We also refer the references in Cuccagna [2008], Kopylova and Komech [2011a,b], and Kowalczyk, Martel, and Muñoz [2017] for related works on scattering of small solutions to Klein-Gordon equations. See also the review Kowalczyk, Martel, and Muñoz [2016-2017] and references therein.

4.4 Blow up profile for the critical wave equation. Recall that Kenig and Merle [2008] provides a classification of all possible behaviors (blow up or scattering) of solutions of (6) whose initial data (u_0,u_1) satisfies $E(u_0,u_1) < E(W,0)$. Next, Duyckaerts and Merle [2008] studies the *threshold case* $E(u_0,u_1) = E(W,0)$ and constructs the stable manifold around W. Then, Duyckaerts, Kenig, and Merle [2011, 2012] proved the following result for solutions slightly above the threshold.

Theorem 6 (Blow up profile for 3D critical NLW, Duyckaerts, Kenig, and Merle [2011, 2012]). *Let $d = 3$. There exists $\delta > 0$ such that if u is a solution of (6) blowing up in finite time $T > 0$ and satisfying the bound*

$$\sup_{[0,T)} \left(\|\nabla u(t)\|_{L^2} + \frac{1}{2}\|\partial_t u(t)\|_{L^2} \right) \le \|\nabla W\|_{L^2} + \delta,$$

then

$$\lim_{t\uparrow T}\left\|(u(t),\partial_t u(t))-(v_0,v_1)\mp\left(\frac{1}{\lambda^{\frac12}(t)}W_{\boldsymbol{\ell}}\left(\frac{\cdot-\sigma(t)}{\lambda(t)}\right),-\frac{1}{\lambda^{\frac32}(t)}(\boldsymbol{\ell}\cdot\nabla W_{\boldsymbol{\ell}})\left(\frac{\cdot-\sigma(t)}{\lambda(t)}\right)\right)\right\|_{L^2}=0$$

for some σ, λ and $\boldsymbol{\ell}\in\mathbb{R}^3$, $|\boldsymbol{\ell}|<1$ and $(v_0,v_1)\in\dot H^1\times L^2$.

We see that the family $\{\pm W_{\boldsymbol{\ell}}\}$ is the universal blow up profile. We refer to the original papers for more results and details.

We refer to Krieger, Nakanishi, and Schlag [2014, 2013] for classification results of solutions with energy at most slightly above the one of the ground state, and to Krieger, Schlag, and Tataru [2009] and Jendrej [2017] for contructions of solutions with prescribed blow up rates (*type II* blow up). We also refer to Martel and Merle [2002], Martel, Merle, and Raphaël [2014], and Merle and Raphael [2004, 2005] and references therein for previous results of blow up profile in the case of mass critical gKdV and NLS equations. Concerning blow up, see also the review Raphaël [2014] and the references therein.

5 Asymptotic multi-solitons

In this section, we discuss results of existence of *asymptotic multi-solitons* for non-integrable models, inspired by Theorem 1 and other explicit constructions of multi-solitons for integrable models, but limited to one direction of time. In particular, these results are valid in asymptotic situations where the distances between all the solitary waves are large enough.

5.1 Multi-solitons with weak interactions.

As a rough idea, weak interaction means that the trajectories of the solitary waves are not affected asymptotically.

Theorem 7 (Existence and uniqueness of gKdV multi-solitons, Martel [2005]). *Let $p=2, 3, 4$ or 5. Let $K\geq 2$, $0<c_K<\cdots<c_1$, and $\sigma_1,\ldots,\sigma_K\in\mathbb{R}$. Let $\mathbf{R}=\sum_{k=1}^K R_k$ where*

$$R_k(t,x)=Q_{c_k}(x-\sigma_k-c_k t).$$

There exists a unique H^1 solution of (3) such that $\lim_{t\to-\infty}\|u(t)-\mathbf{R}(t)\|_{H^1}=0$. Moreover, there exists $\kappa>0$ such that, for all $t\leq 0$, $\|u(t)-\mathbf{R}(t)\|_{H^1}\lesssim e^{-\kappa|t|}$.

Such result shows that the multi-soliton behavior is universal, at least in one direction of time. Observe that it says nothing about the behavior the solution as $t\to+\infty$. The uniqueness statement in the energy space is relevant even in the integrable case since the inverse scattering theory requires *a priori* more space decay. The stability of the multi-soliton structure was studied in the sub-critical case in Martel, Merle, and Tsai [2002].

We observe that a similar existence result also holds for the gKdV super-critical equation ($p > 5$), with a specific classification related to the exponential instability, see Combet [2011] and Côte, Martel, and Merle [2011].

For the NLS equation, we recall the following existence result.

Theorem 8 (Existence of NLS multi-solitary waves, Côte, Martel, and Merle [2011], Martel and Merle [2006], and Merle [1990])**.** *Let $d \geq 1$. Let $p > 1$ for $d = 1, 2$ and $1 < p < \frac{d+2}{d-2}$ for $d \geq 3$. Let $K \geq 2$ and for any $k \in \{1, \ldots, K\}$, let $c_k > 0$, $\beta_k \in \mathbb{R}^d$, $\sigma_k \in \mathbb{R}^d$ and $\gamma_k \in \mathbb{R}$. Assume that, for any $k \neq k'$, $\beta_k \neq \beta_{k'}$. Let $\mathbf{R} = \sum_{k=1}^{K} R_k$ where*

$$R_k(t,x) = Q_{c_k}(x-\sigma_k-\beta_k t)e^{i\Gamma_k(t,x)} \quad \text{and} \quad \Gamma_k(t,x) = \frac{1}{2}(\beta_k \cdot x) - \frac{1}{4}|\beta_k|^2 t + c_k t + \gamma_k.$$

Then, there exist $T_0 \in \mathbb{R}$, $\kappa > 0$ and an H^1 solution u of (4) such that, for all $t \leq T_0$, $\|u(t) - \mathbf{R}(t)\|_{H^1} \lesssim e^{-\kappa|t|}$.

Uniqueness (for critical and sub-critical nonlinearities) or classification (for super-critical nonlinearity) is an open problem. See Combet [2014] for multi-existence in the 1D super-critical case.

Note that the construction of multi-solitons and the study of the stability of the sums of multi-soliton has been extended to several other models, see *e.g.* Côte and Martel [2016] and Côte and Muñoz [2014] for the case of the nonlinear Klein-Gordon equation, and Ming, Rousset, and Tzvetkov [2015] for the water wave model.

For the 5D energy critical wave equation, the following existence result is proved in Martel and Merle [2016].

Theorem 9 (Existence of NLW multi-solitary waves, Martel and Merle [ibid.])**.** *Let $d = 5$. Let $K \geq 2$, and for any $k \in \{1, \ldots, K\}$, let $\lambda_k > 0$, $\boldsymbol{\sigma}_k \in \mathbb{R}^5$, $\iota_k \in \{-1, +1\}$ and $\boldsymbol{\ell}_k \in \mathbb{R}^5$ with $|\boldsymbol{\ell}_k| < 1$. Assume that, for any $k' \neq k$, $\boldsymbol{\ell}_k \neq \boldsymbol{\ell}_{k'}$. Let $\mathbf{R} = \sum_{k=1}^{K} R_k$ where*

$$R_k(t,x) = \frac{\iota_k}{\lambda_k^{\frac{3}{2}}} W_{\boldsymbol{\ell}_k}\left(\frac{x - \boldsymbol{\ell}_k t - \boldsymbol{\sigma}_k}{\lambda_k}\right).$$

Assume that one of the following assumptions holds

(1) Two-solitons: $K = 2$

(2) Collinear speeds: *For all $k \in \{1, \ldots, K\}$, $\boldsymbol{\ell}_k = \ell_k \mathbf{e}_1$ where $\ell_k \in (-1, 1)$.*

Then, there exist $T_0 \in \mathbb{R}$ and a solution u of (6) on $(-\infty, T_0]$ in the energy space such that

$$\lim_{t \to -\infty} \|\nabla_{x,t}(u(t) - \mathbf{R}(t))\|_{L^2} = 0.$$

5.2 Multi-solitons with strong interactions. We state a typical result where the strong interactions between the traveling waves indeed affect their trajectories.

Theorem 10 (Two-solitary waves with logarithmic distance, Nguyen [2016]). *Let* $d \geq 1$. *Let*

$$1 < p < \frac{d+2}{d-2} \ (p > 1 \text{ for } d = 1, 2) \quad \textit{and} \quad p \neq 1 + \frac{4}{d}.$$

There exists a solution u *of* (4) *such that* $|z_1(t) - z_2(t)| \sim 2 \log t$ *as* $t \to -\infty$ *and*

$$\lim_{t \to -\infty} \left\| u(t) - e^{-i\gamma(t)} \sum_{k=1,2} Q(\cdot - z_k(t)) \right\|_{H^1} = 0.$$

As discussed in Section 2.3, such solutions were already known in the integrable case by the inverse scattering theory. The above result means that this behavior is universal for general NLS equations, under the same restriction that the traveling waves have equal scaling. The mass critical case $p = 1 + \frac{4}{d}$ is excluded since it displays a special behavior related to blow up and where the above behavior is visible only in rescaled variables, as previously described in Martel and Raphael [2015]. For the gKdV equation, a result similar to Theorem 10 is given in Nguyen [2017].

We mention a few other previous results of strong interactions: for the Hartree equation Krieger, Martel, and Raphaël [2009], for the energy critical wave equation Jendrej [2016b,a] and Jendrej and Lawrie [2017], for the mass critical gKdV equation Combet and Martel [2017b,a], and for the half-wave equation Gérard, Lenzmann, Pocovnicu, and Raphaël [2018].

5.3 Soliton interaction with the background. Several papers deal with the question of the interaction of a soliton with a changing background or an impurity. See Holmer and Zworski [2007, 2008], Holmer, Marzuola, and Zworski [2007a], and Holmer, Marzuola, and Zworski [2007b] for the interaction of a soliton of NLS with a Dirac mass or a slowly varying potential, and Muñoz [2012b,a] for the interaction of a gKdV soliton with a slowing variable bottom.

6 Decomposition into solitons for the energy critical wave equation

Here, we recall the few existing results of decomposition in solitons in non-integrable cases. First, a complete result of decomposition into solitons for equation (6) was proved in Duyckaerts, Kenig, and Merle [2013] for the radial 3D case.

Theorem 11 (Soliton resolution for the 3D radial critical wave equation, Duyckaerts, Kenig, and Merle [ibid.]). *Let $d = 3$. Let u be a global radial solution of* (6). *Then, there exist a solution v_{L} of the linear wave equation, $K \in \mathbb{N}$, $\iota_k \in \{-1, 1\}$, $\lambda_k > 0$, such that*

$$\lim_{t\to+\infty} \left\| (u(t), \partial_t u(t)) - \left(v_{\mathrm{L}}(t) + \sum_{k=1}^{K} \frac{\iota_k}{\lambda_k^{\frac{1}{2}}(t)} W\left(\frac{\cdot}{\lambda_k(t)}\right), \partial_t v_{\mathrm{L}}(t) \right) \right\|_{\dot H^1 \times L^2} = 0,$$

and $\lambda_1(t) \ll \lambda_2(t) \ll \cdots \ll \lambda_K(t) \ll t$, as $t \to +\infty$.

Note that the above result is in some sense more complete than for gKdV (Section 2.1), since the residue is proved to scatter. A similar result holds for blow up solutions, provided they exhibit type II blow up. The soliton resolution conjecture was later proved in the non-radial case for a subsequence of time for the $3, 4$ and 5D energy critical wave equation in Duyckaerts, Kenig, and Merle [2016] and Duyckaerts, Jia, Kenig, and Merle [2017]. Note that a fundamental idea in the approach of Duyckaerts, Kenig, and Merle [2013] is the introduction of the *method of channels of energy* for the linear wave equation (see Theorem 16 for a typical result in 5D).

See similar results for the wave map problem in Côte [2015] and Côte, Kenig, Lawrie, and Schlag [2015a,b].

7 Collision problem

Concerning the collision problem, we recall the discussion in Craig, Guyenne, Hammack, Henderson, and Sulem [2006] on inelastic collisions. To study the collision problem, it is natural to study the behavior as $t \to +\infty$ of the solutions constructed in Theorems 7, 8, 9. See Craig, Guyenne, Hammack, Henderson, and Sulem [ibid.], page 057106-4 for suggesting this approach which seems more canonical than to study initial data with the sum of two solitons initially distant.

7.1 Collision for the quartic gKdV equation I. We consider the quartic gKdV equation

$$\partial_t u + \partial_x(\partial_x^2 u + u^4) = 0 \quad (t, x) \in \mathbb{R} \times \mathbb{R}. \tag{8}$$

The article Martel and Merle [2011a] (see also Muñoz [2010] for generalization to any gKdV equation) gives the first rigorous results concerning collision of solitons for a non-integrable equation, and in particular the first proof of non-existence of pure two-soliton solutions, in the case where one soliton is much smaller than the other one.

Theorem 12 (Collision of solitons with very different size, Martel and Merle [2011a]). *Assume* $0 < c \ll 1$. *Let* u *be the solution of* (8) *such that*

$$\lim_{t\to-\infty} \left\| u(t) - Q(\cdot - t) - Q_c(\cdot - ct) \right\|_{H^1} \to 0.$$

(i) Global stability of the 2-soliton behavior. *There exist* $c_1^+ \underset{c\to 0}{\sim} 1$, $c_2^+ \underset{c\to 0}{\sim} c$, ρ_1, ρ_2 *such that the function* w^+ *defined by*

$$w^+(t,x) = u(t,x) - Q_{c_1^+}(x - \rho_1(t)) - Q_{c_2^+}(x - \rho_2(t))$$

satisfies

$$\lim_{t\to+\infty} \|w^+(t)\|_{H^1(x\geq \frac{c}{10}t)} = 0 \quad \textit{and} \quad \sup_{t\in\mathbb{R}} \|w^+(t)\|_{H^1} \lesssim c^{\frac{1}{3}}.$$

(ii) Inelasticity of the collision. *Moreover, for* $t \gg 1$,

$$c_1^+ - 1 \gtrsim c^{\frac{17}{6}}, \quad 1 - \frac{c_2^+}{c} \gtrsim c^{\frac{8}{3}}, \quad c^{\frac{17}{12}} \lesssim \|w_x^+(t)\|_{L^2} + c^{\frac{1}{2}} \|w^+(t)\|_{L^2} \lesssim c^{\frac{11}{12}}.$$

The first part of the theorem means that the two solitons are preserved through the collision, even the smallest one. Indeed, for c small, $\sup_t \|w^+(t)\|_{H^1} \lesssim c^{\frac{1}{3}} \ll \|Q_c\|_{H^1} \sim c^{\frac{1}{12}}$.

The second part of the theorem says that the sizes of the final solitons as $t \to +\infty$ are slightly changed with respect to their original sizes as $t \to -\infty$, and that the residue does not converge to zero. In particular, the solution is not a pure 2-soliton as $t \to +\infty$ in this regime. Thus, the collision is *not elastic.*

7.2 Collision for the quartic gKdV equation II. A first intuition on the general problem of two solitons with *almost same sizes* comes from the explicit multi-solitons of the integrable case. From LeVeque [1987], we have a sharp description of the behavior of the multi-soliton of (1) satisfying

$$\lim_{t\to\pm\infty} \left\| u(t) - Q_{c_1}(\cdot - c_1 t - \sigma_1^{\pm}) - Q_{c_2}(\cdot - c_2 t - \sigma_2^{\pm}) \right\|_{H^1} = 0,$$

in the special asymptotics where $0 < \mu_0 = \frac{c_2 - c_1}{c_1 + c_2} \ll 1$, *i.e.* for nearly equal solitons. Indeed, the following global in time estimate is proved for some explicit functions $c_k(t)$, $\sigma_k(t)$:

$$\sup_{t,x\in\mathbb{R}} \left| u(t,x) - Q_{c_1(t)}(x - \sigma_1(t)) - Q_{c_2(t)}(x - \sigma_2(t)) \right| \lesssim \mu_0^2.$$

Moreover, it is proved that $\min_{t\in\mathbb{R}}(\sigma_1(t)-\sigma_2(t)) \sim 2|\ln\mu_0|$. This means that the minimum separation between the two solitons is large when $\mu_0 \ll 1$. In particular, the two solitons never cross and the solution has two maximum points for all time. The interaction is repulsive, the solitons exchange their sizes and speeds at large distance and consequently avoid the collision.

We now recall results from Mizumachi [2003] for the quartic gKdV equation. Let u be a solution of (8) for which the initial data is close to the sum $Q(x)+Q_c(x+Y_0)$, where $Y_0>0$ is large and $0\le c-1\le \exp(-\frac{1}{2}Y_0)$, so that the quicker soliton can be initially *on the left* of the other soliton. It follows from Mizumachi [ibid.] that the interaction of the two solitons is repulsive: the two solitons remain separated for all positive time and eventually $u(t)$ behaves as

$$u(t,x)=Q_{c_1^+}(x-c_1^+t-\sigma_1^+)+Q_{c_2^+}(x-c_2^+t-\sigma_2^+)+w(t,x),$$

for large time, for some $c_1^+>c_2^+$ close to 1 and w small in some sense. The situation for almost equal solitons of the quartic (gKdV) is thus at the main order similar to the one described in the integrable case in LeVeque [1987]. The analysis part in Mizumachi [2003] relies on techniques from Hayashi and Naumkin [1998, 2001] and on the use of spaces introduced in this context in Pego and Weinstein [1992].

Before presenting the main result from Martel and Merle [2011b], for simplicity, we change variables. For $c_2-c_1>0$ small, and any σ_1,σ_2, let $u(t)$ be the unique solution of (8) such that

$$\lim_{t\to-\infty}\|u(t)-Q_{c_1}(\cdot-c_1t-\sigma_1)-Q_{c_2}(\cdot-c_2t-\sigma_2)\|_{H^1}=0.$$

Let

$$c_0=\frac{c_1+c_2}{2},\quad \mu_0=\frac{c_2-c_1}{c_1+c_2},\quad y_1=\sigma_1\sqrt{c_0},\quad y_2=\sigma_2\sqrt{c_0}.$$

Then the function $U(t,x)=c_0^{-1/3}u(c_0^{-3/2}t,c_0^{-1/2}(x+t))$ solves

$$\partial_tU+\partial_x(\partial_x^2U-U+U^4)=0,\quad (t,x)\in\mathbb{R}\times\mathbb{R}, \tag{9}$$

and is the unique solution of (9) satisfying

$$\lim_{t\to-\infty}\left\|U(t)-Q_{1-\mu_0}(\cdot+\mu_0t-y_1)-Q_{1+\mu_0}(\cdot-\mu_0t-y_2)\right\|_{H^1}=0.$$

The next result concerns the asymptotics $\mu_0>0$ small.

Theorem 13 (Inelastic interaction of two nearly equal solitons, Martel and Merle [ibid.]). *Assume*
$0<\mu_0\ll 1$. *Let U be the unique solution of* (9) *such that*

$$\lim_{t\to-\infty}\left\|U(t)-Q_{1-\mu_0}(\cdot+\mu_0t+Y_0+\ln 2)-Q_{1+\mu_0}(\cdot-\mu_0t-Y_0-\ln 2)\right\|_{H^1}=0,$$

where $Y_0 = \frac{1}{2}|\ln(\mu_0^2/\alpha)|$ *and* $\alpha = 12(10)^{2/3}(\int Q^2)^{-1}$. *Then the following holds.*

(i) Global stability of the 2-soliton behavior. *There exist* μ_1, μ_2, y_1, y_2 *such that*

$$w(t,x) = U(t) - Q_{1+\mu_1(t)}(x - y_1(t)) - Q_{1+\mu_2(t)}(x - y_2(t))$$

satisfies $|\min_{t\in\mathbb{R}}(y_1(t) - y_2(t)) - 2Y_0| \lesssim \mu_0$ *and*

$$\lim_{t\to+\infty} \|w(t)\|_{H^1(x>-\frac{9}{10}t)} = 0, \quad \sup_{t\in\mathbb{R}} |w(t)\|_{H^1} \lesssim \mu_0^{\frac{3}{2}}.$$

(ii) Inelasticity of the interaction.

$$\lim_{+\infty} \mu_1 - \mu_0 \gtrsim \mu_0^5, \quad \mu_0 - \lim_{+\infty} \mu_2 \gtrsim \mu_0^5, \quad \liminf_{t\to+\infty} \|w(t)\|_{H^1} \gtrsim \mu_0^3.$$

It follows that *no pure* 2*-soliton exists* also in this regime. The proofs of Theorems 12 and 13 are based on the construction of a refined approximate solution of the two-soliton problem for all t and x.

7.3 Collision for the quartic gKdV equation III.

Still concerning the collision of two solitons for the quartic gKdV equation, we recall from Martel and Merle [2015] the following negative result.

Theorem 14 (Inelasticity of collision for gKdV, Martel and Merle [ibid.])**.** *Let* $K \geq 2$, $0 < c_K < \cdots < c_1 = 1$ *and* $\sigma_1, \ldots, \sigma_K \in \mathbb{R}$. *Let* u *be the solution of* (8) *satisfying*

$$\lim_{t\to-\infty} \left\| u(t) - \sum_{k=1}^{K} Q_{c_k}(\cdot - c_k t - \sigma_k) \right\|_{H^1} = 0.$$

Assume that $\sum_{k=2}^{K} (1 - c_k)^2 < \frac{1}{16}$. *Then,* $u(t)$ *is not an asymptotic multi-soliton as* $t \to +\infty$. *In particular, there exists no pure multi-soliton of* (1) *with the speeds* $c_1, c_2, \ldots, c_K$.

In the case of two solitons, the condition on the speeds reduces to $\frac{3}{4}c_1 < c_2 < c_1$. In contrast with Theorems 12 and 13, the result in Theorem 14 is *not perturbative* and the explicit condition on the speeds seems technical. The strategy of the proof of Theorem 14 is to study the asymptotic behavior of $u(t,x)$ for any t and for any large x (*i.e.* far from the collision region, which seems impossible to describe in the general case) and to find a contradiction with the fact that u is an asymptotic two-soliton in the two directions of time. Being a proof by contradiction, it does not give further information on the collision.

7.4 Collision for the perturbed integrable NLS equation. Let $\beta > 0$ and $0 < c \ll 1$. Under the following assumptions for the perturbation $|f(u)| \lesssim_0 u^2$, $f(u) \lesssim |u|^q$ ($q < 2$), it is proved in Perelman [2011], that there exists a solution u of

$$i\partial_t u + \partial_x^2 u + |u|^2 u + f(|u|^2)u = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{R},$$

satisfying $\lim_{t\to-\infty} \left\| u(t) - e^{it}Q - e^{i\Gamma_\beta(t,\cdot)}Q_c(\cdot - \beta t) \right\|_{H^1} \to 0$, and for which the small soliton splits in two after the collision, in the following sense

$$u(t,x) \sim e^{i\Gamma(t,x)}Q(x - \sigma(t)) + \psi^+(t,x) + \psi^-(t,x),$$

where $\psi^\pm$ are solutions of (2) corresponding to the *transmitted part* and the *reflected part* of the small soliton. The above estimate holds on large time intervals after the collision depending on $1/c$. The splitting means some strong form of inelasticity.

7.5 Collision for the 5D energy critical wave equation. In view of the results described so far, it was natural to search a situation where a non-perturbative approach would allow to treat *all two-soliton collisions*. In the case of the 5D energy critical wave equation, we now recall from Martel and Merle [2017] a result showing the inelastic nature of the collision of any two solitons, except the special case of same scaling and opposite signs.

Theorem 15 (Inelasticity of soliton collisions for NLW, Martel and Merle [ibid.]). *Let $d = 5$. For $k \in \{1,2\}$, let $\lambda_k^\infty > 0$, $\mathbf{y}_k^\infty \in \mathbb{R}^5$, $\epsilon_k \in \{\pm 1\}$, $\boldsymbol{\ell}_k \in \mathbb{R}^5$ with $|\boldsymbol{\ell}_k| < 1$, and*

$$W_k^\infty(t,x) = \frac{\epsilon_k}{(\lambda_k^\infty)^{\frac{3}{2}}} W_{\boldsymbol{\ell}_k}\left(\frac{x - \boldsymbol{\ell}_k t - \mathbf{y}_k^\infty}{\lambda_k^\infty}\right).$$

Assume that $\boldsymbol{\ell}_1 \neq \boldsymbol{\ell}_2$ and $\epsilon_1\lambda_1^\infty + \epsilon_2\lambda_2^\infty \neq 0$. Then, there exists a solution u of (6) in the energy space such that

(i) Two-soliton as $t \to -\infty$

$$\lim_{t\to-\infty} \|\nabla_{t,x}u(t) - \nabla_{t,x}\left(W_1^\infty(t) + W_2^\infty(t)\right)\|_{L^2} = 0.$$

(ii) Dispersion as $t \to +\infty$. *For all $A > 0$ large enough,*

$$\liminf_{t\to+\infty} \|\nabla u(t)\|_{L^2(|x|>t+A)} \gtrsim A^{-\frac{5}{2}}. \tag{10}$$

Note first that the solution constructed in Theorem 15 is a two-soliton asymptotically as $t \to -\infty$ and that it does not necessarily exist for all $t \in \mathbb{R}$. However, by finite speed of propagation and small data Cauchy theory, it is straightforward to justify that it can be extended uniquely as a solution of (6) for all $t \in \mathbb{R}$ in the region $|x| > |t| + A$, provided that A is large enough. Thus, the limit in (10) makes sense. Since the estimate (10) gives an explicit lower bound on the loss of energy as dispersion as $t \to +\infty$, the solution u is not a two-soliton asymptotically as $t \to +\infty$ and *the collision is inelastic*. Note that the two-soliton could have any global behavior, like dislocation of the solitons and dispersion, blow-up or a different multi-soliton plus radiation, but the property obtained is universal.

The only case left open by Theorem 15 corresponds to the *dipole* case. It is the first result proving inelasticity rigorously without restriction on the relative sizes or speeds of the solitons except an exceptional case.

The strategy of the proof is to construct a refined approximate solution of the two-soliton problem for large negative times that displays an explicit dispersive radial part at the leading order and then to propagate the dispersion for any positive time at the exterior of large cones by finite speed of propagation and the method of *channels of energy* from Duyckaerts, Kenig, and Merle [2013] and Kenig, Lawrie, and Schlag [2014]. To finish, we recall such a typical result of channel of energy for the radial linear wave equation in 5D from Kenig, Lawrie, and Schlag [2014], Proposition 4.1 (see also Duyckaerts, Kenig, and Merle [2012, 2013] and Kenig, Lawrie, B. Liu, and Schlag [2015]).

Theorem 16 (Exterior energy estimates for the 5D linear wave equation, Kenig, Lawrie, and Schlag [2014])**.** *Any radial energy solution U_{L} of the 5D linear wave equation*

$$\begin{cases} \partial_t^2 U_{\mathrm{L}} - \Delta U_{\mathrm{L}} = 0, & (t,x) \in \mathbb{R} \times \mathbb{R}^5, \\ U_{\mathrm{L}|t=0} = U_0 \in \dot{H}^1, & \partial_t U_{\mathrm{L}|t=0} = U_1 \in L^2, \end{cases}$$

satisfies, for any $R > 0$,

$$\sum_{\pm} \left\{ \limsup_{t \to \pm\infty} \int_{|x|>|t|+R} \left(|\partial_t U_{\mathrm{L}}(t,x)|^2 + |\nabla U_{\mathrm{L}}(t,x)|^2 \right) dx \right\} \gtrsim$$

$$\gtrsim \|\pi_R^{\perp}(U_0, U_1)\|^2_{(\dot{H}^1 \times L^2)(|x|>R)}$$

where $\pi_R^{\perp}(U_0,U_1)$ denotes the orthogonal projection of (U_0,U_1) onto the complement of the plane $\operatorname{span}\{(|x|^{-3},0),(0,|x|^{-3})\}$ *in* $(\dot{H}^1 \times L^2)(|x| > R)$.

References

M. J. Ablowitz, M. D. Kruskal, and J. F. Ladik (1979). "Solitary wave collisions". *SIAM J. Appl. Math.* 36.3, pp. 428–437. MR: 531606 (cit. on p. 2462).

Mark J. Ablowitz, David J. Kaup, Alan C. Newell, and Harvey Segur (1974). "The inverse scattering transform-Fourier analysis for nonlinear problems". *Studies in Appl. Math.* 53.4, pp. 249–315. MR: 0450815 (cit. on pp. 2457, 2460).

Miguel A. Alejo and Claudio Muñoz (2013). "Nonlinear stability of MKdV breathers". *Comm. Math. Phys.* 324.1, pp. 233–262. MR: 3116324 (cit. on p. 2460).

Marius Beceanu (2008). "A centre-stable manifold for the focussing cubic NLS in $\mathbb{R}^{1+3}$". *Comm. Math. Phys.* 280.1, pp. 145–205. MR: 2391193 (cit. on p. 2466).

– (2012). "A critical center-stable manifold for Schrödinger's equation in three dimensions". *Comm. Pure Appl. Math.* 65.4, pp. 431–507. MR: 2877342 (cit. on p. 2466).

T. B. Benjamin (1972). "The stability of solitary waves". *Proc. Roy. Soc. (London) Ser. A* 328, pp. 153–183. MR: 0338584 (cit. on p. 2463).

H. Berestycki and P.-L. Lions (1983). "Nonlinear scalar field equations. I. Existence of a ground state". *Arch. Rational Mech. Anal.* 82.4, pp. 313–345. MR: 695535 (cit. on p. 2464).

Jerry L. Bona (1975). "On the stability theory of solitary waves". *Proc. Roy. Soc. London Ser. A* 344.1638, pp. 363–374. MR: 0386438 (cit. on p. 2463).

Jerry L. Bona, W. G. Pritchard, and L Ridgway Scott (1980). "Solitary-wave interaction". *The Physics of Fluids* 23.3, pp. 438–441 (cit. on p. 2462).

Jerry L. Bona, P. E. Souganidis, and Walter Strauss (1987). "Stability and instability of solitary waves of Korteweg-de Vries type". *Proc. Roy. Soc. London Ser. A* 411.1841, pp. 395–412. MR: 897729 (cit. on p. 2463).

Michael Borghese, Robert Jenkins, and Kenneth D. T. -R. McLaughlin (Apr. 2016). "Long Time Asymptotic Behavior of the Focusing Nonlinear Schrodinger Equation". arXiv: 1604.07436 (cit. on p. 2461).

V. S. Buslaev and Galina S. Perelman (1992). "Scattering for the nonlinear Schrödinger equation: states that are close to a soliton". *Algebra i Analiz* 4.6, pp. 63–102. MR: 1199635 (cit. on p. 2466).

– (1995). "On the stability of solitary waves for nonlinear Schrödinger equations". In: *Nonlinear evolution equations*. Vol. 164. Amer. Math. Soc. Transl. Ser. 2. Amer. Math. Soc., Providence, RI, pp. 75–98. MR: 1334139 (cit. on p. 2466).

Vladimir S. Buslaev and Catherine Sulem (2003). "On asymptotic stability of solitary waves for nonlinear Schrödinger equations". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 20.3, pp. 419–475. MR: 1972870 (cit. on p. 2466).

T. Cazenave and P.-L. Lions (1982). "Orbital stability of standing waves for some nonlinear Schrödinger equations". *Comm. Math. Phys.* 85.4, pp. 549–561. MR: 677997 (cit. on pp. 2463, 2464).

Thierry Cazenave (2003). *Semilinear Schrödinger equations*. Vol. 10. Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, pp. xiv+323. MR: 2002047 (cit. on pp. 2463, 2464).

Amy Cohen (1979). "Existence and regularity for solutions of the Korteweg-de Vries equation". *Arch. Rational Mech. Anal.* 71.2, pp. 143–175. MR: 525222 (cit. on p. 2460).

Vianney Combet (2011). "Multi-soliton solutions for the supercritical gKdV equations". *Comm. Partial Differential Equations* 36.3, pp. 380–419. MR: 2763331 (cit. on p. 2469).

– (2014). "Multi-existence of multi-solitons for the supercritical nonlinear Schrödinger equation in one dimension". *Discrete Contin. Dyn. Syst.* 34.5, pp. 1961–1993. MR: 3124722 (cit. on p. 2469).

Vianney Combet and Yvan Martel (June 2017a). "Construction of multi-bubble solutions for the critical gKdV equation". arXiv: 1706.09870 (cit. on p. 2470).

– (2017b). "Sharp asymptotics for the minimal mass blow up solution of the critical gKdV equation". *Bull. Sci. Math.* 141.2, pp. 20–103. MR: 3614115 (cit. on p. 2470).

Ovidiu Costin, Min Huang, and Wilhelm Schlag (2012). "On the spectral properties of $L_{\pm}$ in three dimensions". *Nonlinearity* 25.1, pp. 125–164. MR: 2864380 (cit. on p. 2466).

R. Côte (2015). "On the soliton resolution for equivariant wave maps to the sphere". *Comm. Pure Appl. Math.* 68.11, pp. 1946–2004. MR: 3403756 (cit. on p. 2471).

R. Côte, Carlos Kenig, A. Lawrie, and W. Schlag (2015a). "Characterization of large energy solutions of the equivariant wave map problem: I". *Amer. J. Math.* 137.1, pp. 139–207. MR: 3318089 (cit. on p. 2471).

– (2015b). "Characterization of large energy solutions of the equivariant wave map problem: II". *Amer. J. Math.* 137.1, pp. 209–250. MR: 3318090 (cit. on p. 2471).

Raphaël Côte and Yvan Martel (Dec. 2016). "Multi-travelling waves for the nonlinear klein-gordon equation". To appear in Trans. Amer. Math. Soc. arXiv: 1612.02625 (cit. on p. 2469).

Raphaël Côte, Yvan Martel, and Frank Merle (2011). "Construction of multi-soliton solutions for the L^2-supercritical gKdV and NLS equations". *Rev. Mat. Iberoam.* 27.1, pp. 273–302. MR: 2815738 (cit. on p. 2469).

Raphaël Côte and Claudio Muñoz (2014). "Multi-solitons for nonlinear Klein–Gordon equations". *Forum Math. Sigma* 2, e15, 38. MR: 3264254 (cit. on p. 2469).

W. Craig, P. Guyenne, J. Hammack, D. Henderson, and C. Sulem (2006). "Solitary water wave interactions". *Phys. Fluids* 18.5, pp. 057106, 25. MR: 2259317 (cit. on pp. 2458, 2462, 2471).

Scipio Cuccagna (2008). "On asymptotic stability in 3D of kinks for the ϕ^4 model". *Trans. Amer. Math. Soc.* 360.5, pp. 2581–2614. MR: 2373326 (cit. on p. 2467).

– (2014). "On asymptotic stability of moving ground states of the nonlinear Schrödinger equation". *Trans. Amer. Math. Soc.* 366.6, pp. 2827–2888. MR: 3180733 (cit. on p. 2466).

Scipio Cuccagna and Dmitry E. Pelinovsky (2014). "The asymptotic stability of solitons in the cubic NLS equation on the line". *Appl. Anal.* 93.4, pp. 791–822. MR: 3180019 (cit. on p. 2466).

Sara Cuenda, Niurka R. Quintero, and Angel Sánchez (2011). "Sine-Gordon wobbles through Bäcklund transformations". *Discrete Contin. Dyn. Syst. Ser. S* 4.5, pp. 1047–1056. MR: 2754104 (cit. on p. 2461).

Thierry Dauxois and Michel Peyrard (2010). *Physics of solitons*. French. Reprint of the 2006 edition. Cambridge University Press, Cambridge, pp. xii+422. MR: 3642602 (cit. on pp. 2457, 2459–2462, 2464).

Percy Deift and E. Trubowitz (1979). "Inverse scattering on the line". *Comm. Pure Appl. Math.* 32.2, pp. 121–251. MR: 512420 (cit. on p. 2460).

Percy Deift, S. Venakides, and X. Zhou (1994). "The collisionless shock region for the long-time behavior of solutions of the KdV equation". *Comm. Pure Appl. Math.* 47.2, pp. 199–206. MR: 1263128 (cit. on p. 2460).

Percy Deift and X. Zhou (1993). "Long-time asymptotics for integrable nonlinear wave equations". In: *Important developments in soliton theory*. Springer Ser. Nonlinear Dynam. Springer, Berlin, pp. 181–204. MR: 1280475 (cit. on p. 2460).

Percy Deift and Xin Zhou (2002). "Perturbation theory for infinite-dimensional integrable systems on the line. A case study". *Acta Math.* 188.2, pp. 163–262. MR: 1947893 (cit. on p. 2461).

Roger K. Dodd, J. Chris Eilbeck, John D. Gibbon, and Hedley C. Morris (1982). *Solitons and nonlinear wave equations*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, pp. x+630. MR: 696935 (cit. on pp. 2460, 2461).

Thomas Duyckaerts, Hao Jia, Carlos Kenig, and Frank Merle (2017). "Soliton resolution along a sequence of times for the focusing energy critical wave equation". *Geom. Funct. Anal.* 27.4, pp. 798–862. MR: 3678502 (cit. on p. 2471).

Thomas Duyckaerts, Carlos Kenig, and Frank Merle (2011). "Universality of blow-up profile for small radial type II blow-up solutions of the energy-critical wave equation". *J. Eur. Math. Soc. (JEMS)* 13.3, pp. 533–599. MR: 2781926 (cit. on p. 2467).

– (2012). "Universality of the blow-up profile for small type II blow-up solutions of the energy-critical wave equation: the nonradial case". *J. Eur. Math. Soc. (JEMS)* 14.5, pp. 1389–1454. MR: 2966655 (cit. on pp. 2467, 2476).

– (2013). "Classification of radial solutions of the focusing, energy-critical wave equation". *Camb. J. Math.* 1.1, pp. 75–144. MR: 3272053 (cit. on pp. 2470, 2471, 2476).

– (2016). "Solutions of the focusing nonradial critical wave equation with the compactness property". *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* 15, pp. 731–808. MR: 3495446 (cit. on p. 2471).

Thomas Duyckaerts and Frank Merle (2008). "Dynamics of threshold solutions for energy-critical wave equation". *Int. Math. Res. Pap. IMRP*, Art ID rpn002, 67. MR: 2470571 (cit. on pp. 2465, 2467).

W. Eckhaus and Peter C. Schuur (1983). "The emergence of solitons of the Korteweg-de Vries equation from arbitrary initial conditions". *Math. Methods Appl. Sci.* 5.1, pp. 97–116. MR: 690898 (cit. on pp. 2459, 2460).

Shin-Ichiro Ei and Takao Ohta (1994). "Equation of motion for interacting pulses". *Physical Review E* 50.6, pp. 4672–4678 (cit. on p. 2461).

Ludwig D. Faddeev and Leon A. Takhtajan (2007). *Hamiltonian methods in the theory of solitons*. English. Classics in Mathematics. Translated from the 1986 Russian original by Alexey G. Reyman. Springer, Berlin, pp. x+592. MR: 2348643 (cit. on p. 2460).

E. Fermi, J. Pasta, and S. Ulam (1955). "Studies of nonlinear problems, I". *Los Alamos Rep. LA1940*. Reproduced in Nonlinear Wave Motion, A. C. Newell (ed.) (cit. on pp. 2457, 2462).

C. S. Gardner, C. S. Greene, M. D. Kruskal, and R. M. Miura (1967). "Method for solving the Korteweg-de Vries equation". *Phys. Rev. Lett.* 19, pp. 1095–1087 (cit. on p. 2457).

Clifford S. Gardner, John M. Greene, Martin D. Kruskal, and Robert M. Miura (1974). "Korteweg-deVries equation and generalization. VI. Methods for exact solution". *Comm. Pure Appl. Math.* 27, pp. 97–133. MR: 0336122 (cit. on p. 2457).

Patrick Gérard, Enno Lenzmann, Oana Pocovnicu, and Pierre Raphaël (2018). "A Two-Soliton with Transient Turbulent Regime for the Cubic Half-Wave Equation on the Real Line". *Ann. PDE* 4.1, 4:7. MR: 3747579 (cit. on p. 2470).

Pierre Germain, Fabio Pusateri, and Frédéric Rousset (2016). "Asymptotic stability of solitons for mKdV". *Adv. Math.* 299, pp. 272–330. MR: 3519470 (cit. on p. 2466).

J. Ginibre and G. Velo (1979). "On a class of nonlinear Schrödinger equations. I. The Cauchy problem, general case". *J. Funct. Anal.* 32.1, pp. 1–32. MR: 533218 (cit. on p. 2463).

KA Gorshkov and LA Ostrovsky (1981). "Interactions of solitons in nonintegrable systems: direct perturbation method and applications". *Physica D: Nonlinear Phenomena* 3.1-2, pp. 428–438 (cit. on pp. 2460, 2461).

Manoussos Grillakis (1990). "Analysis of the linearization around a critical point of an infinite-dimensional Hamiltonian system". *Comm. Pure Appl. Math.* 43.3, pp. 299–333. MR: 1040143 (cit. on pp. 2464, 2465).

Manoussos Grillakis, Jalal Shatah, and Walter Strauss (1987). "Stability theory of solitary waves in the presence of symmetry. I". *J. Funct. Anal.* 74.1, pp. 160–197. MR: 901236 (cit. on pp. 2463, 2464).

Stephen Gustafson, Kenji Nakanishi, and Tai-Peng Tsai (2004). "Asymptotic stability and completeness in the energy space for nonlinear Schrödinger equations with small solitary waves". *Int. Math. Res. Not.* 66, pp. 3559–3584. MR: 2101699 (cit. on p. 2466).

Joseph Hammack, Diane Henderson, Philippe Guyenne, and Ming Yi (2004). "Solitary-wave collisions". In: *Proceedings of the 23rd ASME Offshore Mechanics and Artic*

Engineering (A symposium to honor Theodore Yao-Tsu Wu), Vancouver, Canada, June 2004. World Scientific, Singapore (cit. on p. 2462).

Nakao Hayashi and Pavel I. Naumkin (1998). "Large time asymptotics of solutions to the generalized Korteweg-de Vries equation". *J. Funct. Anal.* 159.1, pp. 110–136. MR: 1654178 (cit. on p. 2473).

– (2001). "On the modified Korteweg-de Vries equation". *Math. Phys. Anal. Geom.* 4.3, pp. 197–227. MR: 1871426 (cit. on p. 2473).

Daniel B. Henry, J. Fernando Perez, and Walter F. Wreszinski (1982). "Stability theory for solitary-wave solutions of scalar field equations". *Comm. Math. Phys.* 85.3, pp. 351–361. MR: 678151 (cit. on pp. 2464, 2467).

Ryogo Hirota (1971). "Exact solution of the Korteweg—de Vries equation for multiple collisions of solitons". *Physical Review Letters* 27.18, pp. 1192–1194 (cit. on p. 2459).

J. Holmer, J. Marzuola, and M. Zworski (2007a). "Soliton splitting by external delta potentials". *J. Nonlinear Sci.* 17.4, pp. 349–367. MR: 2335125 (cit. on p. 2470).

Justin Holmer, Jeremy Marzuola, and Maciej Zworski (2007b). "Fast soliton scattering by delta impurities". *Comm. Math. Phys.* 274.1, pp. 187–216. MR: 2318852 (cit. on p. 2470).

Justin Holmer and Maciej Zworski (2007). "Slow soliton interaction with delta impurities". *J. Mod. Dyn.* 1.4, pp. 689–718. MR: 2342704 (cit. on p. 2470).

– (2008). "Soliton interaction with slowly varying potentials". *Int. Math. Res. Not. IMRN* 10, Art. ID rnn026, 36. MR: 2429244 (cit. on p. 2470).

Jacek Jendrej (Feb. 2016a). "Construction of two-bubble solutions for energy-critical wave equations". To appear in *Amer. J. Math.* arXiv: 1602.06524 (cit. on p. 2470).

– (2016b). "Construction of two-bubble solutions for some energy-critical wave equations". In: *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2015–2016*. Ed. Éc. Polytech., Palaiseau, Exp. No. XXI, 10. MR: 3616322 (cit. on p. 2470).

– (2017). "Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5". *J. Funct. Anal.* 272.3, pp. 866–917. MR: 3579128 (cit. on p. 2468).

Jacek Jendrej and Andrew Lawrie (May 2017). "Two-bubble dynamics for threshold solutions to the wave maps equation". To appear in SIAM Nonlinear Analysis. arXiv: 1706.00089 (cit. on p. 2470).

Robert Jenkins, Jiaqi Liu, Peter Perry, and Catherine Sulem (Oct. 2017). "Soliton Resolution for the Derivative Nonlinear Schrödinger Equation". arXiv: 1710.03819 (cit. on p. 2461).

V. I. Karpman and V. V. Solovev (1981). "A perturbational approach to the two-soliton systems". *Phys. D* 3.3, pp. 487–502. MR: 631182 (cit. on pp. 2460, 2461).

Tosio Kato (1983). "On the Cauchy problem for the (generalized) Korteweg-de Vries equation". In: *Studies in applied mathematics*. Vol. 8. Adv. Math. Suppl. Stud. Academic Press, New York, pp. 93–128. MR: 759907 (cit. on p. 2462).

Carlos Kenig, Andrew Lawrie, Baoping Liu, and Wilhelm Schlag (2015). "Channels of energy for the linear radial wave equation". *Adv. Math.* 285, pp. 877–936. MR: 3406517 (cit. on p. 2476).

Carlos Kenig, Andrew Lawrie, and Wilhelm Schlag (2014). "Relaxation of wave maps exterior to a ball to harmonic maps for all data". *Geom. Funct. Anal.* 24.2, pp. 610–647. MR: 3192036 (cit. on p. 2476).

Carlos Kenig and Yvan Martel (2009). "Asymptotic stability of solitons for the Benjamin-Ono equation". *Rev. Mat. Iberoam.* 25.3, pp. 909–970. MR: 2590690 (cit. on p. 2466).

Carlos Kenig and Frank Merle (2008). "Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation". *Acta Math.* 201.2, pp. 147–212. MR: 2461508 (cit. on pp. 2465, 2467).

Carlos Kenig, Gustavo Ponce, and Luis Vega (1993). "Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle". *Comm. Pure Appl. Math.* 46.4, pp. 527–620. MR: 1211741 (cit. on p. 2462).

Y. S. Kivshar and B. Malomed (1989). "Dynamics of solitons in nearly integrable models". *Review of Modern Physics* 61, pp. 763–915 (cit. on pp. 2458, 2461).

E. A. Kopylova and A. I. Komech (2011a). "On asymptotic stability of kink for relativistic Ginzburg-Landau equations". *Arch. Ration. Mech. Anal.* 202.1, pp. 213–245. MR: 2835867 (cit. on p. 2467).

– (2011b). "On asymptotic stability of moving kink for relativistic Ginzburg-Landau equation". *Comm. Math. Phys.* 302.1, pp. 225–252. MR: 2770013 (cit. on p. 2467).

Michał Kowalczyk, Yvan Martel, and Claudio Muñoz (2017). "Kink dynamics in the ϕ^4 model: asymptotic stability for odd perturbations in the energy space". *J. Amer. Math. Soc.* 30.3, pp. 769–798. MR: 3630087 (cit. on p. 2467).

– (2016-2017). "On asymptotic stability of nonlinear waves". *Séminaire Laurent Schwartz – EDP et applications*. Exp. No. 18, p. 27 (cit. on p. 2467).

Joachim Krieger, Yvan Martel, and Pierre Raphaël (2009). "Two-soliton solutions to the three-dimensional gravitational Hartree equation". *Comm. Pure Appl. Math.* 62.11, pp. 1501–1550. MR: 2560043 (cit. on p. 2470).

Joachim Krieger, Kenji Nakanishi, and Wilhelm Schlag (2013). "Global dynamics away from the ground state for the energy-critical nonlinear wave equation". *Amer. J. Math.* 135.4, pp. 935–965. MR: 3086065 (cit. on p. 2468).

– (2014). "Threshold phenomenon for the quintic wave equation in three dimensions". *Comm. Math. Phys.* 327.1, pp. 309–332. MR: 3177940 (cit. on p. 2468).

Joachim Krieger, Wilhelm Schlag, and Daniel Tataru (2009). "Slow blow-up solutions for the $H^1(\mathbb{R}^3)$ critical focusing semilinear wave equation". *Duke Math. J.* 147.1, pp. 1–53. MR: 2494455 (cit. on p. 2468).

Martin D. Kruskal (1974). "The Korteweg-de Vries equation and related evolution equations", 61–83. Lectures in Appl. Math., Vol. 15. MR: 0352741 (cit. on p. 2460).

Sergei B. Kuksin (2000). *Analysis of Hamiltonian PDEs*. Vol. 19. Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, pp. xii+212. MR: 1857574 (cit. on p. 2461).

Man Kam Kwong (1989). "Uniqueness of positive solutions of $\Delta u - u + u^p = 0$ in $\mathbf{R}^n$". *Arch. Rational Mech. Anal.* 105.3, pp. 243–266. MR: 969899 (cit. on p. 2464).

George L. Lamb Jr. (1980). *Elements of soliton theory*. Pure and Applied Mathematics, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, pp. xiii+289. MR: 591458 (cit. on pp. 2457, 2460, 2461).

Peter D. Lax (1968). "Integrals of nonlinear equations of evolution and solitary waves". *Comm. Pure Appl. Math.* 21, pp. 467–490. MR: 0235310 (cit. on pp. 2457, 2459, 2460).

Randall J. LeVeque (1987). "On the interaction of nearly equal solitons in the KdV equation". *SIAM J. Appl. Math.* 47.2, pp. 254–262. MR: 881349 (cit. on pp. 2472, 2473).

Yi Li and D. H. Sattinger (1999). "Soliton collisions in the ion acoustic plasma equations". *J. Math. Fluid Mech.* 1.1, pp. 117–130. MR: 1699021 (cit. on p. 2462).

Nicholas Manton and Paul Sutcliffe (2004). *Topological solitons*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, pp. xii+493. MR: 2068924 (cit. on p. 2464).

Yvan Martel (2005). "Asymptotic N-soliton-like solutions of the subcritical and critical generalized Korteweg-de Vries equations". *Amer. J. Math.* 127.5, pp. 1103–1140. MR: 2170139 (cit. on pp. 2466, 2468).

Yvan Martel and Frank Merle (2000). "A Liouville theorem for the critical generalized Korteweg-de Vries equation". *J. Math. Pures Appl. (9)* 79.4, pp. 339–425. MR: 1753061 (cit. on p. 2463).

– (2001). "Asymptotic stability of solitons for subcritical generalized KdV equations". *Arch. Ration. Mech. Anal.* 157.3, pp. 219–254. MR: 1826966 (cit. on p. 2465).

– (2002). "Stability of blow-up profile and lower bounds for blow-up rate for the critical generalized KdV equation". *Ann. of Math. (2)* 155.1, pp. 235–280. MR: 1888800 (cit. on p. 2468).

– (2005). "Asymptotic stability of solitons of the subcritical gKdV equations revisited". *Nonlinearity* 18.1, pp. 55–80. MR: 2109467.

– (2006). "Multi solitary waves for nonlinear Schrödinger equations". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 23.6, pp. 849–864. MR: 2271697 (cit. on p. 2469).

– (2011a). "Description of two soliton collision for the quartic gKdV equation". *Ann. of Math. (2)* 174.2, pp. 757–857. MR: 2831108 (cit. on pp. 2471, 2472).

Yvan Martel and Frank Merle (2011b). "Inelastic interaction of nearly equal solitons for the quartic gKdV equation". *Invent. Math.* 183.3, pp. 563–648. MR: 2772088 (cit. on p. 2473).

– (2015). "On the nonexistence of pure multi-solitons for the quartic gKdV equation". *Int. Math. Res. Not. IMRN* 3, pp. 688–739. MR: 3340333 (cit. on p. 2474).

– (2016). "Construction of multi-solitons for the energy-critical wave equation in dimension 5". *Arch. Ration. Mech. Anal.* 222.3, pp. 1113–1160. MR: 3544324 (cit. on p. 2469).

– (Aug. 2017). "Inelasticity of soliton collisions for the 5D energy critical wave equation". arXiv: 1708.09712 (cit. on p. 2475).

Yvan Martel, Frank Merle, and Pierre Raphaël (2014). "Blow up for the critical generalized Korteweg–de Vries equation. I: Dynamics near the soliton". *Acta Math.* 212.1, pp. 59–140. MR: 3179608 (cit. on pp. 2462, 2468).

Yvan Martel, Frank Merle, and Tai-Peng Tsai (2002). "Stability and asymptotic stability in the energy space of the sum of N solitons for subcritical gKdV equations". *Comm. Math. Phys.* 231.2, pp. 347–373. MR: 1946336 (cit. on p. 2468).

Yvan Martel and Pierre Raphael (Dec. 2015). "Strongly interacting blow up bubbles for the mass critical NLS". To appear in *Ann. Sci. Éc. Norm. Supér.* arXiv: 1512.00900 (cit. on p. 2470).

Frank Merle (1990). "Construction of solutions with exactly k blow-up points for the Schrödinger equation with critical nonlinearity". *Comm. Math. Phys.* 129.2, pp. 223–240. MR: 1048692 (cit. on p. 2469).

Frank Merle and Pierre Raphael (2004). "On universality of blow-up profile for L^2 critical nonlinear Schrödinger equation". *Invent. Math.* 156.3, pp. 565–672. MR: 2061329 (cit. on p. 2468).

– (2005). "Profiles and quantization of the blow up mass for critical nonlinear Schrödinger equation". *Comm. Math. Phys.* 253.3, pp. 675–704. MR: 2116733 (cit. on p. 2468).

Mei Ming, Frederic Rousset, and Nikolay Tzvetkov (2015). "Multi-solitons and related solutions for the water-waves system". *SIAM J. Math. Anal.* 47.1, pp. 897–954. MR: 3315224 (cit. on p. 2469).

Robert M. Miura (1976). "The Korteweg-de Vries equation: a survey of results". *SIAM Rev.* 18.3, pp. 412–459. MR: 0404890 (cit. on pp. 2457, 2459, 2460, 2466).

Robert M. Miura, Clifford S. Gardner, and Martin D. Kruskal (1968). "Korteweg-de Vries equation and generalizations. II. Existence of conservation laws and constants of motion". *J. Mathematical Phys.* 9, pp. 1204–1209. MR: 0252826 (cit. on p. 2457).

Tetsu Mizumachi (2003). "Weak interaction between solitary waves of the generalized KdV equations". *SIAM J. Math. Anal.* 35.4, pp. 1042–1080. MR: 2049033 (cit. on p. 2473).

Claudio Muñoz (2010). "On the inelastic two-soliton collision for gKdV equations with general nonlinearity". *Int. Math. Res. Not. IMRN* 9, pp. 1624–1719. MR: 2643578 (cit. on p. 2471).

– (2012a). "Dynamics of soliton-like solutions for slowly varying, generalized KdV equations: refraction versus reflection". *SIAM J. Math. Anal.* 44.1, pp. 1–60. MR: 2888279 (cit. on p. 2470).

– (2012b). "Inelastic character of solitons of slowly varying gKdV equations". *Comm. Math. Phys.* 314.3, pp. 817–852. MR: 2964776 (cit. on p. 2470).

Kenji Nakanishi and Wilhelm Schlag (2011). *Invariant manifolds and dispersive Hamiltonian evolution equations*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, pp. vi+253. MR: 2847755 (cit. on p. 2466).

Tien Vinh Nguyen (Nov. 2016). "Existence of multi-solitary waves with logarithmic relative distances for the NLS equation". arXiv: 1611.08869 (cit. on p. 2470).

– (2017). "Strongly interacting multi-solitons with logarithmic relative distance for gKdV equation". To appear in Nonlinearity. arXiv: 1705.07319 (cit. on p. 2470).

V. Ju. Novokšenov (1980). "Asymptotic behavior as $t \to \infty$ of the solution of the Cauchy problem for a nonlinear Schrödinger equation". *Dokl. Akad. Nauk SSSR* 251.4, pp. 799–802. MR: 568535 (cit. on p. 2460).

E. Olmedilla (1987). "Multiple pole solutions of the nonlinear Schrödinger equation". *Phys. D* 25.1-3, pp. 330–346. MR: 887466 (cit. on pp. 2460, 2461).

Robert L. Pego and Michael I. Weinstein (1992). "Eigenvalues, and instabilities of solitary waves". *Philos. Trans. Roy. Soc. London Ser. A* 340.1656, pp. 47–94. MR: 1177566 (cit. on pp. 2463, 2473).

– (1994). "Asymptotic stability of solitary waves". *Comm. Math. Phys.* 164.2, pp. 305–349. MR: 1289328 (cit. on p. 2466).

Galina S. Perelman (2011). "Two soliton collision for nonlinear Schrödinger equations in dimension 1". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 28.3, pp. 357–384. MR: 2795711 (cit. on p. 2475).

Pierre Raphaël (2014). "On singularity formation in Hamiltonian evolution equations". In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. III*. Kyung Moon Sa, Seoul, pp. 849–872. MR: 3729055 (cit. on p. 2468).

I. Rodnianski, W. Schlag, and A. Soffer (Sept. 2003). "Asymptotic stability of N-soliton states of NLS". arXiv: math/0309114 (cit. on p. 2466).

Igor Rodnianski, Wilhelm Schlag, and Avraham Soffer (2005). "Dispersive analysis of charge transfer models". *Comm. Pure Appl. Math.* 58.2, pp. 149–216. MR: 2094850 (cit. on p. 2466).

W. Schlag (2007). "Dispersive estimates for Schrödinger operators: a survey". In: *Mathematical aspects of nonlinear dispersive equations*. Vol. 163. Ann. of Math. Stud. Princeton Univ. Press, Princeton, NJ, pp. 255–285. MR: 2333215 (cit. on p. 2466).

W. Schlag (2009). "Stable manifolds for an orbitally unstable nonlinear Schrödinger equation". *Ann. of Math. (2)* 169.1, pp. 139–227. MR: 2480603 (cit. on p. 2466).

Wilhelm Schlag (2006). "Spectral theory and nonlinear partial differential equations: a survey". *Discrete Contin. Dyn. Syst.* 15.3, pp. 703–723. MR: 2220744 (cit. on p. 2466).

Peter C. Schuur (1986). *Asymptotic analysis of soliton problems*. Vol. 1232. Lecture Notes in Mathematics. An inverse scattering approach. Springer-Verlag, pp. viii+180. MR: 874343 (cit. on pp. 2459, 2460).

Harvey Segur (1983). "Wobbling kinks in φ^4 and sine-Gordon theory". *J. Math. Phys.* 24.6, pp. 1439–1443. MR: 708660 (cit. on p. 2461).

L. Y. Shih (1980). "Soliton-like interaction governed by the generalized Korteweg-de Vries equation". *Wave Motion* 2.3, pp. 197–206. MR: 589165 (cit. on p. 2462).

A. Soffer and M. I. Weinstein (1990). "Multichannel nonlinear scattering for nonintegrable equations". *Comm. Math. Phys.* 133.1, pp. 119–146. MR: 1071238 (cit. on p. 2466).

– (1992). "Multichannel nonlinear scattering for nonintegrable equations. II. The case of anisotropic potentials and data". *J. Differential Equations* 98.2, pp. 376–390. MR: 1170476 (cit. on p. 2466).

Terence Tao (2006). *Nonlinear dispersive equations*. Vol. 106. CBMS Regional Conference Series in Mathematics. Local and global analysis. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, pp. xvi+373. MR: 2233925 (cit. on pp. 2463, 2464).

– (2007). "Scattering for the quartic generalised Korteweg-de Vries equation". *J. Differential Equations* 232.2, pp. 623–651. MR: 2286393 (cit. on p. 2466).

– (2009). "Why are solitons stable?" *Bull. Amer. Math. Soc. (N.S.)* 46.1, pp. 1–33. MR: 2457070 (cit. on p. 2466).

Miki Wadati (1973). "The modified Korteweg-de Vries equation". *J. Phys. Soc. Japan* 34, pp. 1289–1296. MR: 0371251 (cit. on p. 2460).

Miki Wadati and Kenji Ohkuma (1982). "Multiple-pole solutions of the modified Korteweg-de Vries equation". *J. Phys. Soc. Japan* 51.6, pp. 2029–2035. MR: 669118 (cit. on p. 2460).

Miki Wadati and M. Toda (1972). "The exact N–soliton solution of the Korteweg–de Vries equation". *J. Phys. Soc. Japan* 32, pp. 1403–1411 (cit. on p. 2466).

Michael I. Weinstein (1985). "Modulational stability of ground states of nonlinear Schrödinger equations". *SIAM J. Math. Anal.* 16.3, pp. 472–491. MR: 783974 (cit. on pp. 2463, 2464).

– (1986). "Lyapunov stability of ground states of nonlinear dispersive evolution equations". *Comm. Pure Appl. Math.* 39.1, pp. 51–67. MR: 820338 (cit. on p. 2463).

G. B. Whitham (1974). *Linear and nonlinear waves*. Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, pp. xvi+636. MR: 0483954 (cit. on p. 2459).

Jianke Yang (2010). *Nonlinear waves in integrable and nonintegrable systems*. Vol. 16. Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, pp. xxvi+430. MR: 2721855 (cit. on pp. 2460–2462).

Norman J Zabusky and Martin D Kruskal (1965). "Interaction of "solitons" in a collisionless plasma and the recurrence of initial states". *Physical review letters* 15.6, pp. 240–243 (cit. on pp. 2457, 2459, 2460, 2462).

V. E. Zakharov and S. V. Manakov (1976). "Asymptotic behavior of non-linear wave systems integrated by the inverse scattering method". *Z. Èksper. Teoret. Fiz.* 71.1, pp. 203–215. MR: 0673411 (cit. on p. 2460).

V. E. Zakharov and A. B. Shabat (1971). "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media". *Ž. Èksper. Teoret. Fiz.* 61.1, pp. 118–134. MR: 0406174 (cit. on pp. 2457, 2460, 2461).

Received 2018-02-09.

Yvan Martel
CMLS, École Polytechnique, CNRS, 91128 Palaiseau Cedex, France
yvan.martel@polytechnique.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2485–2512)

DE GIORGI–NASH–MOSER AND HÖRMANDER THEORIES: NEW INTERPLAYS

Clément Mouhot

Abstract

We report on recent results and a new line of research at the crossroad of two major theories in the analysis of partial differential equations. The celebrated De Giorgi–Nash–Moser theorem provides Hölder estimates and the Harnack inequality for uniformly elliptic or parabolic equations with rough coefficients in divergence form. The theory of hypoellipticity of Hörmander provides general "bracket" conditions for regularity of solutions to partial differential equations combining first and second order derivative operators when ellipticity fails in some directions. We discuss recent extensions of the De Giorgi–Nash–Moser theory to hypoelliptic equations of Kolmogorov (kinetic) type with rough coefficients. These equations combine a first-order skew-symmetric operator with a second-order elliptic operator involving derivatives in only part of the variables, and with rough coefficients. We then discuss applications to the Boltzmann and Landau equations in kinetic theory and present a program of research with some open questions.

Contents

MSC2010: primary 35B65; secondary 82C40, 35R05, 35H10, 35Q20, 82D10.

1 Introduction

1.1 Kinetic theory. Modern physics goes back to Newton and classical mechanics, and was later expanded into the understanding of electric and magnetic forces were (Ampère, Faraday, Maxwell), large velocities and large scales (Lorentz, Poincaré, Minkowski, Einstein), small-scale particle physics and quantum mechanics (Planck, Einstein, Bohr, Heisenberg, Born, Jordan, Pauli, Fermi, Schrödinger, Dirac, De Broglie, Bose, etc.). However, all these theories are classically devised to study one physical system (planet, ship, motor, battery, electron, spaceship, etc.) or a small number of systems (planets in the Solar system, electrons in a molecule, etc.) In many situations though, one needs to deal with an assembly made up of elements so numerous that their individual tracking is not possible: galaxies made of hundreds of billions of stars, fluids made of more than 10^{20} molecules, crowds made of thousands of individuals, etc. Taking such large numbers into account leads to new effective laws of physics, requiring different models and concepts. This passage from microscopic rules to macroscopic laws is the founding principle of statistical physics. All branches of physics (classical, quantum, relativistic, etc.) can be studied from the point of view of statistical physics, in both stationary and dynamical perspectives. It was first done with the laws of classical mechanics, which resulted in kinetic theory, discovered by Maxwell [1867] and Boltzmann [1872] in the 19th century after precursory works by D. Bernoulli, Herapath, Waterston, Joule, König and Clausius.

Kinetic theory replaces a huge number of objects, whose physical states are described by a certain phase space, and whose properties are otherwise identical, by a *statistical distribution* over that phase space. The fundamental role played by the velocity (kinetic) variables inacessible to observation was counter-intuitive, and accounts for the denomination of *kinetic* theory. The theory introduces a distinction between three scales: the macroscopic scale of phenomena which are accessible to observation; the microscopic scale of molecules and infinitesimal constituents; and an intermediate scale, loosely defined and often called *mesoscopic*. This is the scale of phenomena which are not accessible to macroscopic observation but already involve a large number of particles, so that statistical effects are significant.

1.2 Main equations of kinetic theory. Maxwell wrote the first (weak) form of the evolution equation known now as the **Boltzmann equation**: the unknown is a (non-negative) density function $f(t,x,v)$, standing for the density of particles at time t in the phase space (x,v) (equipped with the reference Liouville measure $\mathrm{d}x\,\mathrm{d}v$); the equation, in modern writing and assuming the absence of external forces, is

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = Q(f,f). \tag{1.1}$$

The left-hand side describes the evolution of f under the action of transport with *free streaming operator*. The right-hand side describes elastic collisions with the nonlinear *Boltzmann collision operator*:

$$
\begin{aligned}
(1.2)\quad & Q(f,f) = \\
& = \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B(v-v_*,\omega)\Big(f(t,x,v')\,f(t,x,v'_*) - f(t,x,v)\,f(t,x,v_*)\Big)\,\mathrm{d}v_*\,\mathrm{d}\omega.
\end{aligned}
$$

Note that this operator is localized in t and x, quadratic, and has the structure of a tensor product with respect to $f(t,x,\cdot)$. The velocities v' and v'_* should be thought of as the velocities of a pair of particles before collision, while v and v_* are the velocities after that collision: the formulas are $v' = v - \langle v - v_*, \omega\rangle\,\omega$ and $v'_* = v_* + \langle v - v_*, \omega\rangle\,\omega$. When one computes (v, v_*) from (v', v'_*) (or the reverse), conservation laws of the mass, momentum and energy are not enough to yield the result, with only 4 scalar conservation laws for 6 degrees of freedom. The unit vector $\omega \in \mathbb{S}^2$ removes this ambiguity: in the case of colliding hard spheres, it can be thought of as the direction of the line joining the two centers of the particles. The kernel $B(v - v_*, \omega)$ describes the relative frequency of vectors ω, depending on the relative impact velocity $v - v_*$; it only depends on the modulus $|v - v_*|$ and the deflection angle θ between $v - v_*$ and $v' - v'_*$. Maxwell computed it for hard spheres ($B \sim |v - v_*|\sin\theta$) and for inverse power forces: in the latter case the kernel factorizes as the product of $|v - v_*|^\gamma$ with a function $b(\cos\theta)$; Maxwell showed that if the force is repulsive, proportional to $r^{-\alpha}$ (r the interparticle distance), then $\gamma = (\alpha-5)/(\alpha-1)$ and $b(\cos\theta) \simeq \theta^{-(1+2s)}$ as $\theta \to 0$, where $2s = 2/(\alpha-1)$. In particular, the kernel is usually *nonintegrable* as a function of the angular variable: this is a general feature of *long-range interactions*, nowadays sometimes called "noncutoff property".

The case $\alpha = 5$, $\gamma = 0$ and $2s = 1/2$ is called *Maxwell molecules* Maxwell [1867], the case $\alpha \in (5, +\infty)$, $\gamma > 0$ and $2s \in (0, 1/2)$ is called *hard potentials (without cutoff)*, the case $\alpha \in [3, 5)$, $\gamma \in [-1, 0)$, $2s \in (1/2, 1]$ is called *moderately soft potentials (without cutoff)*, and finally the case $\alpha \in (2, 3)$, $\gamma \in (-3, -1)$, $2s \in (1, 2)$ is called *very soft potentials (without cutoff)*. The limits between hard and soft potentials ($\gamma = 0$) and between moderately and very soft potentials ($\gamma + 2s = 0$) are commonly taken as defining the "hard" / "moderately soft" / "very soft" terminology in any dimension, for kernel of the form $B = |v - v_*|^\gamma b(\cos\theta)$ with $b(\cos\theta) \simeq \theta^{-(1+2s)}$.

In order to find the stationary solutions, that is, time-independent solutions of (1.2), the first step is to identify particular *hydrodynamic* density functions, which make the collision contribution vanish: these are *Gaussian distributions with a scalar covariance* $f(v) = \rho\,(2\pi T)^{-3/2}\,e^{-\frac{|v-u|^2}{2T}}$, where the parameters $\rho > 0$, $u \in \mathbb{R}^3$ and $T > 0$ are the local density, mean velocity, and temperature of the fluid. These parameters can be fixed throughout the whole domain (providing in this case an *equilibrium distribution*), or depend on the position x and time t; in both cases collisions will have no effect. As pointed out in Maxwell's seminar paper, and later proved rigorously at least in some settings Bardos, Golse, and Levermore [1991, 1993] and Golse and Saint-Raymond [2004, 2005, 2009], the Boltzmann equation is connected to classical fluid mechanical equations on ρ, u and T, and one leads to the other in certain regimes. This provides a rigorous connexion between the mesoscopic (kinetic) level and the macroscopic level. At the other end of the scales, a rigorous derivation of the Boltzmann equation from many-body Newtonian mechanics for short time and short-range interactions was obtained by Lanford [1975] for hard spheres; see also King [1975] for an extension to more general short-range interactions, and Gallagher, Saint-Raymond, and Texier [2013] and Pulvirenti, Saffirio, and Simonella [2014] for re-visitation and extension of the initial arguments of Lanford and King. Note however that at the moment the equivalent of Lanford theorem for the Boltzmann equation with long-range interactions is still missing, see Ayi [2017] for partial progresses.

To summarise the key mathematical points: the Boltzmann equation is an integro-(partial)-differential equation with non-local operator in the kinetic variable v. Moreover for long-range interactions with repulsive force $F(r) \sim r^{-\alpha}$, this non-local operator has a singular kernel and shows fractional ellipticity of order $2/(\alpha-1)$. The Boltzmann equation "contains" the hydrodynamic, and it is a *fundamental* equation in the sense that it is derived rigorously, at least in some settings, from microscopic first principles. From now on, we consider the position variable in $\mathbb{R}^3$ or in the periodic box $\mathbb{T}^3$.

In the limit case $s \to 1$ (the Coulomb interactions), the Boltzmann collision operator is ill-defined. Landau [1936] proposed an alternative operator for these Coulomb interactions that is now called the *Landau–Coulomb operator*

$$Q(f,f) =$$
$$= \nabla_v \cdot \left(\int_{\mathbb{R}^3} \mathbf{P}_{(v-v_*)^\perp} \Big(f(t,x,v_*) \nabla_v f(t,x,v) - f(t,x,v) \nabla_v f(t,x,v_*) \Big) |v-v_*|^{\gamma+2} \, \mathrm{d}v_* \right)$$

where $\mathbf{P}_{(v-v_*)^\perp}$ is the orthogonal projection along $(v-v_*)^\perp$ and $\gamma = -3$. It writes as

$$Q(f,f) = \nabla_v \cdot \big(A[f] \nabla_v f + B[f] f \big) \tag{1.3}$$

$$\text{with } \begin{cases} A[f](v) = \displaystyle\int_{\mathbb{R}^3} \left(I - \frac{w}{|w|} \otimes \frac{w}{|w|} \right) |w|^{\gamma+2} f(t,x,v-w) \, \mathrm{d}w, \\ B[f](v) = -\displaystyle\int_{\mathbb{R}^3} |w|^\gamma \, w \, f(t,x,v-w) \, \mathrm{d}w. \end{cases}$$

This operator is a nonlinear drift-diffusion operator with coefficients given by convolution-like averages of the unknown. This is a non-local integro-differential operator, with second-order local ellipticity. The resulting *Landau equation* (1.1)-(1.3) again "contains" the hydrodynamic. It is also considered *fundamental* because of its closed link to the Boltzmann equation for Coulomb interactions (note however that the equivalent to Lanford theorem for the Landau equation is lacking, even at a formal level, see Bobylev, Pulvirenti, and Saffirio [2013] for partial progresses). Because of the difficulty to handle the very singular kernel of the Landau–Coulomb operator, it is common to introduce artificially a scale of models by letting γ vary in $[-3, 1]$ (or even $[-d, 1]$ in general dimension d). The terminology *hard potentials*, *Maxwell molecules*, *soft potentials* are used as for the Boltzmann collision operator when $\gamma > 0$, $\gamma = 0$, $\gamma < 0$ respectively. The terminology *moderately soft potentials* corresponds here (since $s = 1$) to $\gamma \in (-2, 0)$.

1.3 Open problems and conjectures.

1.3.1 The Cauchy problem. The first mathematical question when studying the previous fundamental kinetic equations (Boltzmann and Landau equations) is the Cauchy problem, i.e. existence, uniqueness and regularity of solutions. Short-time solutions have been constructed, as well as global solutions close to the trivial stationary solution or with space homogeneity: see Gualdani, Mischler, and Mouhot [2017] for some of the most recent results and the references therein for the Boltzmann equation with short-range interactions, see Alexandre, Morimoto, Ukai, Xu, and Yang [2012, 2011a] and Gressman and Strain [2011] for the Boltzmann equation with long-range interactions, and see Guo [2002] for the Landau equation. However the construction of solutions "in the large" remains a formidable open problem. Since weak "renormalised" solutions have been constructed by DiPerna and Lions [1989b] that play a similar role to the Leray [1934] solutions in fluid mechanics, this open problem can be compared with the millenium problem of the regularity of solutions to 3D incompressible Navier–Stokes equations.

1.3.2 Study of a priori solutions. Given that the Cauchy problem in the large seems out of reach at the moment, Truesdell and Muncaster [1980] remarked almost 40 years ago that: "*Much effort has been spent toward proof that place-dependent solutions exist for all time. [...] The main problem is really to discover and specify the circumstances that give rise to solutions which persist forever. Only after having done that can we expect to construct proofs that such solutions exist, are unique, and are regular.*" Cercignani then formulated a precise conjecture on the entropy production along this idea in Cercignani [1982]; its resolution lead to precise new quantitative informations on a priori solutions of the Boltzmann and Landau equation (see Desvillettes and Villani [2005], Desvillettes, Mouhot, and Villani [2011], Mouhot [2006], Gualdani, Mischler, and Mouhot [2017], and Carrapatoso and Mischler [2017]). The proof of optimal relaxation rates in physical spaces, conditionnally to some regularity and moments conditions, is now fairly well understood for many interactions. The results obtained along this line of research can all be summarised into the following general form:

Conditional relaxation. Any solution to the Boltzmann (resp. Landau) equation in $L^\infty_x(\mathbb{T}^3; L^1_v(\mathbb{R}^3, (1 + |v|)^k \, \mathrm{d}v))$ (or a closely related functional space as large as possible) converges to the thermodynamical equilibrium with the optimal rate dictated by the linearized equation.

Note however that an interesting remaining open question in this program is to obtain a result equivalent to Gualdani, Mischler, and Mouhot [2017] and Carrapatoso and Mischler [2017] in the case of the Boltzmann equation with long-range interactions (with fractional ellipticity in the velocity variable).

1.3.3 Regularity conjectures for long-range interactions.

In the case of long-range interactions, the Boltzmann and Landau–Coulomb operators show local ellipticity provided the solution enjoys some pointwise bounds on the hydrodynamical fields

$$\rho(t, x) := \int_{\mathbb{R}^3} f \, \mathrm{d}v \qquad\qquad e(t, x) := \int_{\mathbb{R}^3} f |v|^2 \, \mathrm{d}v$$

and the local entropy $h(t, x) := \int_{\mathbb{R}^3} f \ln f \, \mathrm{d}v$. Whereas it is clear in the case of the Landau–Coulomb operator, it was understood almost two decades ago in the case of the Boltzmann collision operator Lions [1998], Villani [1999], and Alexandre, Desvillettes, Villani, and Wennberg [2000a]. This had lead colleagues working on non-local operators and fully nonlinear elliptic problems like L. Silvestre and N. Guillen and co-authors to attempt to use maximum principle techniques à la Krylov and Safonov [1980] in order to obtain pointwise bounds for solutions to these equations. These first attempts, while unsuccessful, later proved crucial in attracting the attention of a larger community on this problem. And these authors rapidly reformulated the initial goal into, again, *conditional* conjectures on the regularity of the form:

Conditional regularity. Consider any solution to the Boltzmann equation with long-range interactions (resp. Landau equation) on a time interval $[0, T]$ such that its hydrodynamical fields are bounded:

$$\forall\, t \in [0, T],\ x \in \mathbb{T}^3, \quad m_0 \le \rho(t, x) \le m_1, \quad e(t, x) \le e_1, \quad h(t, x) \le h_1 \tag{1.4}$$

where $m_0, m_1, e_1, h_1 > 0$. Then the solution is bounded and smooth on $(0, T]$.

Note that this conjecture can be strenghtened by removing the assumption that the mass is bounded from below and replacing it by a bound from below on the total mass

$\int_{\mathbb{T}^3} \rho(t,x)\,\mathrm{d}x \geq M_0 > 0$. Mixing in velocity through collisions combined with transport effects indeed generate lower bounds in many settings, see Mouhot [2005], Filbet and Mouhot [2011], Briant [2015a], and Briant [2015b]; moreover it was indeed proved for the Landau equation with moderately soft potentials in Henderson, Snelson, and Tarfulea [2017].

This conjecture is now been partially solved in the case of the Landau equation, when the interaction is "moderately soft" $\gamma \in (-2,0)$. This result has been the joint efforts of several groups Golse, Imbert, Mouhot, and Vasseur [2017], Henderson and Snelson [2017a], Henderson, Snelson, and Tarfulea [2017], and Imbert and Mouhot [2018], and this is the object of the next section. It is currently an ongoing program of research in the case of the Boltzmann equation with hard and moderately soft potentials, and this is the object of the fourth and last section. The conjecture interestingly remains open in the case of very soft potentials for both equations, and making progress in this setting is likely to require new conceptual tools.

2 De Giorgi–Nash–Moser meet Hörmander

2.1 The resolution of Hilbert 19-th problem. The De Giorgi–Nash–Moser theory De Giorgi [1956, 1957], Nash [1958], Moser [1960], and Moser [1964] was born out of the attempts to answer Hilbert's 19th problem. This problem is about proving the analytic regularity of the minimizers u of an energy functional $\displaystyle\int_U L(\nabla u)\,\mathrm{d}x$, with $u : \mathbb{R}^d \to \mathbb{R}$ and where the *Lagrangian* $L : \mathbb{R}^d \to \mathbb{R}$ satisfies growth, smoothness and convexity conditions and $U \subset \mathbb{R}^d$ is some compact domain. The Euler–Lagrange equations for the minimizers take the form

$$\nabla \cdot \Big[\nabla L(\nabla u)\Big] = 0 \quad \text{i.e.} \quad \sum_{i,j=1}^{d} \underbrace{\big[(\partial_{ij} L)(\nabla u)\big]}_{b_{ij}} \partial_{ij} u = \sum_{i,j=1}^{d} b_{ij} \partial_{ij} u = 0.$$

For instance L can be the Dirichlet energy $L(p) = |p|^2$, or be nonlinear as for instance $L(p) = \sqrt{1+|p|^2}$ for minimal surfaces. With suitable assumptions on L and the domain, the pointwise control of ∇u was known. However the existence, uniqueness and regularity requires more: if $u \in C^{1,\alpha}$ with $\alpha > 0$ then $b_{ij} \in C^{\alpha}$ and Schauder estimates Schauder [1934] imply $u \in C^{2,\alpha}$. Then a bootstrap argument yields higher regularity, and analyticity follows from this regularity Bernstein [1904] and Petrowsky [1939].

Hence, apart from specific result in two dimensions Morrey [1938], the missing piece in solving Hilbert 19th problem, in the 1950s, was the proof of the Hölder regularity of ∇u.

The equation satisfied by a derivative $f := \partial_k u$ is the divergence form elliptic equation:

$$\sum_{i,j=1}^{d} \partial_i \Big[\underbrace{(\partial_{ij} L)(\nabla u)}_{a_{ij}} \partial_j f \Big] = \nabla \cdot (A \nabla f) = 0.$$

De Giorgi [1957] and Nash [1958] independently proved this Hölder regularity of f under the sole assumption that the symmetric matrix $A := (a_{ij})$ satisfies the controls $0 < \lambda \leq A \leq \Lambda$, and is measurable (no regularity is assumed). The proof of Nash uses what is now called the "Nash inequality", an $L \log L$ energy estimate, and refined estimates on the fundamental solution. The proof of De Giorgi uses an iterative argument to gain integrability, and an "isoperimetric argument" to control how oscillations decays when refining the scale of observation. Moser later gave an alternative proof Moser [1960] and Moser [1964] based on one hand on an iterative gain of integrability, formulated differently but similar to that of De Giorgi, and on the other hand on relating positive and negative Lebesgue norms through energy estimate on the equation satisfied by $g := \ln f$ and the use of a *Poincaré inequality*; the proof of Moser had and important further contribution in that it also proved the *Harnack inequality* for the equations considered, i.e. a universal control on the ratio between local maxima and local minima.

Let us mention that the De Giorgi–Nash–Moser (DGNM) theory only considers elliptic or parabolic equations in *divergence form*. An important counterpart result for non-divergence elliptic and parabolic equations was later discovered by Krylov and Safonov [1980]. The extension of the DGNM theory to hypoelliptic equations with rough coefficients that we present in this section requires the equation to be in divergent form. It is an open problem whether the Krylov–Safonov theory extends to hypoelliptic non-divergent equations of the form discussed below.

2.2 The theory of hypoellipticity. The DGNM theory has revolutionised the study of nonlinear elliptic and parabolic partial differential equations (PDEs). However it remained limited to PDEs where the diffusion acts in all directions of the phase space. In kinetic theory, as soon as the solution is non spatially homogeneous, the diffusion or fractional diffusion in velocity is combined to a conservative Hamiltonian dynamic in position and velocity. This structure is called *hypoelliptic*. It can be traced back to the short note of Kolmogoroff [1934]. The latter considered the combination of free transport with drift-diffusion in velocity: the law satisfies what is now sometimes called the *Kolmogorov equation*, that writes $\partial_t f + v \cdot \partial_x f = \Delta_v f$ on $x, v \in \mathbb{R}^d$ in the simpler case. It is the equation satisfied by the law of a Brownian motion integrated in time. Kolmogorov then wrote the fundamental solution associated with a Dirac distribution δ_{x_0,v_0} initial data:

$$G(t,x,v) = \left(\frac{\sqrt{3}}{2\pi t^2}\right)^d \exp\left\{-\frac{3|x-x_0-tv_0+t(v-v_0)/2|^2}{t^3} - \frac{|v|^2}{4t}\right\}.$$

The starting point of Hörmander's seminal paper Hörmander [1967] is the observation that this fundamental solution shows regularisation in all variables, even though the diffusion acts only in the velocity variable. The regularisation in (t,x) is produced by the interaction between the transport operator $v\cdot\nabla_x$ and the diffusion in v. Hörmander's paper then proposes precise geometric conditions for this regularisation, called *hypoelliptic*, to hold, based on commutator estimates. In short, given X_0, X_1, …, X_n a collection of smooth vector fields on $\mathbb{R}^N$ and the second-order differential operator $L = \frac{1}{2}\sum_{i=1}^n X_i^2 + X_0$, then the semigroup e^{tL^*} is regularising (hypoelliptic) as soon as the Lie algebra generated by X_0, …, X_n has dimension N throughout the domain of L.

Let us also mention the connexion with the *Malliavin calculus* in probability, which gives a probabilistic proof to the Hörmander theorem in many settings, see Malliavin [1978] as well as the many subsequent works, for instance Kusuoka and Stroock [1984], Kusuoka and Stroock [1985], Bismut [1981], and Norris [1986].

2.3 Extending the DGNM theory to hypoelliptic settings. The main question of interest here is the extension of the DGNM theory to hypoelliptic PDEs of divergent type. Hypoelliptic PDEs of second order naturally split into two classes: "type I" when the operator is a sum of squares of vector fields ($X_0 = 0$ in the Hörmander form described above), and the "type II" such as the Kolmogorov above, where the operator combines a first-order anti-symmetric operator with some partially diffusive second-order operator. Two main research groups had already been working on the question. On the one hand, Polidoro and collaborators Polidoro [1994], Manfredini and Polidoro [1998], Polidoro and Ragusa [2001], Pascucci and Polidoro [2004], and Di Francesco and Polidoro [2006] had generalised the DGNM theory to the "type I" equations and had obtained the improvement of integrability for the "type II" equations, as well as the Hölder regularity when assuming some continuity property on the coefficients. On the other hand, Wang and Zhang [2009, 2011] and Zhang [2011] had extended the proof of Moser for the "type II" equations to obtain Hölder regularity, with intricate technical calculations that did not seem easy to export. Note also that the use of the DGNM theory in kinetic theory had also been advocated almost a decade before in the premonitory lecture notes Villani [2003].

We present here the work Golse, Imbert, Mouhot, and Vasseur [2017] (see also the two previous related preprints Golse and Vasseur [2015] and Imbert and Mouhot [2015]) that (1) provides an elementary and robust proof of the gain of integrability and Hölder regularity in this "type II" hypoelliptic setting, (2) proves the stronger *Harnack inequality* for these equations (i.e. a quantitative version of the strong maximum principle).

Let us consider the following kinetic Fokker–Planck equation

$$\partial_t f + v \cdot \nabla_x f = \nabla_v \cdot (A \nabla_v f) + B \cdot \nabla_v f + s, \quad t \in (0,T),\ (x,v) \in \Omega, \tag{2.1}$$

where Ω is an open set of $\mathbb{R}^{2d}$, $f = f(t,x,v)$, B and s are bounded measurable coefficients depending in (t,x,v), and the $d \times d$ real matrices A, B and source term s are measurable and satisfy

$$0 < \lambda I \le A \le \Lambda I, \qquad |B| \le \Lambda, \qquad s \text{ essentially bounded} \tag{2.2}$$

for two constants λ, $\Lambda > 0$. Given $z_0 = (x_0, v_0, t_0) \in \mathbb{R}^{2d+1}$, we define the "cylinder" $Q_r(z_0)$ centered at z_0 of radius r that respects the invariances of the equation:

$$Q_r(z_0) := \left\{ (x,v,t) : |x - x_0 - (t-t_0)v_0| < r^3, |v - v_0| < r, t \in \left(t_0 - r^2, t_0\right] \right\}. \tag{2.3}$$

The *weak solutions* to equation (2.1) on $U_x \times U_v \times I$, $U_x \subset \mathbb{R}^d$ open, $U_v \subset \mathbb{R}^d$ open, $I = [a,b]$ with $-\infty < a < b \le +\infty$, are defined as functions $f \in L^\infty_t(I, L^2_{x,v}(U_x \times U_v))) \cap L^2_{x,t}(U_x \times I, H^1_v(U_v))$ such that $\partial_t f + v \cdot \nabla_x f \in L^2_{x,t}(U_x \times I, H^{-1}_v(U_v))$ and f satisfies the equation (2.1) in the sense of distributions.

Theorem 1 (Hölder continuity Golse, Imbert, Mouhot, and Vasseur [2017]). *Let f be a weak solution of (2.1) in $\mathbb{Q}_0 := Q_{r_0}(z_0)$ and $\mathbb{Q}_1 := Q_{r_1}(z_0)$ with $r_1 < r_0$. Then f is α-Hölder continuous with respect to (x,v,t) in $\mathbb{Q}_1$ and*

$$\|f\|_{C^\alpha(\mathbb{Q}_1)} \le C \left(\|f\|_{L^2(\mathbb{Q}_0)} + \|s\|_{L^\infty(\mathbb{Q}_0)} \right)$$

for some $\alpha \in (0,1)$ and $C > 0$ only depending on d, λ, Λ, r_0, r_1 (plus z_0 for C).

In order to prove such a result, we first prove that L^2 sub-solutions are locally bounded; we refer to such a result as an $L^2 - L^\infty$ estimate. We then prove that solutions are Hölder continuous by proving a lemma which is a hypoelliptic counterpart of De Giorgi's "isoperimetric lemma". We moreover prove the Harnack inequality:

Theorem 2 (Harnack inequality Golse, Imbert, Mouhot, and Vasseur [ibid.]). *If f is a non-negative weak solution of (2.1) in $Q_1(0,0,0)$, then*

$$\sup_{Q^-} f \le C \left(\inf_{Q^+} f + \|s\|_{L^\infty(Q_1(0,0,0))} \right) \tag{2.4}$$

where $Q^+ := Q_R(0,0,0)$ and $Q^- := Q_R(0,0,-\Delta)$ and $C > 1$ and $R, \Delta \in (0,1)$ are small (in particular $Q^\pm \subset Q_1(0,0,0)$ and they are disjoint), and universal, i.e. only depend on dimension and ellipticity constants.

Remark 3. Using the transformation $\mathcal{T}_{z_0}(x, v, t) = (x_0 + x + tv_0, v_0 + v, t_0 + t)$, we get a Harnack inequality for cylinders centered at an arbitrary point $z_0 = (x_0, v_0, t_0)$.

Our proof combines the key ideas of De Giorgi and Moser and the *velocity averaging method*, which is a special type of smoothing effect for solutions of the free transport equation $(\partial_t + v \cdot \nabla_x) f = S$ observed for the first time in Agoshkov [1984] and Golse, Perthame, and Sentis [1985] independently, later improved and generalized in Golse, Lions, Perthame, and Sentis [1988] and DiPerna and Lions [1989a]. This smoothing effect concerns averages of f in the velocity variable v, i.e. expressions of the form $\int_{\mathbb{R}^d} f(x, v, t)\, \phi(v)\, \mathrm{d}v$ with, say, $\phi \in C_c^\infty$. Of course, no smoothing on f itself can be observed, since the transport operator is hyperbolic and propagates the singularities. However, when S is of the form $S = \nabla_v \cdot (A(x, v, t) \nabla_v f) + s$, where s is a given source term in L^2, the smoothing effect of velocity averaging can be combined with the H^1 regularity in the v variable implied by the energy inequality in order to obtain some amount of smoothing on the solution f itself. A first observation of this type (at the level of a compactness argument) can be found in Lions [1994]; Bouchut [2002] had then obtained quantitative Sobolev regularity estimates.

Our proof of the $L^2 - L^\infty$ gain of integrability follows the so-called “De Giorgi–Moser iteration”, see Golse, Imbert, Mouhot, and Vasseur [2017] where it is presented in both the equivalent presentations of De Giorgi and of Moser. We emphasize that, in both approaches, the main ingredient is a local gain of integrability of non-negative sub-solutions. This latter is obtained by combining a comparison principle and a Sobolev regularity estimate following from the velocity averaging method discussed above and energy estimates. We then prove the Hölder continuity through a De Giorgi type argument on the decrease of oscillation for solutions. We also derive the Harnack inequality by combining the decrease of oscillation with a result about how positive lower bounds on non-negative solutions deteriorate with time. It is worth mentioning here that our “hypoelliptic isoperimetric argument” is proved non-constructively, by a contradiction method, whereas the original isoperimetric argument of De Giorgi is obtained by a quantitative direct argument. It is an interesting open problem to obtain such quantitative estimates in the hypoelliptic case.

3 Conditional regularity of the Landau equation

3.1 Previous works and a conjecture. The infinite smoothing of solutions to the Landau equation has been investigated so far in two different settings. On the one hand, it has been investigated for weak spatially homogeneous solutions (non-negative in L^1 and with finite energy) see Desvillettes and Wennberg [2004] and the subsequent follow-up papers Alexandre and El Safadi [2005], Huo, Morimoto, Ukai, and Yang [2008], Alexandre, Morimoto, Ukai, Xu, and Yang [2008], Alexandre and Elsafadi [2009], Morimoto, Ukai, Xu, and Yang [2009], Arsen'ev and Buryak [1990], Desvillettes [2004], Villani [1998], and Desvillettes and Villani [2000a], and see also the related entropy dissipation estimates in Desvillettes and Villani [2000b] and Desvillettes [2015], and see the analytic regularisation of weak spatially homogeneous solutions for Maxwellian or hard potentials in H. Chen, Li, and Xu [2010]. Furthermore, Silvestre [2016b] derives an L^∞ bound (gain of integrability) for spatially homogeneous solutions in the case of moderately soft potentials without relying on energy methods. Let us also mention works studying modified Landau equations Krieger and Strain [2012] and Gressman, Krieger, and Strain [2012] and the work Gualdani and Guillen [2016] that shows that any weak radial solution to the Landau–Coulomb equation that belongs to $L^{3/2}$ is automatically bounded and C^2 using barrier arguments. On the other hand, the investigations of the regularity of classical spatially heterogeneous solutions had been more sparse, focusing on the regularisation of classical solutions see Y. Chen, Desvillettes, and He [2009] and Liu and Ma [2014].

The general question of conditional regularity hence suggests the following question in the context of the Landau equation:

Conjecture 1. Any solutions to the Landau equation (1.1)- (1.3) (with Coulomb interaction $\gamma = -3$) on $[0, T]$ satisfying (1.4) is bounded and smooth on $(0, T]$.

An important progress has been made by solving a weaker version of this conjecture when the exponent $\gamma \in (-2, 0)$, which corresponds to *moderately soft potentials*, i.e. $\gamma + 2s > 0$ since here $s = 1$. We describe in this section the different steps and combined efforts of different groups.

3.2 DGNM theory and local Hölder regularity. The first step was the work Golse, Imbert, Mouhot, and Vasseur [2017] already mentioned. A corollary of the general regularity theorem, Theorem 1, is the following:

Theorem 4 (Local Hölder regularity for the LE Golse, Imbert, Mouhot, and Vasseur [ibid.]). *Given any $\gamma \in [-3, 1]$, there are universal constants $C > 0$, $\alpha \in (0, 1)$ such that any f essentially bounded weak solution of* (1.1)- (1.3) *in $B_1 \times B_1 \times (-1, 0]$ satisfying* (1.4) *is α-Hölder continuous with respect to $(x, v, t) \in B_{\frac12} \times B_{\frac12} \times (-\frac12, 0]$ and*

$$\|f\|_{C^\alpha(B_{\frac12}\times B_{\frac12}\times(-\frac12,0])} \le C\left(\|f\|_{L^2(B_1\times B_1\times(-1,0])} + \|f\|^2_{L^\infty(B_1\times B_1\times(-1,0])}\right).$$

Note that this theorem includes the physical case of Coulomb interactions $\gamma = -3$. The adjective "universal" for the constants refers to their independence from the solution.

3.3 Maximum principles and pointwise bounds. This line of research originates in the work of L. Silvestre both on the spatially homogeneous Boltzmann (SHBE) and Landau (SHLE) equations Silvestre [2016a, 2017]. These papers build upon the ideas of "nonlinear maximal principles" introduced in Constantin and Vicol [2012] in the case of the Boltzmann collision operator, and upon the so-called "Aleksandrov–Bakelman–Pucci Maximum Principle" in the case of the Landau collision operator, see for instance Caffarelli and Cabré [1995] and Caffarelli and Silvestre [2009].

The main result of Silvestre [2017] is:

Theorem 5 (Pointwise bound for the SHLE). *Let $\gamma \in [-2, 0]$ (moderately soft potentials) and f a classical spatially homogeneous solution to the Landau equation* (1.1)- (1.3) *satisfying the assumptions* (1.4)*. Then $f \lesssim 1 + t^{-3/2}$ with constant depending only on the bounds* (1.4)*.*

As noted by the author, this estimate implies quite straightforwardly existence, uniqueness and infinite regularity for the spatially homogeneous solution. For the difficult case of very soft potentials $\gamma \in [-3, 2)$, this paper includes a weaker result where the L^∞ bound depends on a certain weighted Lebesgue norms; unfortunately it is not yet known how to control such norm along time. This conceptual barrier when crossing the "very soft potentials threshold" is reminiscent of the situation for the Cauchy theory in Lebesgue and Sobolev spaces of both the spatially homogeneous Boltzmann with long-range interactions Desvillettes and Mouhot [2009] and Landau equation Wu [2014].

The pointwise bounds estimates were then extended to the spatially inhomogeneous case in Cameron, Silvestre, and Snelson [2017]. The main result in this latter paper is:

Theorem 6 (Pointwise bound for the LE). *Let $\gamma \in (-2, 0]$ (moderately soft potentials without the limit case) and f a bounded weak solution to the Landau equation* (1.1)- (1.3) *satisfying the assumptions* (1.4)*. Then $f \lesssim (1 + t^{-3/2})(1 + |v|)^{-1}$ with constant depending only on the bounds* (1.4) *(and* not *on the L^∞ norm of the solution). Moreover if $f_{in}(x, v) \le C_0 e^{-\alpha |v|^2}$, for some $C_0 > 0$ and a sufficiently small $\alpha > 0$ (depending on γ and* (1.4)*), then $f(t, x, v) \le C_1 e^{-\alpha |v|^2}$ with $C_1 > 0$ depending only on C_0, γ and the bounds* (1.4)*.*

The proof relies on using locally the Harnack inequality in Theorem 2 adapted to the Landau equation and on devising a clever change of variable to track how this local estimate behaves at large velocities. The Gaussian bound is then obtained by combining existing maximum principle arguments at large velocities (using that well-constructed Gaussians provide supersolutions at large v) in the spirit of Gamba, Panferov, and Villani [2009], and the previous pointwise bound for not-so-large velocities. Finally the authors remarked that the Hölder regularity estimate of Theorem 4 can be made global using the Gaussian decay bound.

3.4 Schauder estimates and higher regularity. Once the L^∞ norm and the Hölder regularity is under control, the next step is to obtain higher-order regularity. The classical tool is the so-called *Schauder estimates* Schauder [1934]. The purpose of such estimates in general is to show that the solution to an elliptic or parabolic equation whose coefficients are Hölder continuous gains two derivatives with respect to the data (source term, initial data). The gain of the two derivatives is obtained in Hölder spaces: $C^\delta \to C^{2+\delta}$.

Two works have been obtained independently along this line of research. The first one Henderson and Snelson [2017b] focuses on the use of combination of Hölder estimates, maximum principles and Schauder estimates to obtain conditional infinite regularity for solutions to the Landau equation with moderately soft potentials $\gamma \in (-2, 0)$. The second one Imbert and Mouhot [2018] focuses on the use of these ingredients to "break the super-criticality" of the nonlinearity for a toy model of the Landau equation. Both these works develop, in different technical ways, Schauder estimates for this hypoelliptic equation.

The main result in Henderson and Snelson [2017b] is:

Theorem 7 (Conditional regularity for LE)**.** *Let $\gamma \in (-2, 0)$ (moderately soft potentials without the limit case) and f a bounded weak solution to the Landau equation* (1.1)- (1.3) *satisfying the assumptions* (1.4) *and $f_{in}(x, v) \leq C_0 e^{-\alpha|v|^2}$, for some $C_0 > 0$ and a sufficiently small $\alpha > 0$ (depending on γ and* (1.4)*). Then f is smooth and its derivatives have some (possibly weaker) Gaussian decay.*

Note that (1) the regularity and decay bounds are uniform in time, as long as the bounds (1.4) remain uniformly bounded in time, (2) further conditional regularity are given in the paper for very soft potentials $\gamma \in [-3, -2]$ but they require higher $L^\infty_{t,x} L^1_v$ moments and the constants depend on time when $\gamma \in [-3, -5/2]$ in dimension 3, (3) a useful complementary result is provided by Henderson, Snelson, and Tarfulea [2017] where a local existence is proved in weighted locally uniform Sobolev spaces and the lower bound on the mass is relaxed by using the regularity to find a ball where the solution is uniformly positive: the combination of the two papers provide a conditional existence, uniqueness and regularity result for soft potentials, conditionally to upper bounds on the local mass, energy and entropy.

The work Imbert and Mouhot [2018] considers the toy model:

$$\partial f + v \cdot \nabla_x f = \rho[f] \nabla_v \left(\nabla_v f + v f \right), \quad \rho[f] := \int_{\mathbb{R}^d} f \, \mathrm{d}v, \tag{3.1}$$

in $x \in \mathbb{T}^d$, $v \in \mathbb{R}^d$, $d \geq 1$. This model preserves the form of the steady state, the ellipticity in v, the non-locality, the bilinearity and the mass conservation of the LE. It however greatly simplifies the underlying hydrodynamic and the maximum principle structure. Here $H^k(\mathbb{T}^d \times \mathbb{R}^d)$ denotes the standard L^2-based Sobolev space. The main result states (note that solutions are constructed and not conditional here):

Theorem 8. *For all initial data f_{in} such that $f_{in}/\sqrt{\mu} \in H^k(\mathbb{T}^d \times \mathbb{R}^d)$ with $k > d/2$ and satisfying $C_1\mu \le f_{in} \le C_2\mu$, there exists a unique global-in-time solution f of (3.1) with initial data f_{in} satisfying for all time $t > 0$: $f(t)/\sqrt{\mu} \in H^k(\mathbb{T}^d \times \mathbb{R}^d)$ and $C_1\mu \le f \le C_2\mu$ and $f(t,\cdot,\cdot) \in C^\infty$.*

Note that the initial regularity could be relaxed with more work. A key step of the proof is the Schauder estimate. It gives the following additional information on this solution: the *hypoelliptic Hölder norm* $\mathcal{H}^\alpha$ (defined below) of $f/\sqrt{\mu}$ is uniformly bounded in terms of the L^2 norm of $f_{in}/\sqrt{\mu}$ for times away from 0. This norm is defined on a given open connected set $\mathcal{Q}$ by

$$\|g\|_{\mathcal{H}^\alpha(\mathcal{Q})} := \sup_{\mathcal{Q}} |g| + \sup_{\mathcal{Q}} |(\partial_t + v\cdot\nabla_x)g| + \sup_{\mathcal{Q}} |D_v^2 g| + [(\partial_t + v\cdot\nabla_x)g]_{\mathcal{C}^{0,\alpha}(\mathcal{Q})} + [D_v^2 g]_{\mathcal{C}^{0,\alpha}(\mathcal{Q})}$$

where $[\cdot]_{\mathcal{C}^{0,\alpha}(\mathcal{Q})}$ is a Hölder anisotropic semi-norm, i.e. the smallest $C > 0$ such that

$$\forall\, z_0 \in \mathcal{Q},\ r > 0 \text{ s.t. } Q_r(z_0) \subset \mathcal{Q}, \quad \|g - g(z_0)\|_{L^\infty(Q_r(z_0))} \le C r^\alpha$$

where

$$\begin{aligned} Q_r(z_0) &:= \left\{ z : \frac{1}{r}(z_0^{-1} \circ z) \in Q_1 \right\} \\ &= \left\{ (t,x,v) : t_0 - r^2 < t \le t_0,\ |x - x_0 - (t-t_0)v_0| < r^3,\ |v - v_0| < r \right\} \end{aligned}$$

and $rz := (r^2 t, r^3 x, rv)$ and $z_1 \circ z_2 := (t_1 + t_2, x_1 + x_2 + t_2 v_1, v_1 + v_2)$.

The specific contribution of this work is the study of the Cauchy problem: the maximum principle provides Gaussian upper and lower bounds on the solution, and we then provide energy estimates and a blow-up criterion *à la* Beale, Kato, and Majda [1984]. We then use the extensions of the DGNM and Schauder theories to control the quantity governing the blow-up. We prove Hölder regularity through the method of Golse, Imbert, Mouhot, and Vasseur [2017]. We then develop Schauder estimates following the method of Krylov [1996] (see also Polidoro [1994], Manfredini [1997], Di Francesco and Polidoro [2006], Bramanti and Brandolini [2007], Lunardi [1997], Radkevich [2008], and Henderson and Snelson [2017a]). New difficulties arise compared with the parabolic case treated in Krylov [1996] in relation with the hypoelliptic structure and we develop trajectorial hypoelliptic commutator estimates to solve them and also borrow some ideas from hypocoercivity Villani [2009] in the so-called gradient estimate.

Note that it would be interesting to give a proof of Schauder estimates for such hypoelliptic equations that is entirely based on scaling arguments in the spirit Simon [1997] (see also the proof and use of such estimates in Hairer [2014]). This might indeed prove useful for generalising such estimates to the integral Boltzmann collision operator, see the next section.

4 Conditional regularity of the Boltzmann equation

4.1 Previous works and a conjecture. Short time existence of solutions to (1.1)- (1.2) was obtained in Alexandre, Morimoto, Ukai, Xu, and Yang [2010a] for sufficiently regular initial data f_0. Global existence was obtained in Desvillettes and Mouhot [2009] for moderately soft potentials in the spatially homogeneous case. In the next subsections, we present the progresses made so far in the case of moderately soft potentials: the estimate in L^∞ for $t > 0$ was obtained in Silvestre [2016a], the local Hölder regularity in Imbert and Silvestre [2017], and finally the polynomial pointwise decay estimates in Imbert, Mouhot, and Silvestre [2018]. The bootstrap mechanism to obtain higher regularity through Schauder estimates remains however unsolved at now.

Let us briefly review the existing results about regularisation. In Alexandre, Morimoto, Ukai, Xu, and Yang [2010a], the authors prove that if the solution f has five derivatives in L^2, with respect to all variables t, x and v, weighted by $(1 + |v|)^q$ for arbitrarily large powers q, and in addition the mass density is bounded below, then the solution f is C^∞. It is not known however whether these hypotheses are implied by (1.4). Note also the previous partial result Desvillettes and Wennberg [2004] and the subsequent follow-up papers Alexandre and El Safadi [2005], Huo, Morimoto, Ukai, and Yang [2008], Alexandre, Morimoto, Ukai, Xu, and Yang [2008], Alexandre and Elsafadi [2009], and Morimoto, Ukai, Xu, and Yang [2009] in the spatially homogeneous case, with less assumptions on the initial data.

Note that, drawing inspiration from the case of the Landau equation, in order for the iterative gain of regularity in Henderson and Snelson [2017b] to work, it is necessary to start with a solution that decays, as $|v| \to \infty$, faster than any algebraic power rate $|v|^{-q}$. We expect the same general principle to apply to the Boltzmann equation, even the appropriate Schauder type estimates for kinetic integro-differential equations to carry out an iterative gain in regularity are not yet available.

Finally, we highlight the related results of regularisation for the Boltzmann equation with long-range interactions Desvillettes [1995] and Y. Chen and He [2011, 2012], and the related perturbative results for the Landau and (long-range interaction) Boltzmann equation Guo [2002], Gressman and Strain [2011], Alexandre, Morimoto, Ukai, Xu, and Yang [2010b, 2011b], Alexandre [2009], Wu [2014], and Alexandre, Liao, and Lin [2015].

The question of conditional regularity suggests the following conjecture in the context of the Boltzmann equation with long-range interactions:

Conjecture 2. Any solutions to the Boltzmann equation (1.1)- (1.2) with long-range interactions $\gamma \in (-3, 1]$, $s \in (0, 1)$, $\gamma + 2s \in (-1, 1)$) on $[0, T]$ satisfying (1.4) is bounded and smooth on $(0, T]$.

The rest of this section is devoted to describing the partial progresses made in the case of, again, moderately soft potentials $\gamma + 2s > 0$.

4.2 Maximum principle and pointwise L^∞ bound. This first breakthrough is due to Silvestre [2016a]. This article draws inspiration from his own previous works on non-local operators and the "nonlinear maximum principle" of Constantin and Vicol [2012]. It is based on a maximum principle argument for a barrier supersolution that is constant in x, v and blowing-up as $t \to 0^+$; it uses the decomposition of the collision and "cancellation lemma" going back to Alexandre, Desvillettes, Villani, and Wennberg [2000b], the identification of a cone of direction for $(v' - v)$ is order to obtain lower bounds on the f-dependent kernel of the elliptic part of the operator, and finally some Chebycheff inequality and nonlinear lower bound on the collision integral. The main result is:

Theorem 9 (Pointwise bound for the BE Silvestre [2016a])**.** *Let $\gamma \in [-2, 1]$, $s \in (0, 1)$ with $\gamma + 2s > 0$ (moderately soft potentials). Let f be a classical solution to the Boltzmann equation* (1.1)- (1.3) *satisfying the assumptions* (1.4)*. Then $f \le C(1 + t^{-\beta})$ with $C > 0$ and $\beta > 0$ and constant depending only on γ, s and the bounds* (1.4).

Note that the paper also includes further results in the case of very soft potentials but conditionally to additional estimates of the form $L^\infty_{t,x} L^p_v (1 + |v|^q)$ for some $p > 1, q > 0$; it is not known at present how to deduce the latter estimates from the hydrodynamic bounds (1.4).

4.3 Weak Harnack inequality and local Hölder regularity. The second breakthrough is the paper Imbert and Silvestre [2017]. In comparison to the Landau equation, the Boltzmann equation has a more complicated integral structure, that shares similarity with "fully nonlinear" fractional elliptic operators. The main result proved is:

Theorem 10 (Local Hölder regularity for the BE Imbert and Silvestre [ibid.])**.** *Given any $\gamma \in (-3, 1]$ and $s \in (0, 1)$ with $\gamma + 2s > 0$, there are universal constants $C > 0$, $\alpha \in (0, 1)$ such that any f essentially bounded weak solution of* (1.1)- (1.2) *in $B_1 \times \mathbb{R}^3 \times (-1, 0]$ satisfying* (1.4) *is α-Hölder continuous with respect to $(x, v, t) \in B_{\frac{1}{2}} \times B_{\frac{1}{2}} \times (-\frac{1}{2}, 0]$, where $C > 0$ and $\alpha \in (0, 1)$ are constants depending on the L^∞ bound of f and the bounds* (1.4).

The proof goes in two steps. The first step is a local $L^2 \to L^\infty$ gain of integrability, following the approach of De Giorgi and Moser as reformulated in a kinetic context in Pascucci and Polidoro [2004] and Golse, Imbert, Mouhot, and Vasseur [2017]. It requires further technical work to formulate the De Giorgi iteration for such integro-differential equations with degenerate kernels (see also the related works Kassmann [2009], Felsinger and Kassmann [2013], and Bux, Kassmann, and Schulze [2017]). The regularity mechanism at the core of the averaging velocity method is used, however under a different presentation, by relying on explicit calculations on the fundamental solution of the fractional Kolmogorov equation. In the second step of the proof, the authors establish a weak Harnack inequality, i.e. an the local control from above of local $L^{\epsilon}_{t,x,v}$ averages with $\epsilon > 0$ small by a local infimum multiplied by a universal constant. This inequality is sufficient to deduce the Hölder regularity. Two different strategies are used depending on whether $s \in (0, 1/2)$ or $s \in [1/2, 1)$. In the first case, they construct a barrier function to propagate lower bounds as in the method by Krylov and Safonov for nondivergence equations. In the second case, they use a variant of the isometric argument of De Giorgi proved by compactness as in Golse, Imbert, Mouhot, and Vasseur [2017]. Again the regularity of velocity averages plays a crucial role but is exploited by direct calculation on the fundamental solution of the fractional Kolmogorov equation.

4.4 Maximum principle and decay at large velocities. Finally in the paper Imbert, Mouhot, and Silvestre [2018], the nonlinear maximum principle argument of Silvestre [2016a] is refined to obtain "pointwise counterpart" of velocity moments.

Let us recall that in order for an iterative gain of regularity similar to Henderson and Snelson [2017b] and Imbert and Mouhot [2018] to work, it is necessary to start with a solution that decays, as $|v| \to \infty$, faster than any algebraic power rate $|v|^{-q}$, and we expect the same general to be true for the Boltzmann equation. The main result established in this paper is:

Theorem 11 (Decay at large velocities for the BE Imbert, Mouhot, and Silvestre [2018]). *Given any $\gamma \in (-3, 1]$ and $s \in (0, 1)$ with $\gamma + 2s > 0$, there are universal constants $C > 0$, $\alpha \in (0, 1)$ such that for any f classical solution of* (1.1)- (1.2) *in $\mathbb{T}^3 \times \mathbb{R}^3 \times [0, T]$ satisfying* (1.4) *it holds for any $q > 0$: (i) pointwise polynomial decay is propagated: if $f_{in} \lesssim (1+|v|)^{-q}$ then for all $t > 0$, $x \in T^3$ then $f(x, v, t) \leq C(1+|v|)^{-q}$, (ii) if $\gamma > 0$ all the polynomial moments are generated: $f(x, v, t) \leq C'(1+t^{-\beta})(1+|v|)^{-q}$. All the constants depend on γ, s, q and the bounds* (1.4).

The study of large velocity decay, known as the study of *moments*, is an old and important question in kinetic equations. The study of moments was initiated for Maxwellian potentials ($\gamma = 0$) in Ikenberry and Truesdell [1956]. In the case of hard potentials ($\gamma > 0$), Povzner identities Povzner [1962], Elmroth [1983], Wennberg [1996], and Bobylev [1997] play an important role. For instance, Elmroth [1983] used them to prove that if moments are initially bounded, then they remain bounded for all times. Desvillettes [1993] then proved that only one moment of order $s > 2$ is necessary for the same conclusion to hold true. It is explained in Wennberg [1996] and Mischler and Wennberg [1999] that even the condition on one moment of order $s > 2$ can be dispensed with, in both (homogeneous) cutoff and non-cutoff case. These moment estimates were used by Bobylev [1997] in order to derive (integral) Gaussian tail estimates. In the case of soft potentials, Desvillettes [1993] proved for $\gamma \in (-1,0)$ that initially bounded moments grow at most linearly with time and it is explained in Villani [2002] that the method applies to $\gamma \in [-2,0)$. The case of measure-valued solutions is considered in Lu and Mouhot [2012].

However the extension of these integral moments estimates to the spatially inhomogeneous case is a hard and unclear question at the moment. The only result available is Gualdani, Mischler, and Mouhot [2018, Lemma 5.9 & 5.11] which proves the propagation and appearance of certain exponential moments for the spatially inhomogeneous Boltzmann equation for hard spheres (or hard potentials with cutoff), however in a space of the form $W^{3,1}_x L^1(1+|v|^q)$. Another line of research opened by Gamba, Panferov, and Villani [2009] consists in establishing exponential Gaussian pointwise decay by maximum principle arguments (see also Bobylev and Gamba [2017], Alonso, Gamba, and Tasković [2017], and Gamba, Pavlović, and Tasković [2017]). However these works rely on previously establishing exponential integral moments, therefore it is not clear how to use them in this context.

We finally recall that the last part of the research program, the Schauder estimates, is missing for the Boltzmann equation with moderately soft potentials, and is an interesting open question for future researches.

Acknowledgments. This review article greatly benefited from the many discussions of the author with Cyril Imbert and Luis Silvestre, in particular concerning the works on the conditional regularity initiated by Luis Silvestre that form a key part of this review, and concerning their important joint work on the weak Harnack inequality for the Boltzmann equation. Moreover the author gratefully acknowledges the many crucial discussions with Alexis Vasseur about his insights on the De Giorgi method, and with François Golse about averaging lemma and homogeneisation. The author acknowledges partial funding by the ERC grants MATKIT 2011-2016 and MAFRAN 2017-2022.

References

V. I. Agoshkov (1984). “Spaces of functions with differential-difference characteristics and the smoothness of solutions of the transport equation”. *Dokl. Akad. Nauk SSSR* 276.6, pp. 1289–1293. MR: 753365 (cit. on p. 2495).

R. Alexandre, L. Desvillettes, C. Villani, and B. Wennberg (2000a). “Entropy dissipation and long-range interactions”. *Arch. Ration. Mech. Anal.* 152.4, pp. 327–355. MR: 1765272 (cit. on p. 2490).

R. Alexandre, L. Desvillettes, C. Villani, and B. Wennberg (2000b). "Entropy dissipation and long-range interactions". *Arch. Ration. Mech. Anal.* 152.4, pp. 327–355. MR: 1765272 (cit. on p. 2501).

R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang (2008). "Uncertainty principle and kinetic equations". *J. Funct. Anal.* 255.8, pp. 2013–2066. MR: 2462585 (cit. on pp. 2495, 2500).

– (2011a). "The Boltzmann equation without angular cutoff in the whole space: II, Global existence for hard potential". *Anal. Appl. (Singap.)* 9.2, pp. 113–134. MR: 2793203 (cit. on p. 2489).

– (2012). "The Boltzmann equation without angular cutoff in the whole space: I, Global existence for soft potential". *J. Funct. Anal.* 262.3, pp. 915–1010. MR: 2863853 (cit. on p. 2489).

Radjesvarane Alexandre (2009). "A review of Boltzmann equation with singular kernels". *Kinet. Relat. Models* 2.4, pp. 551–646. MR: 2556715 (cit. on p. 2500).

Radjesvarane Alexandre and Mouhamad El Safadi (2005). "Littlewood-Paley theory and regularity issues in Boltzmann homogeneous equations. I. Non-cutoff case and Maxwellian molecules". *Math. Models Methods Appl. Sci.* 15.6, pp. 907–920. MR: 2149928 (cit. on pp. 2495, 2500).

Radjesvarane Alexandre and Mouhamad Elsafadi (2009). "Littlewood-Paley theory and regularity issues in Boltzmann homogeneous equations. II. Non cutoff case and non Maxwellian molecules". *Discrete Contin. Dyn. Syst.* 24.1, pp. 1–11. MR: 2476677 (cit. on pp. 2495, 2500).

Radjesvarane Alexandre, Jie Liao, and Chunjin Lin (2015). "Some a priori estimates for the homogeneous Landau equation with soft potentials". *Kinet. Relat. Models* 8.4, pp. 617–650. MR: 3375485 (cit. on p. 2500).

Radjesvarane Alexandre, Yoshinori Morimoto, Seiji Ukai, Chao-Jiang Xu, and Tong Yang (2010a). "Regularizing effect and local existence for the non-cutoff Boltzmann equation". *Arch. Ration. Mech. Anal.* 198.1, pp. 39–123. MR: 2679369 (cit. on p. 2500).

– (2010b). "Regularizing effect and local existence for the non-cutoff Boltzmann equation". *Arch. Ration. Mech. Anal.* 198.1, pp. 39–123. MR: 2679369 (cit. on p. 2500).

– (2011b). "Global existence and full regularity of the Boltzmann equation without angular cutoff". *Comm. Math. Phys.* 304.2, pp. 513–581. MR: 2795331 (cit. on p. 2500).

Ricardo Alonso, Irene M Gamba, and Maja Tasković (2017). "Exponentially-tailed regularity and time asymptotic for the homogeneous Boltzmann equation". arXiv: 1711.06596 (cit. on p. 2503).

Aleksei A. Arsen'ev and Olga E. Buryak (1990). "On a connection between the solution of the Boltzmann equation and the solution of the Landau-Fokker-Planck equation". *Mat. Sb.* 181.4, pp. 435–446. MR: 1055522 (cit. on pp. 2495, 2496).

Nathalie Ayi (2017). "From Newton's law to the linear Boltzmann equation without cut-off". *Comm. Math. Phys.* 350.3, pp. 1219–1274. MR: 3607474 (cit. on p. 2488).

Claude Bardos, François Golse, and C. David Levermore (1991). "Fluid dynamic limits of kinetic equations. I. Formal derivations". *J. Statist. Phys.* 63.1-2, pp. 323–344. MR: 1115587 (cit. on p. 2488).

– (1993). "Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation". *Comm. Pure Appl. Math.* 46.5, pp. 667–753. MR: 1213991 (cit. on p. 2488).

J. T. Beale, T. Kato, and A. Majda (1984). "Remarks on the breakdown of smooth solutions for the 3-D Euler equations". *Comm. Math. Phys.* 94.1, pp. 61–66. MR: 763762 (cit. on p. 2499).

S. Bernstein (1904). "Sur la nature analytique des solutions des équations aux dérivées partielles du second ordre". *Math. Ann.* 59.1-2, pp. 20–76. MR: 1511259 (cit. on p. 2491).

Jean-Michel Bismut (1981). "Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions". *Z. Wahrsch. Verw. Gebiete* 56.4, pp. 469–505. MR: 621660 (cit. on p. 2493).

A. V. Bobylev (1997). "Moment inequalities for the Boltzmann equation and applications to spatially homogeneous problems". *J. Statist. Phys.* 88.5-6, pp. 1183–1214. MR: 1478067 (cit. on pp. 2502, 2503).

A. V. Bobylev, M. Pulvirenti, and C. Saffirio (2013). "From particle systems to the Landau equation: a consistency result". *Comm. Math. Phys.* 319.3, pp. 683–702. MR: 3040372 (cit. on p. 2489).

Alexander V. Bobylev and Irene M. Gamba (2017). "Upper Maxwellian bounds for the Boltzmann equation with pseudo-Maxwell molecules". *Kinet. Relat. Models* 10.3, pp. 573–585. MR: 3591124 (cit. on p. 2503).

L. Boltzmann (1872). "Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen". *Wiener Berichte, Sitzungsberichte Akademie der Wissenschaften* 66, pp. 275–370 (cit. on p. 2486).

François Bouchut (2002). "Hypoelliptic regularity in kinetic equations". *J. Math. Pures Appl. (9)* 81.11, pp. 1135–1159. MR: 1949176 (cit. on p. 2495).

Marco Bramanti and Luca Brandolini (2007). "Schauder estimates for parabolic nondivergence operators of Hörmander type". *J. Differential Equations* 234.1, pp. 177–245. MR: 2298970 (cit. on p. 2499).

M. Briant (2015a). "Instantaneous filling of the vacuum for the full Boltzmann equation in convex domains". *Arch. Ration. Mech. Anal.* 218.2, pp. 985–1041. MR: 3375544 (cit. on p. 2491).

Marc Briant (2015b). "Instantaneous exponential lower bound for solutions to the Boltzmann equation with Maxwellian diffusion boundary conditions". *Kinet. Relat. Models* 8.2, pp. 281–308. MR: 3356579 (cit. on p. 2491).

Kai-Uwe Bux, Moritz Kassmann, and Tim Schulze (2017). "Quadratic forms and Sobolev spaces of fractional order". arXiv: 1707.09277 (cit. on pp. 2501, 2502).

Luis A. Caffarelli and Xavier Cabré (1995). *Fully nonlinear elliptic equations*. Vol. 43. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, pp. vi+104. MR: 1351007 (cit. on p. 2497).

Luis A. Caffarelli and Luis Silvestre (2009). "Regularity theory for fully nonlinear integro-differential equations". *Comm. Pure Appl. Math.* 62.5, pp. 597–638. MR: 2494809 (cit. on p. 2497).

Stephen Cameron, Luis Silvestre, and Stanley Snelson (2017). "Global a priori estimates for the inhomogeneous landau equation with moderately soft potentials". arXiv: 1701.08215 (cit. on p. 2497).

K. Carrapatoso and S. Mischler (2017). "Landau equation for very soft and Coulomb potentials near Maxwellians". *Ann. PDE* 3.1, Art. 1, 65. MR: 3625186 (cit. on pp. 2489, 2490).

C. Cercignani (1982). "H-theorem and trend to equilibrium in the kinetic theory of gases". *Arch. Mech. (Arch. Mech. Stos.)* 34.3, 231–241 (1983). MR: 715658 (cit. on p. 2489).

Hua Chen, Wei-Xi Li, and Chao-Jiang Xu (2010). "Analytic smoothness effect of solutions for spatially homogeneous Landau equation". *J. Differential Equations* 248.1, pp. 77–94. MR: 2557895 (cit. on p. 2496).

Yemin Chen, Laurent Desvillettes, and Lingbing He (2009). "Smoothing effects for classical solutions of the full Landau equation". *Arch. Ration. Mech. Anal.* 193.1, pp. 21–55. MR: 2506070 (cit. on p. 2496).

Yemin Chen and Lingbing He (2011). "Smoothing estimates for Boltzmann equation with full-range interactions: spatially homogeneous case". *Arch. Ration. Mech. Anal.* 201.2, pp. 501–548. MR: 2820356 (cit. on p. 2500).

– (2012). "Smoothing estimates for Boltzmann equation with full-range interactions: Spatially inhomogeneous case". *Arch. Ration. Mech. Anal.* 203.2, pp. 343–377. MR: 2885564 (cit. on p. 2500).

Peter Constantin and Vlad Vicol (2012). "Nonlinear maximum principles for dissipative linear nonlocal operators and applications". *Geom. Funct. Anal.* 22.5, pp. 1289–1321. MR: 2989434 (cit. on pp. 2497, 2501).

Ennio De Giorgi (1956). "Sull'analiticità delle estremali degli integrali multipli". *Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8)* 20, pp. 438–441. MR: 0082045 (cit. on p. 2491).

– (1957). "Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari". *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)* 3, pp. 25–43. MR: 0093649 (cit. on pp. 2491, 2492).

L. Desvillettes (1993). "Some applications of the method of moments for the homogeneous Boltzmann and Kac equations". *Arch. Rational Mech. Anal.* 123.4, pp. 387–404. MR: 1233644 (cit. on pp. 2502, 2503).

L. Desvillettes and C. Villani (2005). "On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the Boltzmann equation". *Invent. Math.* 159.2, pp. 245–316. MR: 2116276 (cit. on p. 2489).

Laurent Desvillettes (1995). "About the regularizing properties of the non-cut-off Kac equation". *Comm. Math. Phys.* 168.2, pp. 417–440. MR: 1324404 (cit. on p. 2500).

– (2004). "Plasma kinetic models: the Fokker-Planck-Landau equation". In: *Modeling and computational methods for kinetic equations*. Model. Simul. Sci. Eng. Technol. Birkhäuser Boston, Boston, MA, pp. 171–193. MR: 2068584 (cit. on pp. 2495, 2496).

– (2015). "Entropy dissipation estimates for the Landau equation in the Coulomb case and applications". *J. Funct. Anal.* 269.5, pp. 1359–1403. MR: 3369941 (cit. on p. 2496).

Laurent Desvillettes and Clément Mouhot (2009). "Stability and uniqueness for the spatially homogeneous Boltzmann equation with long-range interactions". *Arch. Ration. Mech. Anal.* 193.2, pp. 227–253. MR: 2525118 (cit. on pp. 2497, 2500).

Laurent Desvillettes, Clément Mouhot, and Cédric Villani (2011). "Celebrating Cercignani's conjecture for the Boltzmann equation". *Kinet. Relat. Models* 4.1, pp. 277–294. MR: 2765747 (cit. on p. 2489).

Laurent Desvillettes and Cédric Villani (2000a). "On the spatially homogeneous Landau equation for hard potentials. I. Existence, uniqueness and smoothness". *Comm. Partial Differential Equations* 25.1-2, pp. 179–259. MR: 1737547 (cit. on pp. 2495, 2496).

– (2000b). "On the spatially homogeneous Landau equation for hard potentials. II. H-theorem and applications". *Comm. Partial Differential Equations* 25.1-2, pp. 261–298. MR: 1737548 (cit. on p. 2496).

Laurent Desvillettes and Bernt Wennberg (2004). "Smoothness of the solution of the spatially homogeneous Boltzmann equation without cutoff". *Comm. Partial Differential Equations* 29.1-2, pp. 133–155. MR: 2038147 (cit. on pp. 2495, 2500).

Marco Di Francesco and Sergio Polidoro (2006). "Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form". *Adv. Differential Equations* 11.11, pp. 1261–1320. MR: 2277064 (cit. on pp. 2493, 2499).

Ron J. DiPerna and Pierre-Louis Lions (1989a). "Global weak solutions of Vlasov-Maxwell systems". *Comm. Pure Appl. Math.* 42.6, pp. 729–757. MR: 1003433 (cit. on p. 2495).

– (1989b). "On the Cauchy problem for Boltzmann equations: global existence and weak stability". *Ann. of Math. (2)* 130.2, pp. 321–366. MR: 1014927 (cit. on p. 2489).

T. Elmroth (1983). "Global boundedness of moments of solutions of the Boltzmann equation for forces of infinite range". *Arch. Rational Mech. Anal.* 82.1, pp. 1–12. MR: 684411 (cit. on p. 2502).

Matthieu Felsinger and Moritz Kassmann (2013). "Local regularity for parabolic nonlocal operators". *Comm. Partial Differential Equations* 38.9, pp. 1539–1573. MR: 3169755 (cit. on p. 2501).

Francis Filbet and Clément Mouhot (2011). "Analysis of spectral methods for the homogeneous Boltzmann equation". *Trans. Amer. Math. Soc.* 363.4, pp. 1947–1980. MR: 2746671 (cit. on p. 2491).

Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier (2013). *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, pp. xii+137. MR: 3157048 (cit. on p. 2488).

I. M. Gamba, V. Panferov, and C. Villani (2009). "Upper Maxwellian bounds for the spatially homogeneous Boltzmann equation". *Arch. Ration. Mech. Anal.* 194.1, pp. 253–282. MR: 2533928 (cit. on pp. 2497, 2503).

Irene M Gamba, Nataša Pavlović, and Maja Tasković (2017). "On pointwise exponentially weighted estimates for the Boltzmann equation". arXiv: 1703.06448 (cit. on p. 2503).

François Golse, Cyril Imbert, Clément Mouhot, and A F Vasseur (2017). "Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and application to the Landau equation". 31 pages, 4 figures. (cit. on pp. 2491, 2493–2496, 2499, 2501, 2502).

François Golse, Pierre-Louis Lions, Benoît Perthame, and Rémi Sentis (1988). "Regularity of the moments of the solution of a transport equation". *J. Funct. Anal.* 76.1, pp. 110–125. MR: 923047 (cit. on p. 2495).

François Golse, Benoît Perthame, and Rémi Sentis (1985). "Un résultat de compacité pour les équations de transport et application au calcul de la limite de la valeur propre principale d'un opérateur de transport". *C. R. Acad. Sci. Paris Sér. I Math.* 301.7, pp. 341–344. MR: 808622 (cit. on p. 2495).

François Golse and Laure Saint-Raymond (2004). "The Navier-Stokes limit of the Boltzmann equation for bounded collision kernels". *Invent. Math.* 155.1, pp. 81–161. MR: 2025302 (cit. on p. 2488).

– (2005). "Hydrodynamic limits for the Boltzmann equation". *Riv. Mat. Univ. Parma (7)* 4**, pp. 1–144. MR: 2197021 (cit. on p. 2488).

– (2009). "The incompressible Navier-Stokes limit of the Boltzmann equation for hard cutoff potentials". *J. Math. Pures Appl. (9)* 91.5, pp. 508–552. MR: 2517786 (cit. on p. 2488).

François Golse and Alexis F. Vasseur (2015). "Hölder regularity for hypoelliptic kinetic equations with rough diffusion coefficients". arXiv: 1506.01908 (cit. on p. 2493).

Philip T. Gressman, Joachim Krieger, and Robert M. Strain (2012). "A non-local inequality and global existence". *Adv. Math.* 230.2, pp. 642–648. MR: 2914961 (cit. on p. 2496).

Philip T. Gressman and Robert M. Strain (2011). "Global classical solutions of the Boltzmann equation without angular cut-off". *J. Amer. Math. Soc.* 24.3, pp. 771–847. MR: 2784329 (cit. on pp. 2489, 2500).

Maria P. Gualdani and Nestor Guillen (2016). "Estimates for radial solutions of the homogeneous Landau equation with Coulomb potential". *Anal. PDE* 9.8, pp. 1772–1809. MR: 3599518 (cit. on p. 2496).

Maria P. Gualdani, S. Mischler, and C. Mouhot (2018). "Factorization for non-symmetric operators and exponential H-theorem". *Mémoire de la Société Mathématique de France* (cit. on p. 2503).

Maria P. Gualdani, Stéphane Mischler, and Clément Mouhot (2017). *Factorization for non-symmetric operators and exponential H-theorem*. Mémoire de la Société Mathématique de France. SMF (cit. on pp. 2489, 2490).

Yan Guo (2002). "The Landau equation in a periodic box". *Comm. Math. Phys.* 231.3, pp. 391–434. MR: 1946444 (cit. on pp. 2489, 2500).

M. Hairer (2014). "A theory of regularity structures". *Invent. Math.* 198.2, pp. 269–504. MR: 3274562 (cit. on p. 2499).

Christopher Henderson and Stanley Snelson (2017a). "C^∞ smoothing for weak solutions of the inhomogeneous Landau equation". arXiv: 1707.05710 (cit. on pp. 2491, 2499).

– (2017b). "C^∞ smoothing for weak solutions of the inhomogeneous Landau equation". arXiv: 1707.05710 (cit. on pp. 2498, 2500, 2502).

Christopher Henderson, Stanley Snelson, and Andrei Tarfulea (2017). "Local existence, lower mass bounds, and smoothing for the Landau equation". arXiv: 1712.07111 (cit. on pp. 2491, 2498).

Lars Hörmander (1967). "Hypoelliptic second order differential equations". *Acta Math.* 119, pp. 147–171. MR: 0222474 (cit. on p. 2493).

Zhaohui Huo, Yoshinori Morimoto, Seiji Ukai, and Tong Yang (2008). "Regularity of solutions for spatially homogeneous Boltzmann equation without angular cutoff". *Kinet. Relat. Models* 1.3, pp. 453–489. MR: 2425608 (cit. on pp. 2495, 2500).

E. Ikenberry and C. Truesdell (1956). "On the pressures and the flux of energy in a gas according to Maxwell's kinetic theory. I". *J. Rational Mech. Anal.* 5, pp. 1–54. MR: 0075725 (cit. on p. 2502).

Cyril Imbert and Clément Mouhot (2015). "Hölder continuity of solutions to hypoelliptic equations with bounded measurable coefficients". arXiv: 1505.04608 (cit. on p. 2493).

– (2018). "A toy nonlinear model in kinetic theory". arXiv: 1801.07891 (cit. on pp. 2491, 2498, 2502).

Cyril Imbert, Clément Mouhot, and L. Silvestre (2018). "Decay estimates for large velocities in the Boltzmann equation without cut-off" (cit. on pp. 2500, 2502).

Cyril Imbert and Luis Silvestre (2017). "Weak Harnack inequality for the Boltzmann equation without cut-off". arXiv: 1608.07571 (cit. on pp. 2500, 2501).

Moritz Kassmann (2009). "A priori estimates for integro-differential operators with measurable kernels". *Calc. Var. Partial Differential Equations* 34.1, pp. 1–21. MR: 2448308 (cit. on p. 2501).

Francis Gordon King (1975). *BBGKY hierarchy for positive potentials*. Thesis (Ph.D.)–University of California, Berkeley. ProQuest LLC, Ann Arbor, MI, p. 96. MR: 2625983 (cit. on p. 2488).

A. Kolmogoroff (1934). "Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)". *Ann. of Math. (2)* 35.1, pp. 116–117. MR: 1503147 (cit. on p. 2492).

Joachim Krieger and Robert M. Strain (2012). "Global solutions to a non-local diffusion equation with quadratic non-linearity". *Comm. Partial Differential Equations* 37.4, pp. 647–689. MR: 2901061 (cit. on p. 2496).

N. V. Krylov (1996). *Lectures on elliptic and parabolic equations in Hölder spaces*. Vol. 12. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xii+164. MR: 1406091 (cit. on p. 2499).

N. V. Krylov and M. V. Safonov (1980). "A property of the solutions of parabolic equations with measurable coefficients". *Izv. Akad. Nauk SSSR Ser. Mat.* 44.1, pp. 161–175, 239. MR: 563790 (cit. on pp. 2490, 2492).

S. Kusuoka and D. Stroock (1985). "Applications of the Malliavin calculus. II". *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* 32.1, pp. 1–76. MR: 783181 (cit. on p. 2493).

Shigeo Kusuoka and Daniel Stroock (1984). "Applications of the Malliavin calculus. I". In: *Stochastic analysis (Katata/Kyoto, 1982)*. Vol. 32. North-Holland Math. Library. North-Holland, Amsterdam, pp. 271–306. MR: 780762 (cit. on p. 2493).

L.D. Landau (1936). "Die kinetische Gleichung für den Fall Coulombscher Wechselwirkung". *Phys. Z. Sowjet.* 10. Translation: The transport equation in the case of Coulomb interactions. In D. ter Haar, ed., Collected papers of L. D. Landau, pp.163–170. Pergamon Press, Oxford, 1981., p. 154 (cit. on p. 2488).

Oscar E. Lanford III (1975). "Time evolution of large classical systems". In: *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*. Springer, Berlin, 1–111. Lecture Notes in Phys., Vol. 38. MR: 0479206 (cit. on p. 2488).

Jean Leray (1934). "Sur le mouvement d'un liquide visqueux emplissant l'espace". *Acta Math.* 63.1, pp. 193–248. MR: 1555394 (cit. on p. 2489).

Pierre-Louis Lions (1994). "On Boltzmann and Landau equations". *Philos. Trans. Roy. Soc. London Ser. A* 346.1679, pp. 191–204. MR: 1278244 (cit. on p. 2495).

– (1998). "Régularité et compacité pour des noyaux de collision de Boltzmann sans troncature angulaire". *C. R. Acad. Sci. Paris Sér. I Math.* 326.1, pp. 37–41. MR: 1649477 (cit. on p. 2490).

Shuangqian Liu and Xuan Ma (2014). "Regularizing effects for the classical solutions to the Landau equation in the whole space". *J. Math. Anal. Appl.* 417.1, pp. 123–143. MR: 3191417 (cit. on p. 2496).

Xuguang Lu and Clément Mouhot (2012). "On measure solutions of the Boltzmann equation, part I: moment production and stability estimates". *J. Differential Equations* 252.4, pp. 3305–3363. MR: 2871802 (cit. on p. 2503).

Alessandra Lunardi (1997). "Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbf{R}^n$". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 24.1, pp. 133–164. MR: 1475774 (cit. on p. 2499).

Paul Malliavin (1978). "Stochastic calculus of variation and hypoelliptic operators". In: *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*. Wiley, New York-Chichester-Brisbane, pp. 195–263. MR: 536013 (cit. on p. 2493).

Maria Manfredini (1997). "The Dirichlet problem for a class of ultraparabolic equations". *Adv. Differential Equations* 2.5, pp. 831–866. MR: 1751429 (cit. on p. 2499).

Maria Manfredini and Sergio Polidoro (1998). "Interior regularity for weak solutions of ultraparabolic equations in divergence form with discontinuous coefficients". *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8)* 1.3, pp. 651–675. MR: 1662349 (cit. on p. 2493).

J. C. Maxwell (1867). "On the dynamical theory of gases". *J. Philosophical Transactions of the Royal Society of London* 157, pp. 49–88 (cit. on pp. 2486, 2487).

Stéphane Mischler and Bernt Wennberg (1999). "On the spatially homogeneous Boltzmann equation". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 16.4, pp. 467–501. MR: 1697562 (cit. on p. 2502).

Yoshinori Morimoto, Seiji Ukai, Chao-Jiang Xu, and Tong Yang (2009). "Regularity of solutions to the spatially homogeneous Boltzmann equation without angular cutoff". *Discrete Contin. Dyn. Syst.* 24.1, pp. 187–212. MR: 2476686 (cit. on pp. 2495, 2500).

Charles B. Morrey Jr. (1938). "On the solutions of quasi-linear elliptic partial differential equations". *Trans. Amer. Math. Soc.* 43.1, pp. 126–166. MR: 1501936 (cit. on p. 2491).

Jürgen Moser (1960). "A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations". *Comm. Pure Appl. Math.* 13, pp. 457–468. MR: 0170091 (cit. on pp. 2491, 2492).

Jürgen Moser (1964). "A Harnack inequality for parabolic differential equations". *Comm. Pure Appl. Math.* 17, pp. 101–134. MR: 0159139 (cit. on pp. 2491, 2492).

Clément Mouhot (2005). "Quantitative lower bounds for the full Boltzmann equation. I. Periodic boundary conditions". *Comm. Partial Differential Equations* 30.4-6, pp. 881–917. MR: 2153518 (cit. on p. 2491).

– (2006). "Rate of convergence to equilibrium for the spatially homogeneous Boltzmann equation with hard potentials". *Comm. Math. Phys.* 261.3, pp. 629–672. MR: 2197542 (cit. on p. 2489).

J. Nash (1958). "Continuity of solutions of parabolic and elliptic equations". *Amer. J. Math.* 80, pp. 931–954. MR: 0100158 (cit. on pp. 2491, 2492).

James Norris (1986). "Simplified Malliavin calculus". In: *Séminaire de Probabilités, XX, 1984/85*. Vol. 1204. Lecture Notes in Math. Springer, Berlin, pp. 101–130. MR: 942019 (cit. on p. 2493).

Andrea Pascucci and Sergio Polidoro (2004). "The Moser's iterative method for a class of ultraparabolic equations". *Commun. Contemp. Math.* 6.3, pp. 395–417. MR: 2068847 (cit. on pp. 2493, 2501).

I. G. Petrowsky (1939). "Sur l'analyticité des solutions des systèmes d'équations différentielles". *Rec. Math. N. S. [Mat. Sbornik]* 5(47), pp. 3–70. MR: 0001425 (cit. on p. 2491).

Sergio Polidoro (1994). "On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type". *Matematiche (Catania)* 49.1, 53–105 (1995). MR: 1386366 (cit. on pp. 2493, 2499).

Sergio Polidoro and Maria Alessandra Ragusa (2001). "Hölder regularity for solutions of ultraparabolic equations in divergence form". *Potential Anal.* 14.4, pp. 341–350. MR: 1825690 (cit. on p. 2493).

A. Ja. Povzner (1962). "On the Boltzmann equation in the kinetic theory of gases". *Mat. Sb. (N.S.)* 58 (100), pp. 65–86. MR: 0142362 (cit. on p. 2502).

M. Pulvirenti, C. Saffirio, and S. Simonella (2014). "On the validity of the Boltzmann equation for short range potentials". *Rev. Math. Phys.* 26.2, pp. 1450001, 64. MR: 3190204 (cit. on p. 2488).

E. V. Radkevich (2008). "Equations with nonnegative characteristic form. II". *Sovrem. Mat. Prilozh.* 56, Differentsial'nye Uravneniya s Chastnymi Proizvodnymi, pp. 3–147. MR: 2675371 (cit. on p. 2499).

J. Schauder (1934). "Über lineare elliptische Differentialgleichungen zweiter Ordnung". *Math. Z.* 38.1, pp. 257–282. MR: 1545448 (cit. on pp. 2491, 2498).

Luis Silvestre (2016a). "A new regularization mechanism for the Boltzmann equation without cut-off". *Communications in Mathematical Physics* 348.1, pp. 69–100 (cit. on pp. 2496, 2500–2502).

– (2016b). "Upper bounds for parabolic equations and the Landau equation". arXiv: 1511.03248 (cit. on p. 2496).

– (2017). "Upper bounds for parabolic equations and the Landau equation". *J. Differential Equations* 262.3, pp. 3034–3055. MR: 3582250 (cit. on pp. 2496, 2497).

Leon Simon (1997). "Schauder estimates by scaling". *Calc. Var. Partial Differential Equations* 5.5, pp. 391–407. MR: 1459795 (cit. on p. 2499).

C. Truesdell and R. G. Muncaster (1980). *Fundamentals of Maxwell's kinetic theory of a simple monatomic gas*. Vol. 83. Pure and Applied Mathematics. Treated as a branch of rational mechanics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, pp. xxvii+593. MR: 554086 (cit. on p. 2489).

Cédric Villani (1998). "On the spatially homogeneous Landau equation for Maxwellian molecules". *Math. Models Methods Appl. Sci.* 8.6, pp. 957–983. MR: 1646502 (cit. on pp. 2495, 2496).

– (1999). "Regularity estimates via the entropy dissipation for the spatially homogeneous Boltzmann equation without cut-off". *Rev. Mat. Iberoamericana* 15.2, pp. 335–352. MR: 1715411 (cit. on p. 2490).

Cédric Villani (2002). "A review of mathematical topics in collisional kinetic theory". *Handbook of mathematical fluid dynamics* 1.71-305, pp. 3–8 (cit. on p. 2503).

Cédric Villani (2003). "Cours Peccot at the Collège de France". Unpublished (cit. on p. 2493).

– (2009). "Hypocoercivity". *Mem. Amer. Math. Soc.* 202.950, pp. iv+141. MR: 2562709 (cit. on p. 2499).

Wendong Wang and Liqun Zhang (2009). "The C^α regularity of a class of non-homogeneous ultraparabolic equations". *Sci. China Ser. A* 52.8, pp. 1589–1606. MR: 2530175 (cit. on p. 2493).

– (2011). "The C^α regularity of weak solutions of ultraparabolic equations". *Discrete Contin. Dyn. Syst.* 29.3, pp. 1261–1275. MR: 2773175 (cit. on p. 2493).

Bernt Wennberg (1996). "The Povzner inequality and moments in the Boltzmann equation". In: *Proceedings of the VIII International Conference on Waves and Stability in Continuous Media, Part II (Palermo, 1995)*. 45, part II, pp. 673–681. MR: 1461113 (cit. on p. 2502).

Kung-Chien Wu (2014). "Global in time estimates for the spatially homogeneous Landau equation with soft potentials". *J. Funct. Anal.* 266.5, pp. 3134–3155. MR: 3158719 (cit. on pp. 2497, 2500).

Liqun Zhang (2011). "The C^α regularity of a class of ultraparabolic equations". *Commun. Contemp. Math.* 13.3, pp. 375–387. MR: 2813495 (cit. on p. 2493).

Received 2018-15-01.

Clément Mouhot
University of Cambridge
DPMMS, Centre for Mathematical Sciences
Wilberforce road, Cambridge CB3 0WA, UK
C.Mouhot@dpmms.cam.ac.uk

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2513–2536)

RESONANCES IN HYPERBOLIC DYNAMICS

STÉPHANE NONNENMACHER

Abstract

The study of wave propagation outside bounded obstacles uncovers the existence of resonances for the Laplace operator, which are complex-valued generalized eigenvalues, relevant to estimate the long time asymptotics of the wave. In order to understand distribution of these resonances at high frequency, we employ semiclassical tools, which leads to considering the classical scattering problem, and in particular the set of trapped trajectories. We focus on "chaotic" situations, where this set is a hyperbolic repeller, generally with a fractal geometry. In this context, we derive fractal Weyl upper bounds for the resonance counting; we also obtain dynamical criteria ensuring the presence of a resonance gap. We also address situations where the trapped set is a normally hyperbolic submanifold, a case which can help analyzing the long time properties of (classical) Anosov contact flows through semiclassical methods.

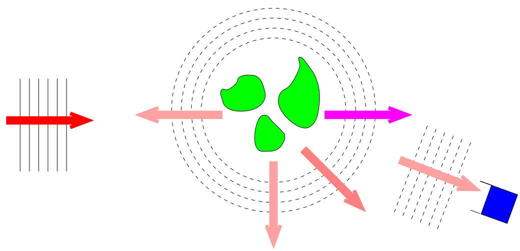
1 Introduction

Spectral geometry attemps to understand the connection between the *shape* (geometry) of a smooth Riemannian manifold (M, g), and the *spectrum* of the positive Laplace-Beltrami operator $-\Delta$ on this manifold. When M is compact, the spectrum is made of discrete eigenvalues of finite multiplicities $(\lambda_k^2)_{k\geq 0}$, associated with an orthonormal basis of smooth eigenfunctions $(\phi_k)_{k\geq 0}$. What is the role of this spectrum? It allows to explicitly describe the time evolution of the waves waves, e.g. evolved through the wave equation $(\partial_{tt}^2 - \Delta)u = 0$. The connection comes as follows: taking as any initial datum $u(0) = 0$, $\partial_t u(0) = u_0 \in L^2(M)$, the wave at any time $t \geq 0$ is given by the exact expansion

$$(1)\quad u(t,x) = \Big(\frac{\sin(t\sqrt{-\Delta})}{\sqrt{-\Delta}}u_0\Big)(x) = \sum_{k\geq 1}\langle \phi_k, u_0\rangle\, \phi_k(x)\,\frac{\sin(t\lambda_k)}{\lambda_k}, \quad x \in M, t \geq 0\,.$$

Hence, any information on the eigenvalues and eigenfunctions allows to better characterize the evolved wave $u(t)$.

Figure 1: Scattering of a wave by obstacles $\Omega = \cup_i \Omega_i \subset \mathbb{R}^d$. Parallel lines indicate incoming and outgoing wave trains (arrows indicate the direction of propagation). The blue box indicates a "detector".



1.1 Scattering. In many physical experiments, the waves (or wavefunctions) are not confined to compact domains, but can spread towards spatial infinity. The ambient manifold (M, g) therefore has infinite volume, and in general its geometry towards infinity is "simple". For instance, a physically relevant situation consists of the case where $M = \mathbb{R}^d \setminus \Omega$, with Ω an open bounded subset of $\mathbb{R}^d$, representing a bounded "obstacle" (or a set of several obstacles). These obstacles will *scatter* an incoming flux of waves arriving from a certain direction at infinity, resulting in a flux of outgoing waves propagating towards infinity along all possible directions (see Figure 1). In actual experiments, the experimentalist can produce incoming waves with definite frequency and direction, and can detect the outgoing waves, along one or several directions. Such an experiment aims at reconstructing the shape of the obstacle, from the analysis of the outgoing waves.

1.2 Resonances. Our objective will not be this ambitious *inverse problem*, but we will try characterize quantitatively this scattering phenomenon, assuming some geometric and dynamical properties of the obstacles. This will imply a spectral study of the Laplacian $-\Delta$ on M (say, with Dirichlet boundary conditions on $\partial\Omega$). Due to the infinite volume of M, the spectrum of $-\Delta$ is purely continuous on $\mathbb{R}_+$ with no embedded eigenvalues. However, one can exhibit a form of discrete expansion resembling (1) by uncovering *resonances* (see e.g. the incoming book Dyatlov and Zworski [2018] on scattering and resonances, or the recent comprehensive review Zworski [2017]).

Let us assume that the initial datum $u_0 \in C_c^\infty(M)$; its time evolution can be expressed through Stone's formula:

$$u(t,x) = \frac{1}{2i\pi} \int_{\mathbb{R}} d\lambda \, e^{-it\lambda} \, R(\lambda) \, u_0 \,, \tag{2}$$

MSC2010: primary 81Q50; secondary 81Q20, 35P25, 37D20, 37D40, 81Q12.

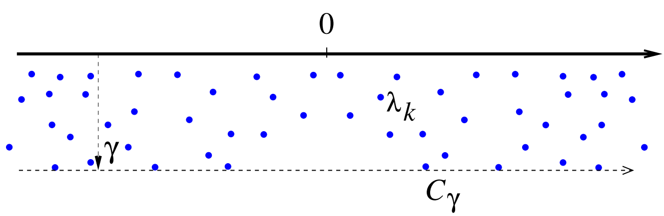
where $R(\lambda)$ is the resolvent operator $(-\Delta - \lambda^2)^{-1}$, first defined in the upper half-plane $\mathrm{Im}\lambda > 0$, and then continued down to $\lambda \in \mathbb{R}$ as an operator $L^2_{comp} \to L^2_{loc}$. $R(\lambda)$ actually admits a meromorphic extension from $\mathrm{Im}\lambda > 0$ to the full lower half-plane $\mathbb{C}_- = \{\mathrm{Im}\lambda < 0\}$ (with a logarithmic singularity at $\lambda = 0$ in even dimensions d), with the possibility of discrete poles $\{\lambda_k \in \mathbb{C}_-\}$ of finite multiplicities, called the resonances of the system.

This meromorphic extension encourages us to deform the contour of the above integral towards a line $C_\gamma = -i\gamma + \mathbb{R}$, thereby collecting the contributions of the residues at the λ_k. Assuming that all resonances have multiplicity 1, we obtain the expansion

$$(3) \qquad u(t) = \sum_{\mathrm{Im}\lambda_k \geq -\gamma} e^{-it\lambda_k}\, \Pi_{\lambda_k} u_0 + I(t, C_\gamma)\,, \quad \text{where } \Pi_{\lambda_k} = \frac{1}{2i\pi} \oint_{\lambda_k} R(\lambda)\, d\lambda,$$

and $I(t, C_\gamma)$ is the integral in (2) taken along the contour C_γ.

Figure 2: Contour deformation uncovering resonances of $-\Delta$.



Resonances come in symmetric pairs $\lambda_k \leftrightarrow -\bar{\lambda}_k$ (see Figure 2). Each λ_k corresponds to a resonant state $u_k \in C_c^\infty(M)$, which satisfies the equation $-\Delta u_k = \lambda_k u_k$ and behaves as $\sim e^{i\lambda_k |x|}$ when $|x| \to \infty$, so it diverges exponentially, showing that $u_k \notin L^2(M)$. If $\mathrm{Re}\lambda_k > 0$ the state u_k is said to be *purely outgoing*; the complex conjugate function $\bar{u}_k(x)$ corresponds to the dual resonance $-\bar{\lambda}_k$ of negative real part: it is purely incoming. The resonant state u_k allows to express the "spectral projector" Π_{λ_k} (which acts $L^2_{comp} \to L^2_{loc}$) as $\Pi_{\lambda_k} u_0 = \langle \bar{u}_k, u_0 \rangle u_k$ (the bracket $\langle \bar{u}_k, u_0 \rangle = \int dx\, u_k(x)\, u_0(x)$ makes sense since u_0 has compact support).

Assuming we control the size of the remainder term (the contour integral $I(t, C_\gamma)$), the expansion (3) provides informations on the shape and intensity of the wave $u(t, x)$, particularly in the asymptotic $t \gg 1$: it can explain at which rate the wave leaks (*disperses*) out of a given bounded region (say, a large ball $B(R)$), by providing some quantitative bounds on $u(t) \restriction_{B(R)}$. To control the remainder $I(t, C_\gamma)$, one needs to control the size of the truncated resolvent operator $\mathbb{1}_{B(R)} R(\lambda) \mathbb{1}_{B(R)}$ for $\lambda \in C_\gamma$, in particular the contour should avoid hitting resonances, which requires to control the location of the resonance cloud in the vicinity of C_γ.

1.3 Semiclassical regime. These arguments hint at our main objective: to determine, as precisely as possible, the distribution of the resonances $\{\lambda_k\}$, and possibly also obtain bounds on the meromorphically continued resolvent $R(\lambda)$. We will be mostly interested in the *high frequency regime* $|\mathrm{Re}\lambda| \gg 1$, which we choose to rephrase as a *semiclassical regime* with small parameter $h \ll 1$. To avoid having to deal with both signs of $\mathrm{Re}\lambda$, we replace the wave equation by the half-wave equation, written in this semiclassical setting as:

$$(4)\qquad ih\partial_t u(t) = P_h u(t), \quad \text{with the semiclassical operator } P_h = \sqrt{-h^2\Delta}\,.$$

The small parameter $0 < h \ll 1$ is usually called "Planck's constant", since the above equation has the form of a semiclassical Schrödinger equation (see below). Here h is just a bookkeeping parameter: we will study the resonances $z_k = z_k(h) \stackrel{\mathrm{def}}{=} h\lambda_k$ of the operator P_h near some fixed energy $E > 0$ (typically $E = 1$ for the above half-wave equation), indicating that $\mathrm{Re}\lambda_k \sim h^{-1}$.

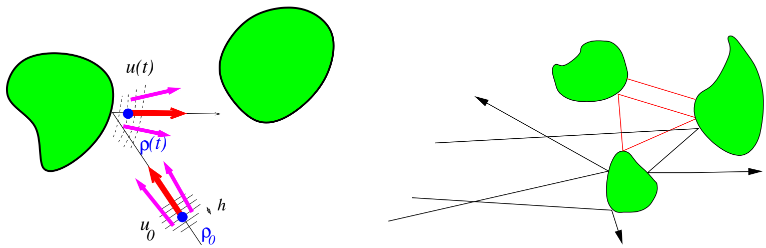
We will use the same notations when considering the "true" semiclassical Schrödinger equation, describing the evolution of a quantum particle on M, subject to an electric potential $V(x)$:

$$(5)\qquad ih\partial_t u(t) = P_h u(t), \quad P_h = -h^2\Delta + V(x)\,, \quad V \in C_c^\infty(M,\mathbb{R})\,.$$

The Schrödinger operator P_h also admits resonances $z_k(h)$ in the lower half-plane, obtained as the poles of the resolvent $(P_h - z)^{-1}$, meromorphically extended from $\{\mathrm{Re}z > 0, \mathrm{Im}z > 0\}$ to $\{\mathrm{Im}z < 0\}$; now the $z_k(h)$ depend nontrivially of h. In these semiclassical notations, the time evolution operator now reads $e^{-itP_h/h}$, so each term $\langle \bar{u}_k, u_0\rangle u_k$ in (3) will evolve at a rate $e^{-itz_k/h}$, hence decay at a rate $e^{t\mathrm{Im}z_k/h}$. The deeper the resonance ($\equiv$ the larger $|\mathrm{Im}z_k|$), the faster this term will decay. We call $\tau_k(h) \stackrel{\mathrm{def}}{=} \frac{h}{|\mathrm{Im}z_k|}$ the lifetime of the resonance. As we will see below, we will be mostly interested in resonances with lifetimes bounded from below, $\tau_k \geq c > 0$, which corresponds to studying the resonances in strips of width $\{\mathrm{Im}z_k = \mathcal{O}(h)\}$.

1.3.1 Semiclassical evolution of wavepackets. This semiclassical regime allows us to use the powerful machinery of semiclassical/microlocal analysis Zworski [2012], which relates the Schrödinger evolution (4) with the evolution of classical particles through the Hamiltonian flow φ_p^t on the phase space $T^*M \ni (x,\xi)$. This flow is generated by the classical Hamiltonian $p(x,\xi)$, given by the principal symbol of the operator P_h (in the

Figure 3: Left: a wavepacket of wavelength h is scattered by an obstacle. Right: scattering of classical trajectories (light rays following broken geodesics).



above examples $p(x, \xi) = |\xi|$, respectively $p(x, \xi) = |\xi|^2 + V(x)$). To illustrate this connection, we represent on the left of Figure 3 the propagation of a minimum-uncertainty wavepacket $u_0(x)$ through the half-wave equation on $M = \mathbb{R}^d \setminus \Omega$. The wavepacket can be chosen for instance as a minimum-uncertainty Gaussian wavepacket, also called a *coherent state*

$$u_0(x) = C_h \, e^{-\frac{|x-x_0|^2}{2h}} \, e^{i\xi_0 \cdot x/h}.$$

This wavepacket is essentially localized in an $h^{1/2}$-neighbourhood of the point x_0, while its semiclassical Fourier transform $\mathfrak{F}_h u_0(\xi)$ is localized in an $h^{1/2}$-neighbourhood of the momentum ξ_0 (materialized by the red and pink arrows in the Figure); we say that this state is *microlocalized* (or centered) on the phase space point $\rho_0 = (x_0, \xi_0)$. Heisenberg's uncertainty principle shows that the concentration of such a wavepacket is maximal, equivalently the "uncertainty" in its position and momentum is minimal. For a given time window $t \in [0, T]$, in the semicassical limit the evolved state $u(t) = e^{-itP_h/h}u_0$ will remain a microscopic wavepacket, centered at the point $\rho(t) = \varphi^t(\rho_0)$, where φ^t is the broken geodesic flow shown on the figure. If we replace the hard obstacles by a smooth potential, the geodesic flow will be replaced by the Hamiltonian flow φ_p^t.

1.3.2 Introducing the trapped set. In order to analyze the quantum scattering and its associated resonances, it will be crucial to understand the corresponding classical dynamical system, that is the scattering of classical particles induced by obstacles, potentials or metric perturbations, as sketched on the right of Figure 3. In particular, the distribution of resonances will depend on the dynamics of the trajectories remaining in a bounded region of phase space for very long times. For a given energy value $E > 0$, we thus introduce the set of points which are trapped forever in the past (resp. in the future, resp. in both time directions):

(6) $$\Gamma_E^{\pm} \stackrel{\text{def}}{=} \{\rho \in p^{-1}(E),\ \varphi_p^t(\rho) \not\to \infty,\ t \to \mp\infty\}, \quad K_E = \Gamma_E^+ \cap \Gamma_E^- .$$

Our assumptions on the structure of M near infinity will always imply that the *trapped set* K_E is a compact subset of the energy shell $p^{-1}(E)$; this set is invariant through the flow φ_p^t. The distribution of the resonances in the semiclassical limit will be impacted by the dynamics of the flow φ_p^t on K_E. The punchline of the present notes could be:

> *In the semiclassical regime, the distribution of the resonance $\{z_k(h)\}$ near the energy E strongly depends on the structure of the trapped set K_E, and of the dynamical properties of the flow φ_p^t near K_E.*

1.4 Hyperbolicity. In these notes dedicated to "quantum chaos", we will mostly focus on systems for which the flow $\varphi_p^t \restriction_{K_E}$ is hyperbolic (Section 5 will contain examples of partial hyperbolicity). What does hyperbolicity mean? It describes the rate at which nearby trajectories depart from each other: for a hyperbolic flow, they separate at an exponential rate, either in the past direction, or in the future, or (most commonly) in both time directions. The trajectories are therefore unstable w.r.t. perturbations of the initial conditions. More precisely, an orbit $\mathcal{O}(\rho_0) = (\varphi^t(\rho_0))_{t\in\mathbb{R}} \subset p^{-1}(E)$ is hyperbolic if and only if, at each point $\rho \in \mathcal{O}(\rho_0)$, the $2d-1$-dimensional tangent space $T_\rho p^{-1}(E)$ splits into three subspaces,

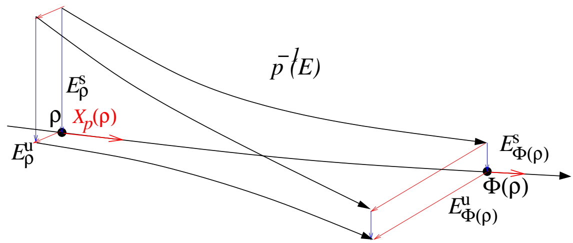
$$T_\rho p^{-1}(E) = \mathbb{R}X_p(\rho) \oplus E^u(\rho) \oplus E^s(\rho),$$

where $X_p(\rho)$ is the Hamiltonian vector field generating the flow, $E^s(\rho)$ (resp. $E^u(\rho)$) is the stable (resp. unstable) subspace at ρ, characterized by the following contraction properties in the future, resp. in the past:

(7) $$\exists C, \mu > 0, \quad \forall t \geq 0, \quad \|d\varphi_p^t \restriction_{E^s(\rho)}\| \leq C\, e^{-\mu t}, \quad \|d\varphi_p^{-t} \restriction_{E^u(\rho)}\| \leq C\, e^{-\mu t} .$$

The trapped set K_E is said to be (uniformly) hyperbolic if each orbit $\mathcal{O}(\rho) \subset K_E$ is hyperbolic, with the coefficients C, μ being uniform w.r.t. $\rho \in K_E$. In general the unstable subspaces E_ρ^u are only Hölder-continuous w.r.t. $\rho \in K_E$, even if the flow φ^t is smooth; this poor regularity jumps to a smooth (actually, real analytic) dependence in the setting of hyperbolic surfaces described in the next section. Such a uniformly hyperbolic flow $\varphi_p^t \restriction_{K_E}$ satisfies Smale's Axiom A; its long time dynamical properties have been studied since the 1960s, using the tools of symbolic dynamics and the thermodynamical formalism Bowen and Ruelle [1975]. Below we will use some "thermodynamical" quantities associated to the flow, namely the topological entropy and pressures. The Anosov flows we will mention in the last section are particular examples of such Axiom A flows.

Figure 4: Hyperbolicity of the orbit $\mathcal{O}(\rho)$, with the stable an unstable subspaces transverse to the vector $X_p(\rho)$.



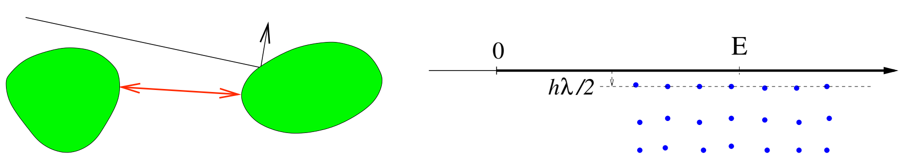
2 Examples of hyperbolic flows

2.1 A single hyperbolic periodic orbit. The simplest example of hyperbolic set occurs in the scattering by the union of two disjoint strictly convex obstacles in $\mathbb{R}^d$: in that case the trapped set is made of a single orbit bouncing periodically between the two obstacles (see Figure 5). For this simple situation, the resonances of $P_h = -h^2\Delta$ can be computed very precisely in the semiclassical limit Ikawa [1983] and Gérard and Sjöstrand [1987]; in dimension $d = 2$, in a small neighbourhood of the classical energy $E = 1$, they asymptotically form a half-lattice:

(8)
$$z_{\ell,k}(h) = E(h) + \frac{2\pi hk}{T} - ih\lambda(1/2+\ell) + \mathcal{O}(h^2), \quad \ell \in \mathbb{N},\ k \in \mathbb{Z},\ E(h) = 1 + \mathcal{O}(h)\,.$$

Here T is the period of the bouncing orbit, while $\lambda > 0$ is the rate of unstability along

Figure 5: Left: the simplest case of hyperbolic set: scattering between two strictly convex obstacles. Right: semiclassical resonances for this system



the orbit, meaning that $\|d\varphi_p^T \restriction_{E^u(\rho)}\| = e^{\lambda T}$. Obtaining such explicit formulas for the resonances is specific to this very simple situation, but it already presents two interesting features. First, the number of resonances in any rectangle $R(E, Ch, \gamma h)$ of the type (11) is uniformly bounded when $h \to 0$, and it is nonzero if γ and C are large enough. Second, if $\gamma < \lambda/2$ (and if h is small enough), the box $R(E, Ch, \gamma h)$ will be empty of resonances: this is the first instance of a *resonance gap* connected with the hyperbolicity of the flow on the trapped set.

2.2 Fully developed chaos: fractal hyperbolic trapped set. Beside hyperbolicity, the second ingredient of "chaos" is the complexity of the flow, which can be characterized by a positive topological entropy, indicating an exponential proliferation of long periodic orbits:

$$H_{top}(\varphi^t \restriction_{K_1}) = \lim_{T\to\infty} \frac{1}{T} \log \#\{\gamma \in \mathrm{Per}(K_1),\ T \le T_\gamma \le T+1\}\,, \tag{9}$$

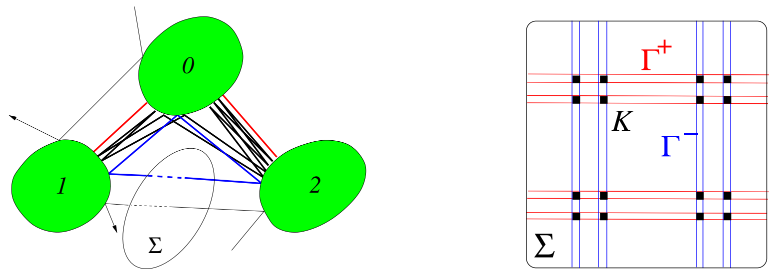
where $\mathrm{Per}(K_1)$ denotes the set of periodic orbits in K_1, and T_γ is the period of the orbit γ. A simple example of system featuring such a chaotic trapped set is obtained by adding one convex obstacle to the 2-obstacle example of the previous paragraph. Provided this third obstacle is well-placed with respect to the other two (so that the three obstacles satisfy a "no-eclipse condition", like in Figure 6, left), the trapped trajectories at energy $E = 1$ form a hyperbolic set K_1, which contains a countable number of periodic orbits, and uncountably many nonperiodic ones. A way to account for this complexity is to construct a symbolic representation of the orbits. Label each obstacle by a number $\alpha \in \{0, 1, 2\}$; then to each bi-infinite word $\cdots\alpha_{-1}\alpha_0\alpha_1\alpha_2\cdots$ such that $\alpha_i \neq \alpha_{i+1}$, corresponds a unique trapped orbit in K_1, which hits the obstacles sequentially in the order indicated by the word. Periodic words correspond to periodic orbits, nonperiodic words to nonperiodic orbits. This correspondence between words and orbits allows to quantitatively estimate the complexity of the flow on K_1. In turn, the strict convexity of the obstacles ensures that all trapped orbits are hyperbolic, the instability arising at the bounces.

The trapped set K_1 has a fractal geometry, which can be described by some fractal dimension. It is foliated by the trajectories (which accounts for one "smooth" dimension), so its fractal nature occurs in the transverse directions to the flow, visible in its intersection with a Poincaré section $\Sigma \subset S^*X$ (see Figure 6). This intersection $K_E \cap \Sigma$ (represented by the union of black squares) has the structure of a horseshoe; as the intersection of stable (Γ^-) and unstable (Γ^+) manifolds, it locally has a "product structure".

In space dimension $d = 2$, the dimension of K_1 can be expressed by using a *topological pressure*. This pressure, a "thermodynamical" quantity of the flow, is defined in terms of the *unstable Jacobian* of the flow, $J^u_t(\rho) = |\det(d\varphi^t \restriction_{E^u(\rho)})|$. For a periodic orbit γ of period T_γ, we denote $J^u(\gamma) = J^u_{T_\gamma}(\rho_\gamma)$, where ρ_γ is any point in γ. Now, for any $s \in \mathbb{R}$, we may define the pressure as

$$\mathcal{P}(s) = \mathcal{P}(s, \varphi^t \restriction_{K_1}) = \lim_{T\to\infty} \frac{1}{T} \log \sum_{T \le T_\gamma \le T+1} J^u(\gamma)^{-s}\,, \tag{10}$$

Figure 6: Left: three convex obstacles on $\mathbb{R}^2$, leading to a fractal hyperbolic repeller. Right: intersection of K_E with a Poincaré section Σ.



where the sum runs over all periodic orbits $\gamma \in \mathrm{Per}(K_1)$ of periods in the interval $[T, T+1]$. $\mathcal{P}(0)$ is equal to the topological entropy (9), which is positive. When increasing s, the factors $J^u(\gamma)^{-s}$ decay exponentially when $T \to \infty$, hence the hyperbolicity embodied by these factors balances the complexity characterized by the large number of orbits. The pressure $\mathcal{P}(s)$ is smooth and strictly decreasing with s, and one can show that $\mathcal{P}(1) < 0$; hence, it vanishes at a single value $\delta \in (0, 1)$. In the 2-dimensional setting (for which $E^{u/s}(\rho)$ are 1-dimensional), the Hausdorff dimension of K_1 is given by Bowen's formula:

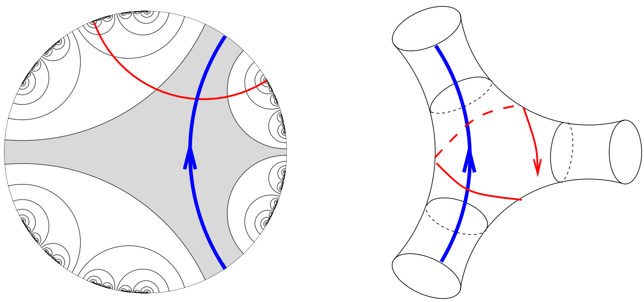
$$\dim K_1 = 1 + 2\delta \iff \mathcal{P}(\delta) = 0 .$$

The topological pressure will pop up again when studying resonance gaps, see Theorem 2.

2.3 An interesting class of examples: hyperbolic surfaces of infinite area. We have mentioned above that one way to "scatter" a wave, or a classical particle, was to modify the metric on M in some compact neighbourhood. Because we are interested in hyperbolic dynamics, an obvious way to generate hyperbolicity is to consider metrics g of negative sectional curvature (giving M locally the surface the aspect of a "saddle"). Such a metric automatically induces the hyperbolicity of the orbits, the instability rate being proportional to the square-root of the curvature.

Such surfaces can be constructed Borthwick [2016] by starting from the Poincaré hyperbolic disk $\mathbb{D} = \{z \in \mathbb{C},\ |z| < 1\}$, equipped with the metric $g = 4\frac{dz\,d\bar{z}}{(1-z\bar{z})^2}$: the curvature is then equal to -1 everywhere. The Lie group $SL(2,\mathbb{R})$ acts on this disk isometrically. By choosing a discrete subgroup $\Gamma < SL(2,\mathbb{R})$ of the Schottky type, the quotient $M = \Gamma\backslash\mathbb{D}$ is a smooth surface of infinite volume, without cusps. On the left of Figure 6 we represent the Poincaré disk, tiled by fundamental domains of such a Schottky subgroup Γ (the grey area is one fundamental domain), the boundaries of the domains being given by geodesics

Figure 7: Construction of a hyperbolic surface $M = \Gamma\backslash\mathbb{C}$ of infinite volume. Left: fundamental domains of the action of Γ on $\mathbb{D}$. Right: representation of M.



on $\mathbb{D}$ (which corresond to Euclidean circles hitting $\partial\mathbb{D}$ orthogonally). On the figure we also notice the accumulation of small circles towards a subset $\Lambda_\Gamma \subset \partial\mathbb{D}$, called the limit set of the group Γ. This limit set is a fractal set of dimension $\delta = \delta_\Gamma \in (0,1)$.

On the right of the figure we plot the quotient surface $M = \Gamma \setminus \mathbb{D}$, composed of a compact part (the "core") and of three "hyperbolic funnels" leading to infinity. The trapped geodesics of M are fully contained in the compact core, they can be represented by geodesics on $\mathbb{D}$ connecting two points of Λ_Γ (red geodesic on the figure). On the opposite, geodesics on $\mathbb{D}$ crossing $\partial\mathbb{D} \setminus \Lambda_\Gamma$ correspond to transient geodesics on M (blue geodesic on the figure) which start and end in a funnel. The trapped set can therefore be identified as $K_1 \equiv \Lambda_\Gamma \times \Lambda_\Gamma \times \mathbb{R}$, and its Hausdorff dimension $\dim K_1 = 1 + 2\delta$.

The Laplace-Beltrami operator $-\Delta_M$ has a continuous spectrum on $[1/4, \infty)$, which is usually represented by the values $s(1-s)$, for a spectral parameter $s \in \frac{1}{2} + i\mathbb{R}$. The resolvent operator $R(s) = (-\Delta_M - s(1-s))^{-1}$ can be meromorphically extended from $\{\mathrm{Re} s > 1/2\}$ to $\{\mathrm{Re} s < 1/2\}$. The resonances are given by a discrete set $\{s_k\}$ in the half-space $\{\mathrm{Re} s < 1/2\}$. A huge advantage of this model, is that these resonances are given by the zeros of the Selberg zeta function

$$Z_\Gamma(s) \stackrel{\text{def}}{=} \prod_{\gamma\in\mathrm{Per}^*} \prod_{m=0}^{\infty} (1 - e^{-(s+m)|\gamma|}),$$

where Per^* denotes the set of primitive periodic geodesics on M. This exact connection between *geometric data* (lengths of the periodic geodesics) and *spectral data* (resonances of $-\Delta$) is specific to the case of surfaces of constant curvature. Another particular feature of the constant curvature is the fact that the stable/unstable directions $E^{s/u}(\rho)$ can be defined at any point $\rho \in M$, and depend smoothly on the base point ρ.

The identification of the resonances with the zeros of $Z_\Gamma(s)$ provides powerful techniques to study their distribution, with purely "classical" techniques, without any use of PDE methods. These zeros can be obtained by studying a 1-dimensional map on the circle, called the Bowen-Series map, constructed from the generators of the group Γ Naud [2005]. This map induces a family of transfer operators $\mathfrak{L}_s$ indexed by the spectral parameter; one shows that these operators, when acting on appropriate spaces of analytic functions, are nuclear (in the sense of Grothendieck), and that the Selberg zeta function can be obtained as the Fredholm determinant $Z_\Gamma(s) = \det(1-\mathfrak{L}_s)$. The spectral study of the classical transfer operators $\mathfrak{L}_s$ can therefore deliver informations on the resonance spectrum, which are often more precise than what is achievable through PDE techniques.

3 Fractal Weyl upper bounds

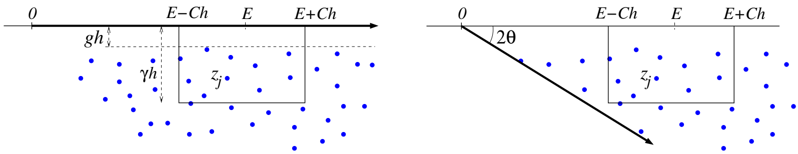
3.1 Counting long living resonances. We are interested in the distribution of the resonances (λ_j) (for $-\Delta$) or $(z_k(h))$ (for P_h) in the lower half-plane. Because we want to use these resonances in dynamics estimates as in (3), we will focus on the *long living resonances*, such that $\mathrm{Im} z_k(h) \geq -\gamma h$ for some fixed $\gamma > 0$, or equivalently such that the corresponding lifetimes $\tau_k(h) \geq 1/\gamma > 0$. We will also focus on resonances such that $\mathrm{Re} z_k$ lies in some small energy window $[E-\epsilon, E+\epsilon]$: this will allow us to connect their distribution with the properties of the classical flow at energy E. Figure 8 sketches the more precise spectral region we will study, centered at $E > 0$: we will count the resonances in rectangles of the type

$$(11) \qquad R(E, Ch, \gamma h) = [E - Ch, E + Ch] - i[0, \gamma h], \quad C, \gamma > 0 \text{ independent of } h.$$

In the present section, our main result is a *fractal Weyl upper bound* (see Theorem 1) for the number of resonances in those rectangles. In the next section we will be especially interested in situations for which such a rectangle contains *no* resonance, like in the rectangle $R(E, Ch, gh)$ of Figure 8: we will then speak of a (semiclassical) *resonance gap* near the energy E.

3.2 Complex deformation of P_h: turning resonances into eigenvalues. For simplicity we consider manifolds M which, outside some big ball $B(R_0/2)$, is equal to the Euclidean space $\mathbb{R}^d \setminus B(R_0/2)$. To analyze the resonances of P_h in $R(E, Ch, \gamma h)$, a convenient method consists in twisting the selfadjoint operator P_h into a nonselfadjoint operator $P_{h,\theta}$, through a "complex deformation" procedure Aguilar and Combes [1971]. Outside a large ball $B(R_0)$, the differential operator $P_{h,\theta}$ is equal to $-h^2 e^{-2i\theta}\Delta$, while it is equal to the original P_h inside $B(R_0/2)$. In our applications the angle parameter $\theta \in (0, \pi/4)$ will be assumed small. Through the twisting $P_h \to P_{h,\theta}$, the continuous spectrum has

Figure 8: Resonances of a semiclassical operator P_h in the rectangle $R(E, Ch, \gamma h)$. Right: spectrum of the twisted operator $P_{h,\theta}$.



been tilted from $\mathbb{R}_+$ to $e^{-2i\theta}\mathbb{R}_+$, and by doing so has *uncovered* the resonances $z_j(h)$ contained in this corresponding sector: these resonances have been turned into eigenvalues, with eigenfunctions $\tilde{u}_j \in L^2$. For $h > 0$ small enough, the rectangle $R(E, Ch, \gamma h)$ will be contained in the $e^{-2i\theta}$ sector, so we are lead to analyze the (discrete) L^2 spectrum of the nonselfadjoint semiclassical operator $P_{h,\theta}$ inside this rectangle.

Let us analyze the twisted Schrödinger evolution. We have seen in Section 1.3 that a wavepacket u_{ρ_0} centered at a phase space point ρ_0 is transported by the unitary Schrödinger propagator $e^{-itP_h/h}$ along the trajectory $\rho(t) = \varphi_p^t(\rho_0)$. The twisted propagator $\mathcal{U}_\theta^t = e^{-itP_{h,\theta}/h}$ also transports the wavepacket along the trajectory $(\rho(t))$, but the nonselfadjoint character of $P_{h,\theta}$ will have the effect to modify the norm of the wavepacket:

$$\frac{d}{dt}\|u(t)\|^2 = \frac{2}{h}\mathrm{Im}\langle u(t), P_{h,\theta}u(t)\rangle \approx \frac{2\mathrm{Im}\, p_\theta(\rho(t))}{h}\|u(t)\|^2\,,$$

where p_θ is the principal symbol of $P_{h,\theta}$. When $x(t)$ is outside $B(R_0)$, this symbol reads $p_\theta(x,\xi) = e^{-2i\theta}|\xi|^2$, so at the point $\rho(t)$ its imaginary part is $-\sin(2\theta)E < 0$. As a result, the norm of $u(t)$ decreases very fast: its norm is reduced to $\mathcal{O}(h^\infty)$ as soon as $\rho(t)$ exits $B(R_0)$: the twisted propagator is *strongly absorbing* outside $B(R_0)$.

3.3 Resonances vs. classical trapped set. As explained before, the distribution of resonances in rectangles $R(E, Ch, \gamma h)$ depend crucially on the dynamics of φ_p^t on the trapped set K_E. Let us explain more precisely how this connection operates, starting with the simple case of a nontrapping dynamics.

3.3.1 Case of a nontrapping dynamics. If $K_E = \emptyset$, any point $\rho_0 \in p^{-1}(E)$ will leave $B(R_0)$ within a finite time T_0. As a result, a wavepacket u_{ρ_0} microlocalized on ρ_0 will be transported by $\mathcal{U}_\theta^t$ outside of $B(R_0)$, and will be absorbed. Let us now assume that $v_z \in L^2(M)$ satisfies $(P_{h,\theta} - z)v_z = 0$, for some $z \in R(E, Ch, \gamma h)$. Elliptic estimates show that v_z can be decomposed as a sum of (normalized) coherent states centered inside a small neighbourhood $U(E)$ of $p^{-1}(E) \cap T^*B(R_0)$:

$$v_z = \int_{U(E)} \frac{d\rho}{(2\pi h)^d} \langle u_\rho, v_z \rangle\, u_\rho + \mathcal{O}(h^\infty)\,. \tag{12}$$

Let us apply the propagator $\mathcal{U}_\theta^{T_0}$ to the above equality. On the right hand side each evolved wavepacket $\mathcal{U}_\theta^{T_0} u_\rho = \mathcal{O}(h^\infty)$ from the above discussion, while on the left hand side we get $\mathcal{U}_\theta^{T_0} v_z = e^{-izT_0/h} v_z$. The equality between both sides contradicts our assumption $\mathrm{Im} z \geq -\gamma h$. This argument shows that if $K_E = \varnothing$, deeper rectangles $R(E, Ch, \gamma h|\log h|)$ are also empty of resonances Martinez [2002].

3.3.2 Fractal hyperbolic trapped set. We now consider a nontrivial hyperbolic trapped set K_E. In this cases resonances generally exist in $R(E, Ch, \gamma h)$, at least when C and γ are large enough. In Section 2 we have mentioned the case where K_E is composed of a single hyperbolic periodic orbit, for which one can derive explicit asymptotic expressions for the resonances. In case of a more complex, fractal chaotic trapped set, we don't have any explicit expressions at our disposal. Yet, semiclassical methods provide upper bounds for the number of resonances inside $R(E, Ch, \gamma h)$, in terms of the Minkowski dimension of the trapped set K_E.

Theorem 1 (Fractal Weyl upper bound)**.** *Assume the trapped set K_E is a hyperbolic repeller of upper Minkowski dimension $1 + 2\delta$. Then, for any $C, \gamma > 0$, there exits $C_{C,\gamma} > 0$ and h_0 such that*

$$\forall h < h_0, \qquad \#\mathrm{Res}(P_h) \cap R(E, Ch, \gamma h) \leq C_{C,\gamma} h^{-\delta}\,. \tag{13}$$

The Minkowski dimension is a type of fractal dimension, often called "box dimension". Essentially, it indicates that the volumes of the ϵ-neighbourhoods of K_E (inside $p^{-1}(E)$) decay as $\epsilon^{2d-1-(1+2\delta)}$ when $\epsilon \to 0$.

The above theorem was first proved in Sjöstrand [1990] (for wider rectangles), and then refined by Sjöstrand and Zworski [2007], both in the case of smooth symbols $p(x, \xi)$. The case of Schottky hyperbolic surfaces was addressed by Zworski [1999] using semiclassical methods, and generalized to hyperbolic manifolds of higher dimension in Guillopé, Lin, and Zworski [2004] by using transfer operators. The case of scattering by $N \geq 3$ convex obstacles was tackled in Nonnenmacher, Sjöstrand, and Zworski [2014], using quantum monodromy operators (quantizations of Poincaré maps).

The bound (13) is called a *fractal Weyl upper bound*, by analogy with the selfadjoint semiclassical Weyl's law. Indeed, assume we add to P_h a confining potential $\tilde{V}(x)$, so that any energy shell $\tilde{p}^{-1}(E)$ is compact. The spectrum of $\tilde{P}_h$ is then discrete, and the

following semiclassical Weyl's law holds near noncritical energies E:
(14)
$$\mathrm{Spec}(\tilde{P}_h) \cap [E - Ch, E + Ch] = \frac{1}{(2\pi h)^d} \mathrm{Vol}\big(\tilde{p}^{-1}([E - Ch, E + Ch])\big) + \mathcal{O}(h^{-d+1}) .$$

The volume on the right hand side behaves as $C\mathcal{V}_E h^{-d+1}$ for some $\mathcal{V}_E > 0$, while the trapped set $\tilde{K}_E$ has dimension $1 + 2(d-1)$, so the power in the above estimate agrees with (13).

The result (13) and the above selfadjoint Weyl's law differ on several aspects:

1. (13) is an upper bound, not an asymptotics. Numerical studies have suggested that this upper bound should be sharp at the level of the order $h^{-\delta}$, at least if γ is large enough. Yet, proved lower bounds for the counting function are of smaller order $\mathcal{O}(1)$, similar with the case of a single hyperbolic orbit. A counting function $\asymp h^{-\delta}$ could already be called a *fractal Weyl's law*.

2. If a more precise estimate should hold, what could be the optimal constant $C_{C,\gamma}$? How does it depend on the depth γ? This question is related with the gap question discussed in the next section.

This conjectural fractal Weyl's law has been tested numerically on various chaotic systems, with variable success: Schrödinger operator with a smooth potential Lin [2002], hyperbolic surfaces by Guillopé, Lin, and Zworski [2004] and Borthwick [2014], discrete time analogues of scattering systems (quantized open maps) in Nonnenmacher and Zworski [2007], and even experimentally in the case of the scattering by N disks, see Potzuweit, Weich, Barkhofen, Kuhl, Stöckmann, and Zworski [2012].

3.3.3 Sketch of the proof of the Fractal Weyl upper bound. The spectrum of a nonselfadjoint operator Q is notoriously harder to identify than in the selfadjoint case. To study the spectrum of Q near some value z_0, one method is to "hermitize" the operator Q, namely study the bottom of the spectrum of the positive operator $(Q - z_0)^*(Q - z_0)$, or equivalently the small *singular values* of the operator $Q - z_0$; estimates on the number of singular values will then, through Weyl's inequalities, deliver upper bounds on the number of small eigenvalues of $Q - z_0$. It is much more difficult to obtain lower bounds on the number of eigenvalues: this difficulty explains the large gap between upper and lower bounds.

In our problem, to obtain a sharp upper bound we need to twist again the operator $P_{h,\theta}$, by conjugating it with an operator G_h obtained by quantizing a well-chosen *escape function* $g(x, \xi)$:
$$P_{h,G} \stackrel{\text{def}}{=} e^{-G_h} P_{h,\theta}\, e^{G_h} .$$

Through this conjugation, the symbol of the operator can be expanded as

$$p_G = p_\theta - ih\{p_\theta, g\} + smaller\,,$$

where the Poisson bracket $\{p_\theta, g\}$ represents the time derivative of $g(\rho(t))$. Using the hyperbolicity of the flow, for any $\gamma > 0$ it is possible to construct a function g such that $\{p_\theta, g\}(\rho) \geq 2\gamma h$ as soon as $\mathrm{dist}(\rho, K_E) \geq h^{1/2}$: this function is called an "escape function", because it grows along the flow, strictly so outside of the neighbourhood

$$K_E(h^{1/2}) \stackrel{\mathrm{def}}{=} \{\rho \in p^{-1}(E); \mathrm{dist}(\rho, K_E) \leq h^{1/2}\}.$$

As a result, $\mathrm{Im}\, p_G(\rho) \leq -3/2\gamma h$ for ρ outside $K_E(h^{1/2})$. The above hermitization techniques imply that the eigenstates of $P_{h,G}$ with eigenvalues $z \in R(E, Ch, \gamma h)$ must be concentrated in $K_E(h^{1/2})$. Applying the selfadjoint Weyl's law (14) to this set (thickened to an h-energy slab), and expressing its volume in terms of the Minkowski dimension of K_E, leads to the bound (13). □

3.3.4 Improved fractal upper bounds on hyperbolic surfaces. Eventhough the dynamics of φ_p^t on K_E is used to construct the escape function, the upper bound (13) only depends on the geometry of K_E, and not really on the flow φ_p^t itself. More recently, finer techniques have been developed in the special case of hyperbolic surfaces, taking into account more efficiently the dynamics on K_E Naud [2014] and Dyatlov [2015]. In this case the Minkowski dimension of K_E is given by $1 + 2\delta$, with $\delta \in (0, 1)$ the dimension of the limit set Λ_Γ. The upper bound now has a threshold at the value $\gamma_{th} = \frac{1-\delta}{2}$, which corresponds to the decay rate of a cloud of classical particles. For $\gamma \geq \gamma_{th}$ ("deep resonances") the upper bound remains $\mathcal{O}(h^{-\delta})$, but for $\gamma < \gamma_{th}$ ("shallow resonances") the upper bound is of the form $\mathcal{O}(h^{-\alpha(\gamma)})$, with $\alpha(\gamma) < \delta$ an explicit function, which decreases when $\gamma \searrow 0$. Jakobson and Naud [2012] have actually conjectured that for $\gamma < \gamma_{th}$ and h small enough, the rectangle $R(E, Ch, \gamma h)$ should be empty of resonances. This conjectured gap has not been confirmed numerically.

4 Dynamical criteria for resonance gaps

Let us now come to the question of resonance gaps. As explained in the introduction (see (3)), in the case of the wave equation in odd dimension, a global resonance gap ensures that the time evolved wave locally decays at a precise rate. Such a gap therefore reflects the phenomenon of dispersion of the wave, which spreads (leaks) outside any given ball. In the semiclassical setting, we have seen in Section 3.3.1 that this leakage is easy to understand if the classical flow is nontrapping: in that case the leakage operates in a finite time T_0, following the classical escape of all trajectories.

When there exist trapped trajectories, the explanation of this leakage is more subtle, and requires to take into account the dynamics for long times. In the present situation, this dispersion is induced by a combination of two factors: the hyperbolicity of the classical flow on K_E, and Heisenberg's uncertainty principle, which asserts that a quantum state cannot be localized in a phase space ball of radius smaller than $h^{1/2}$.

Our main result, reproduced from Nonnenmacher and Zworski [2009], shows that the rate of this dispersion can be estimated by a certain topological pressure of the flow $\varphi_p^t \restriction_{K_E}$ (see (10)), which combines both the unstability of the flow with its complexity.

Theorem 2 (Pressure gap). *Assume that the trapped set K_E is a hyperbolic repeller, and that the topological pressure $\mathcal{P}(1/2) < 0$. Then, for any $\epsilon > 0$, $C > 0$, and for $h > 0$ small enough, the operator P_h has no resonance in the rectangle $R(E, Ch, (|\mathcal{P}(1/2)| - \epsilon)h)$.*

According to our discussion in Section 2.2, the pressure $\mathcal{P}(1/2)$ can take either positive or negative values, respectively in the case of "thick" or "thin" trapped sets. So the condition $\mathcal{P}(1/2) < 0$ characterizes systems with a "thin" enough trapped set. We notice that this bound is sharp in the case K_E consists in a single hyperbolic orbit (Section 2.1): in dimension $d = 2$, the pressure $\mathcal{P}(1/2) = -\lambda/2$, which asymptotically corresponds to the first line of resonances.

This pressure bound was proved by Patterson [1976] in the case of hyperbolic surfaces, by showing that the zeros of the Selberg zeta function satisfy $\mathrm{Re} s_j \leq \delta$. In this case, the negativity of the pressure is equivalent with the bound $\delta < 1/2$ (see Section 2.3 for the notations).

This pressure bound was proved in the case of scattering by $N \geq 3$ disks in $\mathbb{R}^2$, almost simultaneously and independently by Ikawa [1988] and by Gaspard and Rice [1989] (although the latter article does not satisfy the standards of mathematical rigour, it contains the crucial ideas of the proof, and was the first one to identify the pressure). The method used in Nonnenmacher and Zworski [2009], which we sketch below, relies on similar ideas as these articles, carried out in the general setting of a Schrödinger operator P_h.

4.1 Evolution of an individual wavepacket. Our aim is to show that if v_z is an eigenstate of $P_{h,\theta}$ with eigenvalue $z \approx E$, then $\mathrm{Im} z/h \leq \mathcal{P}(1/2) + \epsilon$. To do so we will study the propagation of v_z by the Schrödinger flow $\mathcal{U}_\theta^t = e^{-itP_{h,\theta}/h}$ for long times (we will need to push the evolution up to *logarithmic times* $t \sim C|\log h|$, with $C > 0$ independent of h). From the decomposition (12) into wavepackets, we see that it makes sense to study in a first step the evolution of individual wavepackets u_ρ, centered at some point ρ in the neighbourhood $U(E)$.

4.1.1 Hyperbolic dispersion of a wavepacket. Take a wavepacket u_0 centered on a point $\rho_0 \in K_E$; its semiclassical evolution transports it along $\varphi_p^t(\rho_0)$, but also stretches the wavepacket along the unstable direction $E^u(\rho(t))$, following the linearized evolution $d\varphi_p^t(\rho_0)$.This spreading can be understood from a simple 1-dimensional toy model, namely the Hamiltonian $q(x,\xi) = \lambda x \xi$, generating the Hamiltonian flow $x(t) = e^{\lambda t} x_0$, $\xi(t) = e^{-\lambda t}\xi_0$, a clearly hyperbolic dynamics. The quantum evolution is generated by $P_h = \lambda(x\frac{h}{i}\partial_x - ih/2)$; its propagator is a unitary dilation:

$$e^{-itP_h/h}u_0(x) = e^{-t\lambda/2}u_0(e^{-t\lambda}x)\,. \tag{15}$$

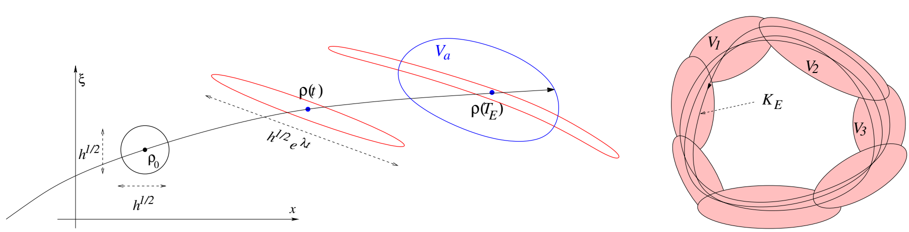
If we start from a the coherent state $u_0(x) = C_h e^{-\frac{x^2}{2h}}$ centered at the origin, the wavepacket at time $t > 0$ will have a horizontal (=unstable) width $e^{t\lambda}h^{1/2}$, while its amplitude will be reduced by a factor $e^{-t\lambda/2}$. The dynamics has dispersed the wavepacket along E^u.

Let us come back to our flow φ_p^t, and assume for simplicity that all the orbits of K_E have the same expansion rate $\lambda > 0$, in all unstable directions; this is the case for instance for the geodesic flow in constant curvature $\kappa = -\lambda^2$. In that case, the evolved wavepacket $u(t)$ spreads on a length $\sim e^{t\lambda}h^{1/2}$ along the unstable directions. By the time

$$T_E = \frac{|\log h|}{2\lambda}, \quad \text{which we call the } \textit{Ehrenfest time}, \tag{16}$$

the wavepacket $u(t)$ spreads on a distance ~ 1 along the unstable manifold $W^u(\rho(t))$, it is no more microscopic but becomes macroscopic. Some parts of $u(t)$ are now at finite distance from K_E; after a few time steps they will exit the ball $B(R_0)$ and hence be absorbed by the nonunitary propagator (see the left of Figure 9 for a sketch of this evolution).

Figure 9: Left: evolution of a minimal-uncertainty wavepacket: the evolved state stretches exponentially along the unstable directions. By the time T_E the state spreads outside a single cell V_a. Right: sketch of the partition (V_a), representing only the elements covering K_E.



4.1.2 Introducing a quantum partition. In order to precisely estimate the decay of $\|u(t)\|$, one needs to partition the phase space, such as to keep track of the portions of $u(t)$ which exit $B(R_0)$ (and are absorbed), and the ones which stay near K_E. One cooks up a finite partition $(V_a)_{a\in A}$ of the phase space T^*M (making it more precise near K_E), and quantizes the functions $\mathbb{1}_{V_a}$ to produce a family of microlocal truncations Π_a, satisfying $\sum_{a\in A}\Pi_a = Id_{L^2}$. The family $(\Pi_a)_{a\in A}$ is called a quantum partition.

We may insert this quantum partition at each integer step of the evolution: calling $\mathcal{U}_\theta = e^{-iP_{h,\theta}/h}$, we have for any time $N \in \mathbb{N}$:

$$(\mathcal{U}_\theta)^N = \sum_{\vec{a}=a_0,\cdots,a_N} \mathcal{U}_{\vec{a}}, \qquad \mathcal{U}_{\vec{a}} = \Pi_{a_N}\mathcal{U}_\theta \cdots \Pi_{a_2}\mathcal{U}_\theta\Pi_{a_1}\mathcal{U}_\theta\Pi_{a_0}\,,$$

where we sum over all possible words $\vec{a}$ of length $N+1$. We can control the action of the truncated propagators $\mathcal{U}_{\vec{a}}$ on our wavepacket. For times $N < T_E$, the evolved state $u(N) = \mathcal{U}_\theta^N u_0$ is dominated by a single term $\mathcal{U}_{\vec{a}}\, u_0$, where the word $\vec{a}$ is such that each point $\rho(j) \in V_{a_j}$. Around the Ehrenfest time $u(T_E)$ becomes macroscopic, so it is no more concentrated inside a single set V_a; the truncations Π_a will cut this state into several pieces, each one carrying a reduced norm. At each following time step, the evolution $\mathcal{U}_\theta$ continues to stretch the pieces $\mathcal{U}_{\vec{a}}\, u_0$ by a factor e^λ along the unstable directions, so several truncations will again act nontrivially. The norms of the pieces $\mathcal{U}_{\vec{a}}\, u_0$ can be estimated by the decay of the amplitude of the wavepacket, similarly as in the linear model (15) (there are now $(d-1)$ unstable directions):

(17)

$$\|\mathcal{U}_{\vec{a}}\, u_0\| \le \exp\Big(-\frac{\lambda(d-1)}{2}(N-T_E)\Big)\,\|u_0\| + \mathcal{O}(h^\infty), \quad N \ge T_E, \quad \vec{a} = a_0\cdots a_N\,.$$

For most words $\vec{a}$, this bound is not sharp. For instance, the symbols a_j corresponding to partition elements V_{a_j} outside of $B(R_0)$ indicate that the state is absorbed fast, and lead to $\mathcal{O}(h^\infty)$ terms. As a result, for $N > T_E$ the nonnegligible pieces correspond to words $\vec{a}$ such that almost all the elements V_{a_j} intersect the trapped set. Keeping only those "trapped" words, we obtain

$$\mathcal{U}_\theta^N u_0 = \sum_{\vec{a}\ \text{trapped}} \mathcal{U}_{\vec{a}}\, u_0 + \text{negligible}\,,$$

with each term bounded as in (17). A more careful analysis (involving a "good" choice of partition) shows that for long logarithmic times $N = C|\log h|$, $C \gg 1$, the number of relevant words is bounded above by $\exp(N(H_{top}+\epsilon))$, where H_{top} is the topological entropy (9), and $\epsilon > 0$ can be made arbitrary small by taking C large enough.

4.2 Evolving a general state. Take an eigenstate v_z with eigenvalue z near E. Being microlocalized near $p^{-1}(E)$, v_z can be decomposed into wavepackets u_ρ as in (12). By linearity, we find

$$\mathcal{U}_\theta^N v_z = \frac{1}{(2\pi h)^d} \sum_{\vec{a} \text{ trapped}} \int_{U(E)} d\rho \, \langle u_\rho, v_z \rangle \, \mathcal{U}_{\vec{a}} \, u_\rho \, + \mathcal{O}(h^\infty) \,,$$

where each term $\mathcal{U}_{\vec{a}} \, u_\rho$ is bounded as in (17). Applying the triangle inequality, we find

$$\| \mathcal{U}_\theta^N v_z \| \leq \frac{\mathrm{Vol}(U(E))}{(2\pi h)^d} e^{N(H_{top}+\epsilon)} e^{-\frac{\lambda(d-1)}{2}(N-T_E)} \,.$$

For a constant expansion rate, $\mathcal{P}(1/2) = H_{top} - \frac{\lambda(d-1)}{2}$, so the above bound can be recast as $h^{-\beta} e^{N(\mathcal{P}(1/2)+\epsilon)}$ for some $\beta > 0$. Taking $N = C|\log h|$ with C large enough, we can have $h^\beta \leq e^{N\epsilon}$, thereby giving a bound $e^{N(\mathcal{P}(1/2)+2\epsilon)}$. This bound is nontrivial if $\mathcal{P}(1/2)$ is negative. Using the fact that v_z is an eigenstate of eigenvalue $z \approx E$, we get for such a time N:

$$|e^{N \mathrm{Im} z/h}| \leq e^{N(\mathcal{P}(1/2)+2\epsilon)} \Longrightarrow \mathrm{Im} z/h \leq \mathcal{P}(1/2) + 2\epsilon \,.$$

□

4.3 Improving the pressure gap. In the case of a fractal hyperbolic repeller, the pressure bound of Theorem 2 is believed to be nonoptimal, at least for generic hyperbolic systems. Estimating $\|u(N)\|$ by adding the norms of the terms $\mathcal{U}_{\vec{a}} u_\rho$ does not take into account the partial cancellations between these terms. Indeed, when $N = C|\log h|$ with $C \gg 1$, many of those terms are almost proportional to each other, essentially differing by complex valued prefactors. The norm of their sum is hence governed by a sum of many complex factors, which is generally much smaller than the sum of their moduli.

Such partial cancellations (or "destructive interferences") are at the heart of Dologopyat's proof of the exponential decay of correlations for Anosov flows Dolgopyat [1998], when analyzing the spectrum of a family of transfer operators. Naud [2005] adapted Dolgopyat's method to show an improved high frequency resonance gap for the Laplacian on Schottky hyperbolic surfaces, still working at the level of transfer operators. By a similar (yet, more involved) method, Petkov and Stoyanov improved the high frequency resonance gap for scattering by convex obstacles on $\mathbb{R}^d$; these authors managed to establish a semiclassical connection between the quantum propagator and a transfer operator, thereby applying Dolgopyat's method to the former. All the above works improve the pressure bound by some small, not very explicit $\epsilon_1 > 0$.

In the case of hyperbolic surfaces, a recent breakthrough was obtained by Dyatlov and his collaborators. 2016 showed that a nontrivial gap for a hyperbolic surface with parameter $\delta \in (0, 1)$ results from a *fractal uncertainty principle* (FUP), a new type of estimate in 1-dimensional harmonic analysis. This FUP states that if $K \subset [0,1]$ is a Cantor set of dimension δ and $K(h)$ its h-neighbourhood, then there exists $\beta > 0$ such that

$$\|\mathbb{1}_{K(h)} \mathfrak{F}_h \mathbb{1}_{K(h)}\|_{L^2 \to L^2} \leq C h^{\beta} ,$$

where $\mathfrak{F}_h$ is the semiclassical Fourier transform. This estimate shows that a function $u \in L^2(\mathbb{R})$ and its semiclassical Fourier transform cannot be both concentrated on $K(h)$. This FUP obviously holds when $\delta < 1/2$, giving back the pressure bound. In a ground-breaking work Bourgain and Dyatlov [2016] managed to prove this FUP in the full range $\delta \in (0,1)$, thereby showing a resonance gap on any Schottky hyperbolic surface. The improved gap is not very explicit, it is much smaller than the gap $\frac{1-\delta}{2}$ conjectured by Jakobson-Naud.

Although the methods of Dyatlov and Zahl [2016] strongly rely on the constant negative curvature, it seems plausible to prove a resonance gap for any hyperbolic repeller in two space dimensions. On the other hand, the extension of an FUP to higher dimensional systems is at present rather unclear, partly due to the more complicated structure of the trapped sets.

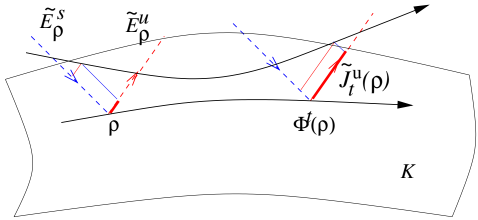
5 Normally hyperbolic trapped set

In this last section, we focus on a different type of trapped set. We assume that for some energy window $[E_1, E_2]$, the trapped set $K = K_{[E_1,E_2]} = \cup_{E \in [E_1,E_2]} K_E$ is a smooth, normally hyperbolic, symplectic submanifold of the energy slab $p^{-1}([E_1, E_2])$. What does this all mean? If K is a symplectic submanifold of T^*M, at each point $\rho \in K$ the tangent space $T_\rho(T^*M)$ splits into $T_\rho K \oplus (T_\rho K)^\perp$, where both are symplectic subspaces. Normal hyperbolicity means that the flow φ_p^t is hyperbolic transversely to K: the transverse subspace $(T_\rho K)^\perp = \tilde{E}^s(\rho) \oplus \tilde{E}^u(\rho)$, such that $d\varphi_p^t \restriction_{TK^\perp}$ contracts exponentially along $\tilde{E}^s(\rho)$, and expands along $\tilde{E}^u(\rho)$ (see Figure 10). We denote by $\tilde{J}_t^u(\rho) = |\det(d\varphi_p^t \restriction_{\tilde{E}_\rho^u})|$ the normal unstable Jacobian.

5.1 Examples of normally hyperbolic trapped sets.

5.1.1 Examples in chemistry and general relativity. This dynamical situation may occur in quantum chemistry, when modeling certain reaction dynamics. The reactants and products of the chemical reaction are two parts of phase space, connected by a hyperbolic "saddle" along two conjugate coordinates (x_1, ξ_1), similar with the linear dynamics of Section 4.1.1, while the evolution of the other coordinates remains bounded Waalkens, Schubert, and Wiggins [2008]. The trapped set $K_{[E_1,E_2]}$ is then a bounded piece of the space $\{x_1 = \xi_1 = 0\}$.

Figure 10: Sketch of a normally hyperbolic trapped set K.



This dynamical situation also occurs in general relativity, namely when describing time-like trajectories in the Kerr or Kerr-de Sitter black holes Wunsch and Zworski [2011] and Dyatlov [2012]. The trapped set is a normally hyperbolic manifold diffeomorphic to T^*S^2. In this situation resonances are replaced by quasinormal modes, obtained by solving a generalized spectral problem $P(z)u = 0$. Yet, the semiclassical methods sketched below can be easily adapted to this context.

5.1.2 From classical to quantum resonances. An original application of this dynamical assumption concerns the study of contact Anosov flows. A flow ϕ^t defined on a compact manifold M is said to be Anosov if at any point $x \in M$, the tangent space $T_x M$ splits into $\mathbb{R}\Xi(x) \oplus E^u(x) \oplus E^s(x)$, where $\Xi(x)$ is the vector generating the flow, while $E^s(x)$, $E^u(x)$ are the stable/unstable subspaces, satisfying the properties (7). The assumption that ϕ^t preserves a contact 1-form α, implies that the subspace $E^u(x) \oplus E^s(x)$, which forms the kernel of $d\alpha(x)$, depends smoothly on x.

The long time properties of such a flow are governed by a set of so-called Ruelle-Pollicott (RP) resonances $\{\lambda_k \subset \mathbb{C}_-\}$, which share many properties with the quantum resonances we have studied so far. Considering two test functions $u, v \in C^\infty(M)$, their *correlation function* $C_{v,u}(t) \overset{\text{def}}{=} \int_M dx\, v(x) u(\phi^t(x)) - \int dx\, v(x) \int dx\, u(x)$ can be expanded in terms of these RP resonances:

$$C_{v,u}(t) = \sum_{\mathrm{Im}\lambda_k \geq -\gamma} e^{-i\lambda_k t} \langle v, \Pi_{\lambda_k} u \rangle + \mathcal{O}_{u,v}(e^{-\gamma t}), \tag{18}$$

Hence, if the RP resonances λ_k satisfy a uniform gap, the correlation decays exponentially (one speaks of *exponential mixing*). Such a resonance gap has been first proved by Dolgopyat [1998] and Liverani [2004], while Tsujii [2010] proved an explicit bound for the high frequency gap.

Comparing (18) with (3), Faure and Sjöstrand [2011] had the idea to interpret the RP resonances (or rather $z_k = h\lambda_k$) as the "quantum resonances" of the "quantum Hamiltonian" $P_h = -ih\Xi$. Notice that $e^{-itP_h/h}u(x) = u(\phi^{-t}(x))$. What do we gain from this interpretation? The principal symbol of P_h, $p(x,\xi) = \xi(\Xi(x))$, generates on T^*M the symplectic lift of ϕ^t : $\varphi_p^t(x,\xi) = (\phi^t(x), {}^T d\phi^t(x)^{-1}\xi)$. As opposed to the scattering situation, each energy shell $p^{-1}(E)$ goes to infinity along the fibers of T^*M. Hence, for any energy $E \in \mathbb{R}$, the trapped set K_E is given by the points $\rho = (x,\xi) \in p^{-1}(E)$ such that ${}^T d\phi^t(x)^{-1}\xi$ remains bounded when $t \to \pm\infty$. From the hyperbolicity structure, this is possible only if $\xi = E\alpha_x$. Hence, $K_E = \{(x, \xi = E\alpha_x),\ x \in M\}$, a smooth submanifold of $p^{-1}(E)$. It is easy to check that $K = \cup_E K_E$ is symplectic, and normally hyperbolic (the subspaces $\tilde{E}^{s/u}$ are lifts of the subspaces $E^{s/u}$ of TM). The resonances of the quantum Hamiltonian P_h can thus be connected with the properties of this trapped set.

The main difficulty when analyzing this classical dynamical problem as a "quantum scattering" one Faure and Sjöstrand [ibid.], is to twist the selfadjoint operator P_h, such as to transform the resonances into eigenvalues. This was done by constructing spaces of anisotropic distributions $\mathcal{H}^m \subset \mathfrak{D}'(M)$, such that $P_h : \mathcal{H}^m \to \mathcal{H}^m$ has discrete spectrum in $\{\mathrm{Im} z \geq -mh\}$, made of "uncovered" Ruelle-Pollicott resonances. We will not detail this construction, which can also be presented as a twist of the operator P_h into a nonselfadjoint operator $P_{h,m}$ on $L^2(M)$.

5.2 An explicit resonance gap for normal hyperbolic trapped sets. Let us come back to our general setting, and start again to propagate minimum-uncertainty wavepackets u_ρ centered on a point $\rho \in K$. Due to the normal hyperbolicity, the state $e^{-itP_h/h}u_\rho$ spreads exponentially fast along the transverse unstable direction $\tilde{E}^u$. Similarly as what we did in Section 3.3.3, one can twist the operator P_h by a microlocal weight G_h, such that the twisted operator $P_{h,G}$ is absorbing outside the neighbourhood $K(Ch^{1/2})$. After a few time steps, the evolved wavepacket will leak outside of this neighbourhood, and be partially absorbed: their norms will decay at the rate

$$\|e^{-itP_{h,G}/h}u_\rho\| \leq C\, \tilde{J}_t^u(\rho)^{-1/2}, \qquad t > 0.$$

If we call $\tilde{\Lambda}_{\min} = \liminf_{t\to\infty} \frac{1}{t} \inf_{\rho\in K} \log \tilde{J}_t^u(\rho)$ the minimal growth rate of the transverse unstable Jacobian, for any $\epsilon > 0$ and $t > t_\epsilon$ large enough, the above right hand sides are bounded by $e^{-t(\tilde{\Lambda}_{\min}/2-\epsilon)}$. With more work, one can show that this uniform decay of our individual wavepackets induces the same decay of any state microlocalized on K, in particular of any eigenstate v_z of $P_{h,G}$. One then obtains the following gap estimate for the eigenvalues of $P_{h,G}$, or equivalently the resonances of P_h Nonnenmacher and Zworski [2015]:

Theorem 3 (Resonance gap, normally hyperbolic trapped set)**.** *Assume the trapped set $K = K_{[E-c,E+c]}$ is normally hyperbolic, with minimal transverse growth rate $\tilde{\Lambda}_{\min}$. Then, for any $\epsilon > 0$ and $h > 0$ small enough, the rectangle $R(E, c, (\tilde{\Lambda}_{\min}/2 - \epsilon)h)$ contains no resonance.*

Like in the case of Theorem 2 and its improvements, we also obtain a bound for the truncated resolvent operator inside the rectangle, of the form $\|\chi(P_h - z)^{-1}\chi\| \leq h^{-\beta}$, $\chi \in C_c^\infty(M)$. When applying this result to the situation of Section 5.1.2 (mixing of contact Anosov flows), we exactly recover Tsujii's gap for the high frequency RP resonances.

In two of the settings presented above (the resonances of Kerr-de Sitter spacetimes Dyatlov [2016], respectively the Ruelle-Pollicott for contact Anosov flows Faure and Tsujii [2013], the spectrum of resonances has been shown to enjoy a richer structure, provided certain bunching conditions on the rates of expansion are satisfied. Namely, beyond the first gap stated in the above theorem, resonances are gathered in a (usually finite) sequence of parallel strips, separated by secundary resonance free strips. The widths of the strips are expressed in terms of maximal and minimal expansion rates similar with $\Lambda_{\min}$. Besides, the number of resonances along each of the strips satisfies a Weyl's law, corresponding to the volume of the $h^{1/2}$ neighbourhood of K.

Acknowledgments. The author has benefitted from invaluable interactions with many colleagues, in particular M.Zworski who introduced him to the topic of chaotic scattering, as well as V.Baladi, S.Dyatlov, F.Faure, F.Naud, J.Sjöstrand. In the past few years he has been partially supported by the grant Gerasic-ANR-13-BS01-0007-02 awarded by the Agence Nationale de la Recherche.

References

J. Aguilar and J. M. Combes (1971). "A class of analytic perturbations for one-body Schrödinger Hamiltonians". *Comm. Math. Phys.* 22, pp. 269–279. MR: 0345551 (cit. on p. 2523).

David Borthwick (2014). "Distribution of resonances for hyperbolic surfaces". *Exp. Math.* 23.1, pp. 25–45. MR: 3177455 (cit. on p. 2526).

– (2016). *Spectral theory of infinite-area hyperbolic surfaces*. Second. Vol. 318. Progress in Mathematics. Birkhäuser/Springer, [Cham], pp. xiii+463. MR: 3497464 (cit. on p. 2521).

Jean Bourgain and Semyon Dyatlov (Dec. 2016). "Spectral gaps without the pressure condition". To appear in *Ann. of Math.* arXiv: 1612.09040 (cit. on p. 2532).

Rufus Bowen and David Ruelle (1975). "The ergodic theory of Axiom A flows". *Invent. Math.* 29.3, pp. 181–202. MR: 0380889 (cit. on p. 2518).

Dmitry Dolgopyat (1998). "On decay of correlations in Anosov flows". *Ann. of Math. (2)* 147.2, pp. 357–390. MR: 1626749 (cit. on pp. 2531, 2533).

Semyon Dyatlov (2012). "Asymptotic distribution of quasi-normal modes for Kerr–de Sitter black holes". *Ann. Henri Poincaré* 13.5, pp. 1101–1166. MR: 2935116 (cit. on p. 2533).

– (Dec. 2015). "Improved fractal Weyl bounds for hyperbolic manifolds, with an appendix with D. Borthwick and T. Weich". To appear in *J.Eur.Math.Soc.* arXiv: 1512.00836 (cit. on p. 2527).

– (2016). "Spectral gaps for normally hyperbolic trapping". *Ann. Inst. Fourier (Grenoble)* 66.1, pp. 55–82. MR: 3477870 (cit. on p. 2535).

Semyon Dyatlov and Joshua Zahl (2016). "Spectral gaps, additive energy, and a fractal uncertainty principle". *Geom. Funct. Anal.* 26.4, pp. 1011–1094. MR: 3558305 (cit. on pp. 2531, 2532).

Semyon Dyatlov and Maciej Zworski (Apr. 2018). "Mathematical theory of scattering resonances" (cit. on p. 2514).

Frédéric Faure and Johannes Sjöstrand (2011). "Upper bound on the density of Ruelle resonances for Anosov flows". *Comm. Math. Phys.* 308.2, pp. 325–364. MR: 2851145 (cit. on p. 2534).

Frédéric Faure and Masato Tsujii (2013). "Band structure of the Ruelle spectrum of contact Anosov flows". *C. R. Math. Acad. Sci. Paris* 351.9-10, pp. 385–391. MR: 3072166 (cit. on p. 2535).

Pierre Gaspard and Stuart A. Rice (1989). "Semiclassical quantization of the scattering from a classically chaotic repellor". *J. Chem. Phys.* 90.4, pp. 2242–2254. MR: 980393 (cit. on p. 2528).

Christian Gérard and Johannes Sjöstrand (1987). "Semiclassical resonances generated by a closed trajectory of hyperbolic type". *Comm. Math. Phys.* 108.3, pp. 391–421. MR: 874901 (cit. on p. 2519).

Laurent Guillopé, Kevin K. Lin, and Maciej Zworski (2004). "The Selberg zeta function for convex co-compact Schottky groups". *Comm. Math. Phys.* 245.1, pp. 149–176. MR: 2036371 (cit. on pp. 2525, 2526).

Mitsuru Ikawa (1983). "On the poles of the scattering matrix for two strictly convex obstacles". *J. Math. Kyoto Univ.* 23.1, pp. 127–194. MR: 692733 (cit. on p. 2519).

– (1988). "Decay of solutions of the wave equation in the exterior of several convex bodies". *Ann. Inst. Fourier (Grenoble)* 38.2, pp. 113–146. MR: 949013 (cit. on p. 2528).

Dmitry Jakobson and Frédéric Naud (2012). "On the critical line of convex co-compact hyperbolic surfaces". *Geom. Funct. Anal.* 22.2, pp. 352–368. MR: 2929068 (cit. on p. 2527).

Kevin K. Lin (2002). "Numerical study of quantum resonances in chaotic scattering". *J. Comput. Phys.* 176.2, pp. 295–329. MR: 1894769 (cit. on p. 2526).

Carlangelo Liverani (2004). "On contact Anosov flows". *Annals of Math. (2)* 159.3, pp. 1275–1312. MR: 2113022 (cit. on p. 2533).

André Martinez (2002). "Resonance free domains for non globally analytic potentials". *Ann. Henri Poincaré* 3.4, pp. 739–756. MR: 1933368 (cit. on p. 2525).

Frédéric Naud (2005). "Expanding maps on Cantor sets and analytic continuation of zeta functions". *Ann. Sci. École Norm. Sup. (4)* 38.1, pp. 116–153. MR: 2136484 (cit. on pp. 2523, 2531).

– (2014). "Density and location of resonances for convex co-compact hyperbolic surfaces". *Invent. Math.* 195.3, pp. 723–750. MR: 3166217 (cit. on p. 2527).

Stéphane Nonnenmacher, Johannes Sjöstrand, and Maciej Zworski (2014). "Fractal Weyl law for open quantum chaotic maps". *Ann. of Math. (2)* 179.1, pp. 179–251. MR: 3126568 (cit. on p. 2525).

Stéphane Nonnenmacher and Maciej Zworski (2007). "Distribution of resonances for open quantum maps". *Comm. Math. Phys.* 269.2, pp. 311–365. MR: 2274550 (cit. on p. 2526).

– (2009). "Quantum decay rates in chaotic scattering". *Acta Math.* 203.2, pp. 149–233. MR: 2570070 (cit. on p. 2528).

– (2015). "Decay of correlations for normally hyperbolic trapping". *Invent. Math.* 200.2, pp. 345–438. MR: 3338007 (cit. on p. 2534).

Samuel J. Patterson (1976). "The limit set of a Fuchsian group". *Acta Math.* 136.3-4, pp. 241–273. MR: 0450547 (cit. on p. 2528).

Vesselin Petkov and Luchezar Stoyanov (2010). "Analytic continuation of the resolvent of the Laplacian and the dynamical zeta function". *Anal. PDE* 3.4, pp. 427–489. MR: 2718260.

Alexander Potzuweit, Tobias Weich, Sonia Barkhofen, Ulrich Kuhl, Hans-Jürgen Stöckmann, and Maciej Zworski (Dec. 2012). "Weyl asymptotics: From closed to open systems". *Phys. Rev. E* 86 (6), p. 066205 (cit. on p. 2526).

Johannes Sjöstrand (1990). "Geometric bounds on the density of resonances for semiclassical problems". *Duke Math. J.* 60.1, pp. 1–57. MR: 1047116 (cit. on p. 2525).

Johannes Sjöstrand and Maciej Zworski (2007). "Fractal upper bounds on the density of semiclassical resonances". *Duke Math. J.* 137.3, pp. 381–459. MR: 2309150 (cit. on p. 2525).

Masato Tsujii (2010). "Quasi-compactness of transfer operators for contact Anosov flows". *Nonlinearity* 23.7, pp. 1495–1545. MR: 2652469 (cit. on p. 2533).

Holger Waalkens, Roman Schubert, and Stephen Wiggins (2008). "Wigner's dynamical transition state theory in phase space: classical and quantum". *Nonlinearity* 21.1, R1–R118. MR: 2376314 (cit. on p. 2532).

Jared Wunsch and Maciej Zworski (2011). "Resolvent estimates for normally hyperbolic trapped sets". *Ann. Henri Poincaré* 12.7, pp. 1349–1385. MR: 2846671 (cit. on p. 2533).

Maciej Zworski (1999). "Dimension of the limit set and the density of resonances for convex co-compact hyperbolic surfaces". *Invent. Math.* 136.2, pp. 353–409. MR: 1688441 (cit. on p. 2525).

– (2012). *Semiclassical analysis*. Vol. 138. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xii+431. MR: 2952218 (cit. on p. 2516).

– (2017). "Mathematical study of scattering resonances". *Bull. Math. Sci.* 7.1, pp. 1–85. MR: 3625851 (cit. on p. 2514).

Received 2018-02-28.

Stéphane Nonnenmacher
Laboratoire de Mathématique d'Orsay
Univ. Paris-Sud, CNRS
Université Paris-Saclay
91405 Orsay
France
stephane.nonnenmacher@u-psud.fr

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2537–2558)

FLUIDS, WALLS AND VANISHING VISCOSITY

HELENA J. NUSSENZVEIG LOPES AND MILTON C. LOPES FILHO

Abstract

The vanishing viscosity problem consists of understanding the limit, or limits, of solutions of the Navier-Stokes equations, with viscosity ν, as ν tends to zero. The Navier-Stokes equations are a model for real-world fluids and the parameter ν represents the ratio of friction, or resistance to shear, and inertia. Ultimately, the relevant question is whether a real-world fluid with very small viscosity can be approximated by an ideal fluid, which has no viscosity. In this talk we will be primarily concerned with the classical open problem of the vanishing viscosity limit of fluid flows in domains with boundary. We will explore the difficulty of this problem and present some known results. We conclude with a discussion of criteria for the vanishing viscosity limit to be a solution of the ideal fluid equations.

Introduction

The vanishing viscosity problem is a classical one in fluid dynamics. In its simplest form, the question is to understand under which circumstances the behavior of real world fluids can be well-approximated by that of ideal, or frictionless, fluids. Said differently, when can small viscosity be realistically neglected? The purpose of this article is to discuss some of the current knowledge concerning this problem.

More precisely, we will be focusing on incompressible Newtonian fluids. In addition, we are specifically interested in the interaction of fluid flows with rigid boundaries. Finally, without ignoring the physics, we will be primarily concerned with the mathematical treatment of this problem. We will ignore the important issues surrounding the computational modeling of such flows.

We begin our discussion with the physical description of the small viscosity flow regime and, in particular, with Ludwig Prandtl's contributions. After discussing the discrepancy between slightly viscous and non-viscous flow near a solid boundary, we will explore what

M.C.L.F. was partially supported by CNPq grant #306886/2014-6. H.J.N.L. was supported in part by CNPq grant #307918/2014-9 and FAPERJ grant #E-26/202.950/2015.

MSC2010: primary 35Q31; secondary 35Q30, 76D10, 76F06.

is known if this discrepancy propagates into the bulk of the fluid, and we conclude with criteria under which the small viscosity limit may still be considered ideal, despite this discrepancy.

Hereafter, *viscosity* is understood as the inverse of the Reynolds number, a non-dimensional constant which parameterizes incompressible flow. Roughly, the Reynolds number captures the relative importance of inertia and friction, with large Reynolds number flows dominated by inertia. It incorporates the scale and characteristic speed of interest, along with the kinematic viscosity of the fluid.

1 L. Prandtl and boundary layers

"... the behaviour of a fluid of small viscosity μ may, on account of boundary layer separation, be completely different to that of a (hypothetical) fluid of no viscosity at all." Acheson [p. 30 1990].

A fundamental part of the study of fluid motion is understanding the interaction of fluids with solid objects. A natural point of departure for this discussion is the fact that the interaction of incompressible flow with a solid object is completely different if the flow has very small viscosity or none at all. This fact was observed in experiments, long before a consensus physical theory for it became available, see Acheson [p. 264 ibid.].

The relevant physical theory was proposed in a 7.5-page paper delivered at the Third ICM in Heidelberg, in 1904, by Ludwig Prandtl [1905]. At the time, Prandtl was a young fluid dynamicist transitioning from the University of Hannover to Gottingen. This remarkable short paper contains several new ideas, among which are the foundations of Boundary Layer Theory. In this section we briefly present its main ideas. We refer the reader to the classical text Schlichting [1960] for a broad discussion.

Prandtl assumes that a viscous fluid does not slip along the boundary, something which was still controversial at the time. Further, Prandtl's model describes the flow as two separate, yet interacting parts: in one part, far from the boundary, the flow can be treated as inviscid, and satisfies, in particular, the conservation laws of ideal fluid theory. The second part, localized near the rigid boundary, is where viscous effects are important. Prandtl refers to the region near the boundary as a *transition* or *boundary layer* and he suggests that, within this layer, the tangential velocity varies rapidly in the normal direction, while the normal velocity is slowly varying; together, they vanish at the boundary and interpolate the inviscid flow in the bulk of the fluid domain. Prandtl derives a system of partial differential equations which is an asymptotic model for the flow in the boundary layer; these equations are now known as the *Prandtl equations*. Furthermore, he estimates the thickness of the boundary layer as being of the order of the square-root of viscosity. Lastly, Prandtl notes that the boundary layer may separate, or detach, from the boundary and entrain into the bulk of the flow, even for flows with small, yet positive, viscosity.

Let us be more precise. The standard mathematical model for incompressible viscous flows is given by the ν-Navier-Stokes equations, which we write as

$$\begin{cases} \partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla p + \nu\Delta\mathbf{u} + \mathbf{f}, & \text{in } (0,+\infty)\times\mathfrak{D}; \\ \operatorname{div}\mathbf{u} = 0, & \text{in } [0,+\infty)\times\mathfrak{D}. \end{cases} \tag{1-1}$$

Above, $\mathfrak{D}$ is the physical domain of the fluid, $\mathbf{u} = \mathbf{u}(t,\mathbf{x}) = (u_1,u_2,u_3)(t,\mathbf{x})$ represents the fluid velocity at time $t \geq 0$ and at the point $\mathbf{x} \in \mathfrak{D}$, and $p = p(t,\mathbf{x})$ is the scalar pressure. The viscosity is $\nu > 0$, and $\mathbf{f}$ represents a given external force, which we assume to vanish throughout this paper. The *no slip* boundary condition translates to

$$\mathbf{u} = 0 \text{ on } (0,+\infty)\times\partial\mathfrak{D}. \tag{1-2}$$

The equations for inviscid, or ideal, fluid flow are known as the Euler equations and are given by setting $\nu = 0$ in (1-1). The boundary condition for the Euler equations in a domain with boundary is the *non-penetration* boundary condition:

$$\mathbf{u}\cdot\mathbf{n} = 0 \text{ on } (0,+\infty)\times\partial\mathfrak{D}, \tag{1-3}$$

where $\mathbf{n}$ is the unit normal exterior to $\partial\mathfrak{D}$.

The mismatch between (1-2) and (1-3) is at the heart of the difficulty in addressing the small viscosity problem.

For simplicity, let us assume that $\mathfrak{D}$ is the two-dimensional half-plane $\mathbb{H} = \{\mathbf{x} = (x,y) \in \mathbb{R}^2 \mid y > 0\}$. Prandtl's theory applies to situations where the flow $\mathbf{u} = (u,v)$ can be described as a boundary layer flow $\mathbf{u}^{BL}$ inside a layer of thickness $\delta = \delta(\nu)$, superimposed to a mainstream, inviscid flow $\mathbf{U} = (U,V)$ at a distance greater than δ from the boundary. Asymptotic matching may be used to derive an approximate model for the behavior of the flow in the boundary layer. Rescaling the problem with respect to δ and introducing the variable $Y = y/\delta$ yields $\delta = \sqrt{\nu}$ and leads to the boundary layer equations below.

Let $\mathbf{u}^P = (u^P, v^P)$. The Prandtl equations are given by

$$\begin{cases} \partial_t u^P + (\mathbf{u}^P\cdot\widetilde{\nabla})u^P = -\partial_x p^P + \partial_Y^2 u^P, & \text{in } (0,+\infty)\times\mathbb{H}; \\ \partial_Y p^P = 0, & \text{in } [0,+\infty)\times\mathbb{H}; \\ \widetilde{\operatorname{div}}\,\mathbf{u}^P = 0, & \text{in } [0,+\infty)\times\mathbb{H}, \end{cases} \tag{1-4}$$

where $\widetilde{\nabla} = (\partial_x, \partial_Y)$ and $\widetilde{\operatorname{div}}\,\mathbf{u}^P = \widetilde{\nabla}\cdot\mathbf{u}^P$.

We expect that the solution $\mathbf{u} = (u, v)$ of (1-1) is well-approximated, for small $\nu > 0$, by

$$\mathbf{u} = (u, v) \sim \begin{cases} \mathbf{U} = (U, V), \text{ if } y >> \sqrt{\nu}, \\ \mathbf{u}^P = (u^P, \sqrt{\nu} v^P), \text{ if } 0 < y << \sqrt{\nu}. \end{cases} \tag{1-5}$$

The boundary conditions for the Prandtl solution should be:

$$\mathbf{u}^P(t, x, 0) = (0, 0) \text{ and } \lim_{Y \to +\infty} \mathbf{u}^P(t, x, Y) = (U(t, x, 0), 0), \quad \text{on } (0, +\infty) \times \mathbb{R}. \tag{1-6}$$

In addition, since the Prandtl pressure is independent of the vertical variable Y we have, taking the limit $Y \to \infty$, that the Prandtl pressure is matched to the Euler pressure at the boundary. Therefore,

$$p^P(t, x, Y) = p^E(t, x, 0), \quad \text{in } (0, +\infty) \times \mathbb{H},$$

so that

$$-\partial_x p^P = (\partial_t U + U \partial_x U)\big|_{(t,x,0)}, \quad \text{in } (0, +\infty) \times \mathbb{H}. \tag{1-7}$$

Finally, the initial data for the Prandtl equations is chosen so that the approximation in (1-5) is verified at $t = 0$. We write:

$$\mathbf{u}_0^P(x, Y) \equiv \mathbf{u}^P(0, x, Y). \tag{1-8}$$

In summary, given an initial flow $\mathbf{u}_0$, Prandtl's boundary layer theory, therefore, is based on the *ansatz* that the viscous flow whose initial velocity is $\mathbf{u}_0$ is well-approximated by the inviscid flow with the same initial velocity, far from the boundary, superimposed with an interpolating field near the boundary. As Prandtl himself notes, the agreement of this theory with experiments happens only in very particular situations, such as special laminar flows. The key observation is that the boundary layer is a thin region of intense shear next to the boundary.

This theory is not valid, even for laminar flows, when *boundary layer separation* occurs. This is when the boundary layer *detaches* from the boundary and affects the inviscid downstream flow. It is the case with flow past a cylinder, flow past a finite flat plate, past a corner, etc.

From a mathematical point-of-view, Prandtl's boundary layer theory introduces two natural questions. First, under which conditions are the Prandtl equations well-posed, and, second, given solutions of Prandtl's equations, when can the validity of the large Reynolds number asymptotics be rigorously verified. These two questions are obviously related, but much more is known regarding the former.

The rigorous study of the Prandtl equations began with the work of O. Oleinik, in the 1960s, where the steady and time-dependent problems were studied, in several scenarios, see for instance Oleĭnik [1963, 1966] and Oleĭnik and Samokhin [1999]. The following result is particularly noteworthy.

Theorem 1.1. (O. A. Oleinik, 1967, Oleĭnik [1966]) *Let* $u_0^P = u^P(0,x,Y)$ *satisfy the monotonicity condition* $\partial_Y u_0^P > 0$ *in* $\mathbb{H}$*, and assume that* $U = U(t,x,0) > 0$ *for all* $t \geq 0$*,* $x \in \mathbb{R}$*. Then there exists a local, classical, solution of the Prandtl equations* (1-4)*, subject to the initial and boundary conditions* (1-8)*,* (1-6)*, with pressure given by* (1-7)*.*

The proof of this theorem is based on a clever time-dependent change of variables called *Crocco* transform. The condition $U = U(t,x,0) > 0$ is the "no back-flow" condition, known to prevent boundary layer separation, as does the monotonicity condition. We observe, however, that there is no proof that the Prandtl approximation holds under these conditions. We refer the reader to Kelliher [2017] and Constantin, Kukavica, and Vicol [2015] for a discussion of the inviscid limit in this context.

A different, energy-based proof of Oleinik's theorem, still assuming the Oleinik monotonicity condition and no back-flow, was obtained in Masmoudi and Wong [2015] and in Alexandre, Y.-G. Wang, Xu, and Yang [2015], see also Kukavica, Masmoudi, Vicol, and Wong [2014].

Well-posedness of Prandtl's equations has also been studied in the analytic setting. R. Caflisch and M. Sammartino studied flows in a half-plane and proved, see Sammartino and Caflisch [1998b], that both Euler and Prandtl are locally well-posed for real analytic data. In Sammartino and Caflisch [1998a] they went on to show that, under the assumption of real-analyticity of the initial Euler and Prandtl velocities, the solution of the Navier-Stokes equations is well-approximated as in (1-5), assuming the initial data for the Navier-Stokes equations satisfies the same asymptotics which, in certain contexts, reads as $\mathbf{u}_0^\nu = \mathbf{u}_0^\nu(x,y) = \mathbf{u}_0^E(x,y) + (u_0^P, \sqrt{\nu} v_0^P)(x,Y) + \mathcal{O}(\sqrt{\nu})$. In addition, results on local well-posedness for Prandtl, assuming only tangential analyticity have been obtained, see Lombardo, Cannone, and Sammartino [2003] and Kukavica and Vicol [2013].

The problem in the analytic setting is, therefore, well-understood. However, as pointed out in Grenier, Guo, and Nguyen [2015], analytic regularity is too much to expect in real-world flows. In Maekawa [2013, 2014] it was shown that the Prandtl approximation is valid if the initial vorticity is compactly supported away from the boundary. More precisely, the author assumes that the Navier-Stokes and Euler initial velocities are the same and their initial curl is supported far from the boundary and, additionally, the curl is Sobolev regular. The author establishes local-in-time well-posedness for the Prandtl equations and shows that, in L^∞, the Prandtl approximation is valid. Note that, in particular, the initial velocity is assumed to be analytic in a neighborhood of the boundary, but this analyticity is lost at positive time and the author carefully estimates how.

The mathematical difficulty in treating the Prandtl equations stems from the loss of one derivative in x, which cannot be recovered due to lack of diffusion in the horizontal direction. One realization of this difficulty is the fact that the Prandtl equations have been shown to be linearly ill-posed in Sobolev spaces, see Gérard-Varet and Dormy [2010]. Finite-time blow-up for smooth solutions of Prandtl's equation goes back to E and Engquist [1997], see also Kukavica, Vicol, and F. Wang [2017] for a more physically motivated example.

As we have already observed, the mismatch between the no-slip and non-penetration boundary conditions is largely responsible for the difficulty in studying the vanishing viscosity limit. Assuming settings for which Prandtl's *ansatz*, that the difference between viscous and inviscid flow is confined to a small region near the boundary, holds true then, at small viscosity, this mismatch corresponds to the formation of a thin layer of intense shear near the boundary. In Grenier, Guo, and Nguyen [2015] this was explored by explicitly connecting the validity of the Prandtl asymptotic model to the stability of viscous shear flows. More precisely, the authors of Grenier, Guo, and Nguyen [ibid.] conjecture that shear flows are typically unstable for the Navier-Stokes equations and, therefore, that the Prandtl approximation does not hold in Sobolev spaces; their conjecture is the subject of ongoing work, see Grenier, Guo, and Nguyen [2016].

In the shear layer discussed above, vorticity, the curl of velocity, tends to be very large. If we consider, instead, the infinite Reynolds number limit, then the mismatch between viscous and ideal flow boundary conditions will lead to a *vortex sheet*, understood as an idealization of a thin region of intense shear, forming at the boundary. This was noted more precisely in Kelliher [2008, 2017], where (a variant of) the result below was proved.

A word on notation: $L^2_\sigma(\mathfrak{D})$ refers to divergence-free vector fields whose components are square-integrable. We recall that a vector field in $L^2_\sigma(\mathfrak{D})$ has a well-defined trace of normal component at $\partial\mathfrak{D}$. In addition, if $\mathbf{x} = (x, y)$ then $\mathbf{x}^\perp \equiv (-y, x)$, and $\nabla^\perp \equiv (-\partial_y, \partial_x)$.

Fix $T > 0$ and assume $\mathfrak{D}$ is connected and simply connected.

Proposition 1.2. (See Kelliher [2008, 2017].) *For each* $\nu > 0$, *let* $\mathbf{u}^\nu \in L^2_\sigma(\mathfrak{D}) \cap H^1_0(\mathfrak{D})$. *Assume that there exists* $\mathbf{v} \in L^2_\sigma(\mathfrak{D})$ *such that* $\mathbf{u}^\nu \rightharpoonup \mathbf{v}$ *weakly in* $L^2_\sigma(\mathfrak{D})$ *as* $\nu \to 0$, *and that the trace of* $\mathbf{v}$, *at* $\partial\mathfrak{D}$, *is well-defined. Then, if* $!^\nu \equiv \operatorname{curl} \mathbf{u}^\nu$, *it follows that*

1. *If* $\mathfrak{D} \subset \mathbb{R}^2$, *then* $!^\nu \rightharpoonup \operatorname{curl} \mathbf{v} - (\mathbf{v} \cdot \tau)\mu$, *weak-$*$* $(H^1(\mathfrak{D}))^*$, *as* $\nu \to 0$, *where* $\tau = \mathbf{n}^\perp$ *and* μ *is the* 1*-dimensional Hausdorff measure on* $\partial\mathfrak{D}$.

2. *If* $\mathfrak{D} \subset \mathbb{R}^3$, *then* $!^\nu \rightharpoonup \operatorname{curl} \mathbf{v} + (\mathbf{v} \times \mathbf{n})\mu$, *weak-$*$* $(H^1(\mathfrak{D}))^*$, *as* $\nu \to 0$, *where* μ *is the* 2*-dimensional Hausdorff measure on* $\partial\mathfrak{D}$.

This result can be derived from Stokes' theorem in a straightforward manner.

Proposition 1.2 implies that, if $\mathbf{u}^\nu \rightharpoonup \mathbf{u}^E$, with $\mathbf{u}^E$ a solution of the Euler equations, and if the trace of tangential component of $\mathbf{u}^E$ is well-defined on $\partial\mathfrak{D}$, then a vortex sheet will form on the boundary, with strength given by $\mathbf{u}^E \cdot \tau$.

2 Vortex sheets in ideal fluid flow

As we have seen, in the vanishing viscosity limit, thin shear layers arise near the boundary and are expected be unstable, detach and affect the bulk of the fluid. In these thin shear layers vorticity is very intense and concentrated, and it is the nature of ideal flow that vorticity can neither be created nor destroyed. An idealization of these thin shear layers are the "Helmholtz surfaces of discontinuity" (see Prandtl [1905]), also known as vortex sheets, which are surfaces across which the velocity field has a tangential discontinuity, while the normal component is continuous. The tangential discontinuity of the flow is read at the level of vorticity as a Dirac delta supported on the surface of discontinuity. In this section we will discuss the evolution of vortex sheets in ideal fluids. We emphasize that, at this point, the problems are purely mathematical, as no real-world fluid has zero viscosity, nor do vortex sheets exist in nature.

2.1 Full plane. We will continue to carry out our discussion in two dimensions. We introduce the vorticity $\omega \equiv \operatorname{curl}\mathbf{u} = \nabla^\perp \cdot \mathbf{u}$ as a new dynamical variable. Taking the curl $\equiv \nabla^\perp\cdot$ of the Euler equations yields:

$$\text{(2-1)} \qquad \begin{cases} \partial_t\omega + (\mathbf{u}\cdot\nabla)\omega = 0, & \text{in } (0,+\infty)\times\mathfrak{D}, \\ \operatorname{div}\mathbf{u} = 0, & \text{in } [0,+\infty)\times\mathfrak{D}, \\ \operatorname{curl}\mathbf{u} \equiv \nabla^\perp\cdot\mathbf{u} = \omega, & \text{in } [0,+\infty)\times\mathfrak{D}. \end{cases}$$

The study of ideal fluid flow from the point-of-view of the evolution of vorticity is called *vortex dynamics*. It is a particularly useful stance given that vorticity is transported. This is true even in three dimensional space, although the transport is more complicated. We note that, if $\mathfrak{D}$ is a simply connected domain, then (2-1) is actually a closed evolution equation for the dynamic variable ω. In this case, the *elliptic system* formed by the last two equations in (2-1) is explicitly solvable:

$$\text{(2-2)} \qquad \mathbf{u}(t,\cdot) = \nabla^\perp(\Delta_0^{\mathfrak{D}})^{-1}\omega(t,\cdot) \equiv K^{\mathfrak{D}}[\omega(t,\cdot)],$$

where $\Delta_0^{\mathfrak{D}}$ is the homogeneous Dirichlet Laplacian on $\mathfrak{D}$. This is called the *Biot-Savart law* and the kernel in the integral operator $K^{\mathfrak{D}}[\cdot]$ is the *Biot-Savart kernel* $K^{\mathfrak{D}} = K^{\mathfrak{D}}(\mathbf{x},\mathbf{y})$.

Let $\mathbf{u}_0$ be a vector field which is irrotational on either side of a given, smooth, curve $\mathcal{C}_0 \subset \mathbb{R}^2$ and such that there is a jump in the tangential component across $\mathcal{C}_0$. It is easy to see, in this case, that the vorticity $\omega_0 = \mathrm{curl}\,\mathbf{u}_0$ is a measure concentrated on the curve $\mathcal{C}_0$, with density, or *vortex sheet strength* $\gamma_0 = [\mathbf{u}_0]_{\mathcal{C}_0}$, that is, $\omega_0 = \gamma_0 \delta_{\mathcal{C}_0}$.

There are two points-of-view used to describe the evolution of a vortex sheet. One gives rise to the *explicit description*, where we seek a time-dependent parameterization of the curve of discontinuity of the flow. To begin with, we parameterize the initial curve and assume that the vortex sheet structure is preserved under the flow, so that $\omega(t,\cdot) = \gamma_t \delta_{\mathcal{C}_t}$, a reasonable assumption given that vorticity is transported by $\mathbf{u} = \mathbf{u}(t,\cdot)$. This *ansatz* leads to the Birkhoff–Rott equations, derived explicitly by G. Birkhoff [1962], implicit in the work of N. Rott [1956]. Using the identification $\mathbb{R}^2 \sim \mathbb{C}$, $z = x + iy$, the Birkhoff–Rott equations are written as

$$\frac{\partial}{\partial t}\overline{z}(t,\Gamma) = \frac{1}{2\pi i}\, pv \int \frac{1}{z(t,\Gamma) - z(t,\Gamma')}\, d\Gamma'. \tag{2-3}$$

The parameter Γ is called the *circulation variable*.

The Birkhoff–Rott equations encode both the motion of the sheet and the time-dependent evolution of the density, or sheet strength, or yet, the magnitude of the jump in tangential velocity across the sheet. The density can be recovered through $\gamma_t(\cdot) = (\partial_\Gamma z(t,\cdot))^{-1}$.

The study of vortex sheet motion through the Birkhoff–Rott equations has a long history. As an idealization of intense thin shear layers, it is expected that vortex sheets develop a complicated motion through spontaneous generation of small scales. This can be illustrated by performing a periodic perturbation on a stationary flat vortex sheet and observing the exponentially growing modes that ensue, see Marchioro and Pulvirenti [1994]. The linear instability observed in the Birkhoff–Rott equations is called *Kelvin-Helmholtz instability* and it manifests itself macroscopically as a tendency of the sheet to roll-up into spirals.

In C. Sulem, P.-L. Sulem, Bardos, and Frisch [1981] short time existence was established assuming the initial vortex sheet and sheet strength were real analytic, since real analyticity implies exponential decay of high Fourier modes. In Moore [1978] and Moore [1979] sophisticated asymptotic calculations were performed which suggested the appearance of a singularity in finite time for analytic vortex sheets and, moreover, he described the expected singularity as a blow-up in curvature. Moore's calculations were rigorously confirmed by Caflisch and Orellana [1986, 1989], who showed that, for an analytic perturbation of amplitude $\mathcal{O}(\varepsilon)$ the time-of-existence is $\mathcal{O}(\log \varepsilon)$. For further work see Duchon and Robert [1988] and Lebeau [2002]. The state-of-the-art result regarding ill-posedness is due to S. Wu [2002, 2006].

An alternative point-of-view in the description of vortex sheet evolution is to embed the discontinuity curve and density in a solution of the Euler equations, whose evolution should carry the information along. In this *implicit description* we make no assumption on the structure of the solution at future time. The tools used when taking this approach are

PDE methods, and the relevant information is the regularity space in which the equations are studied. From this standpoint a Dirac delta or a general bounded Radon measure are indistinguishable. We refer to initial velocities whose curl is a bounded Radon measure as *vortex sheet initial data*.

Let us proceed with a precise definition of a weak solution in a general fluid domain $\mathfrak{D}$.

Definition 2.1. Let $\mathbf{u}_0 \in L^2_{\text{loc}}(\mathfrak{D})$ and assume that $\omega_0 = \operatorname{curl} \mathbf{u}_0 \in \mathfrak{BM}(\mathfrak{D})$. The vector field $\mathbf{u} = \mathbf{u}(t, \mathbf{x}) \in L^\infty_{\text{loc}}(\mathbb{R}_+; L^2_{\text{loc}}(\mathfrak{D}))$ is a *weak solution* of the incompressible Euler equations in $\mathfrak{D}$, (1-1) with $\nu = 0$, with initial data $\mathbf{u}_0$, if the following conditions hold.

1. For every divergence-free test vector field $\Phi = \Phi(t, \mathbf{x}) \in C^\infty_c(\mathbb{R}_+ \times \mathfrak{D})$ the identity below holds true:

$$\int_0^{+\infty} \int_{\mathfrak{D}} \{\partial_t \Phi \cdot \mathbf{u} + [(\mathbf{u} \cdot \nabla)\Phi] \cdot \mathbf{u}\} \, \mathrm{d}\mathbf{x}\mathrm{d}t + \int_{\mathfrak{D}} \Phi(0, \mathbf{x}) \cdot \mathbf{u}_0(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 0, \tag{2-4}$$

2. $\operatorname{div} \mathbf{u}(t, \cdot) = 0$ in $\mathfrak{D}'(\mathbb{R}_+ \times \mathfrak{D})$,

3. if $\partial\mathfrak{D} \neq \varnothing$ then $\mathbf{u} \cdot \mathbf{n} = 0$ in the trace sense on $\partial\mathfrak{D}$, a.e. t.

The study of weak solutions of the Euler equations with vortex sheet initial data was pioneered by R. DiPerna and A. Majda in a series of papers, see DiPerna and Majda [1987a,b, 1988], where they developed the framework and criteria to establish existence. J.-M. Delort [1991], proved the existence of a weak solution with vortex sheet initial data provided the vorticity has a *distinguished sign*. Let us briefly recall Delort's result, for $\mathfrak{D} = \mathbb{R}^2$, and revisit the proof.

Theorem 2.2. (Delort [ibid.]) *Let* $\mathbf{u}_0 \in L^2_{\text{loc}}(\mathbb{R}^2)$ *and assume that* $\omega_0 = \operatorname{curl} \mathbf{u}_0 \in \mathfrak{BM}_{c,+}(\mathbb{R}^2)$. *Then there exists a weak solution in the sense of Definition 2.1.*

The proof is obtained by means of a compensated compactness argument. We discuss an alternative proof, given by Schochet [1995], which involves rewriting the weak formulation in terms of vorticity and then symmetrizing the integration kernels which arise. This has turned out to be a very useful technique, the source of a number of additional results.

The weak vorticity formulation is obtained by first multiplying, formally, the vorticity equation (2-1) by a test function $\varphi \in C^\infty_c(\mathbb{R}_+ \times \mathbb{R}^2)$, then integrating by parts so as to throw all derivatives onto the test function, thus finding

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}^2} [(\partial_t \varphi)\omega + (\mathbf{u} \cdot \nabla\varphi)\omega] \, \mathrm{d}\mathbf{x}\mathrm{d}t + \int_{\mathbb{R}^2} \varphi(0, \mathbf{x})\omega_0(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 0.$$

One then recalls that the velocity $\mathbf{u}$ can be recovered from the vorticity ω by means of the *Biot-Savart* law (2-2) which, in the full plane, reads:

$$\mathbf{u} = \mathbf{u}(t, \mathbf{x}) = \int_{\mathbb{R}^2} K(\mathbf{x} - \mathbf{y}) \omega(t, \mathbf{y}) \, \mathrm{d}\mathbf{y},$$

with

$$K(\mathbf{z}) = \frac{\mathbf{z}^\perp}{2\pi |\mathbf{z}|^2} \equiv \nabla^\perp(\Delta^{-1}).$$

Indeed, since $\operatorname{div} \mathbf{u} = 0$ it follows that $\mathbf{u} = \nabla^\perp \psi$ and, therefore, $\omega = \operatorname{curl} \mathbf{u} = \nabla^\perp \cdot \nabla^\perp \psi = \Delta \psi$. We then substitute $\mathbf{u}$ for the Biot-Savart law in the nonlinear term of the weak vorticity formulation, we symmetrize with respect to $\mathbf{x}$ and $\mathbf{y}$ and use the anti-symmetry of the kernel K to get:

$$\begin{aligned}
\int_{\mathbb{R}_+} \int_{\mathbb{R}^2} (\mathbf{u} \cdot \nabla \varphi) \omega \, \mathrm{d}\mathbf{x} \mathrm{d}t &= \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} K(\mathbf{x} - \mathbf{y}) \omega(t, \mathbf{y}) \, \mathrm{d}\mathbf{y} \right) \cdot \nabla \varphi(t, \mathbf{x}) \omega(t, \mathbf{x}) \, \mathrm{d}\mathbf{x} \mathrm{d}t \\
&= \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} K(\mathbf{x} - \mathbf{y}) \omega(t, \mathbf{x}) \, \mathrm{d}\mathbf{x} \right) \cdot \nabla \varphi(t, \mathbf{y}) \omega(t, \mathbf{y}) \, \mathrm{d}\mathbf{y} \mathrm{d}t \\
&\equiv \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} H_\varphi(t, \mathbf{x}, \mathbf{y}) \omega(t, \mathbf{x}) \omega(t, \mathbf{y}) \, \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y} \mathrm{d}t,
\end{aligned}$$

where the *auxiliary test function* H_φ is given by
(2-5)

$$H_\varphi(t, \mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y}) \cdot \frac{(\nabla \varphi(t, \mathbf{x}) - \nabla \varphi(t, \mathbf{y}))}{2} \equiv \frac{(\mathbf{x} - \mathbf{y})^\perp}{|\mathbf{x} - \mathbf{y}|} \cdot \frac{\nabla \varphi(t, \mathbf{x}) - \nabla \varphi(t, \mathbf{y})}{4\pi |\mathbf{x} - \mathbf{y}|}.$$

With this notation the weak vorticity formulation becomes

$$\begin{aligned}
\text{(2-6)} \qquad \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} (\partial_t \varphi) \omega \, \mathrm{d}\mathbf{x} \mathrm{d}t + \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} H_\varphi(t, \mathbf{x}, \mathbf{y}) \omega(t, \mathbf{x}) \omega(t, \mathbf{y}) \, \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y} \mathrm{d}t \\
+ \int_{\mathbb{R}^2} \varphi(0, \mathbf{x}) \omega_0(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 0.
\end{aligned}$$

The main advantage of this formulation is that it makes sense for any flow whose vorticity is a bounded Radon measure which is *continuous*, that is, which does not contain an atomic part. To see this one first observes that, for any test function $\varphi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R}^2)$, the auxiliary test function H_φ is a *bounded* function on $\mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}^2$. In addition, H_φ is discontinuous *only* on the set $\{(t, \mathbf{x}, \mathbf{y}) \mid \mathbf{x} = \mathbf{y}\}$. Now, if the measure $\omega(t, \cdot)$ is a continuous bounded measure on $\mathbb{R}^2$, then the tensor product $\omega(t, \cdot) \otimes \omega(t, \cdot)$ is a continuous bounded measure on $\mathbb{R}^2 \times \mathbb{R}^2$, which can be integrated against $H_\varphi(t, \cdot, \cdot)$. Moreover, if

$\omega^n \rightharpoonup \omega$ weak-$*$ in $L^\infty(\mathbb{R}_+; \mathfrak{BM}(\mathbb{R}^2))$, and if, uniformly in n and $t \geq 0$, $\omega^n(t,\cdot)$ does not attribute mass to points (no atomic part, uniformly in n), that is, if

$$\sup_n \sup_t \sup_{\mathbf{x}} \int_{B(\mathbf{x};r)} |\omega^n(t,\mathbf{y})|\, d\mathbf{y} \to 0 \text{ as } r \to 0, \tag{2-7}$$

then the weak limit commutes with the nonlinearity.

To prove Delort's theorem one chooses a smooth approximation of the initial data and solves the Euler equations with the smooth initial data. This procedure yields a sequence of exact solutions $(\mathbf{u}^n, \omega^n = \operatorname{curl} \mathbf{u}^n)$ satisfying uniform bounds in $L^\infty_{\text{loc}}(\mathbb{R}_+; L^2_{\text{loc}}(\mathbb{R}^2))$ for velocity and in $L^\infty(\mathbb{R}_+; L^1(\mathbb{R}^2))$ for vorticity. The distinguished sign of vorticity, together with the respective bounds above, yield the condition that $\omega^n(t,\cdot)$ does not attribute mass to points, uniformly in n, in the sense of (2-7). Thus, upon extracting a weakly convergent subsequence, one can pass to the weak limit in all the terms of the weak vorticity formulation (2-6). To conclude the proof one shows that (2-6) is equivalent to items 1. and 2. in Definition 2.1.

We note in passing that it is a well-known fact that sequences of *nonnegative* bounded Radon measures, uniformly bounded in H^{-1}_{loc}, satisfy (2-7).

Delort obtained his existence result for an initial velocity in $L^2_{\text{loc}}(\mathbb{R}^2)$ whose curl belongs to $L^p_c(\mathbb{R}^2) + \mathfrak{BM}_{c,+}(\mathbb{R}^2)$, for some $1 < p \leq \infty$, that is, whose vorticity is such that only the singular part is of distinguished sign. In Vecchi and S. Wu [1993] existence was extended to $p = 1$ using the Dunford-Pettis theorem for uniformly integrable functions.

Versions of Delort's theorem have been obtained for other approximations, such as solutions of the Navier-Stokes equations in the full plane, see Majda [1993], numerical approximations using the vortex blob method, see Liu and Xin [1995], approximations generated by truncation, see Lions [1996] and by central difference schemes, see Lopes Filho, Nussenzveig Lopes, and Tadmor [2000]. An alternative proof, highlighting the compensated compactness aspect of the result, was given in Evans and Müller [1994].

It should be noted that a comparison between the Birkhoff–Rott (explicit) and weak Euler (implicit) mathematical models for the evolution of vortex sheets was carried out in Lopes Filho, Nussenzveig Lopes, and Schochet [2007], where it was shown that a weak solution of the Euler equations whose vorticity is a Dirac delta on a curve of finite length $\mathcal{C}_t$ and with density γ_t is a solution of the Birkhoff–Rott equations if and only if the density is integrable along the curve $\mathcal{C}_t$. This establishes a restricted equivalence between the two descriptions.

2.2 Domains with boundary. In view of the focus of this paper, it is necessary to consider how vortex sheet initial data flow interacts with a rigid boundary. Delort [1991] studied flows with vortex sheet initial data of distinguished sign in bounded domains with smooth boundary, and he established existence of a weak solution in much the same way as for the full plane. Delort explored the fact that his proof, for the full plane, was *local*, since the test vector fields were assumed to have compact support; this made it possible to use the same proof for a bounded domain. However, ideal flows in bounded domains must satisfy the non-penetration condition $\mathbf{u} \cdot \mathbf{n} = 0$ on $\partial\mathfrak{D}$, see (1-3). This is a *linear* condition for velocity (understood in the trace sense), which is trivially continuous with respect to weak convergence in $L^\infty(\mathbb{R}_+; L^2_\sigma(\mathfrak{D}))$. Hence, using the same strategy as for the full plane, namely smoothing out initial data and exactly solving the Euler equations, with the smooth data, *in the bounded domain* $\mathfrak{D}$, yields a sequence of approximations $\mathbf{u}^n$ which converge weak-$*$ in $L^\infty(\mathbb{R}_+; L^2(\mathfrak{D}))$, which are all tangent to $\mathfrak{D}$, so that the weak limit satisfies the non-penetration condition in the trace sense.

The simplest two-dimensional fluid domain with a boundary is the half-plane, $\mathbb{H} \equiv \mathbb{R}^2_+ = \{\mathbf{x} = (x, y) \,|\, y > 0\}$. If we use the same arguments as Delort, but for the (unbounded) half-plane, we also obtain a weak solution with vortex sheet initial data assuming the (singular part of the) initial vorticity is of distinguished sign. Now, the *image method* is well-known in fluid dynamics as a means to extend to the full plane an ideal fluid flow in a half-plane: one simply reflects the half-plane flow, by mirror-symmetry, with respect to the boundary of the half-plane. In $\mathbb{H}$ this means that the first component of the velocity field is even with respect to y, and the second is odd. This symmetry induces the corresponding vorticity to be odd with respect to y, so that, necessarily, it must change sign. If, however, one attempts to use the image method for the weak solution obtained using Delort's proof, one *does not* find a weak solution in the full plane, with this odd vorticity. The reason is that the non-penetration condition along the boundary of $\mathbb{H}$, $y = 0$, is assumed only in the trace sense, and this is too weak for the image method to work. This observation is at the heart of the main result in Lopes Filho, Nussenzveig Lopes, and Xin [2001], where Delort's theorem is extended to include flows whose vorticity is an odd bounded Radon measure, single-signed on each side of a line, plus an arbitrary L^p_c function, $p \geq 1$.

Theorem 2.3. (See Lopes Filho, Nussenzveig Lopes, and Xin [ibid.].) *Let* $\mathbf{u}_0 \in L^2(\mathbb{R}^2)$ *be a divergence-free vector field such that* $\omega_0 = \operatorname{curl} \mathbf{u}_0 \in \mathfrak{BM}_c(\mathbb{R}^2) + L^p_c(\mathbb{R}^2)$*, for some* $p \geq 1$*. Assume that* ω_0 *is odd with respect to* y *and single-signed on* $\mathbb{H}$*. Then there exists a weak solution of the incompressible 2D Euler equations, in the sense of Definition 2.1, with initial data* $(\mathbf{u}_0, \omega_0)$*.*

A key point in the proof of Theorem 2.3 is the *a priori* estimate below, estabished in Lopes Filho, Nussenzveig Lopes, and Xin [ibid.]. Fix $T > 0$, $L > 0$. There exists $C = C(T, L, \|\mathbf{u}_0\|_{L^2} \|\omega_0\|_{\mathfrak{BM}}) > 0$ such that, if $\mathbf{n} = (0, -1)$, then

$$\int_0^T \int_{-L}^{L} |\mathbf{u}^n \cdot \mathbf{n}^\perp|^2(t, x, 0)\, \mathrm{d}x \mathrm{d}t \leq C. \tag{2-8}$$

The new estimate (2-8) implies, together with the estimates with which Delort worked, that, for any compact subset $\mathfrak{K} \subset \mathbb{R}^2$,

$$\sup_n \int_0^T \left(\sup_{\mathbf{x}\in\mathfrak{K}} \int_{B(\mathbf{x};r)} |\omega^n(\mathbf{y},t)|\mathrm{d}\mathbf{y} \right) \mathrm{d}t \leq C|\log r|^{-1/2} \to 0 \text{ as } r \to 0.$$

This is a slightly weaker condition than (2-7), but still sufficient to pass to the limit in the nonlinear term. Moreover, (2-8) makes it possible to establish the validity of the image method for weak solutions in the half plane, since it allows one to prove that Delort-type weak solutions actually satisfy a stronger notion of weak solution, called *boundary-coupled weak solution*, for which the non-penetration condition is assumed in a stronger way than simply the trace sense.

Let us now seek a weak vorticity formulation in a general bounded domain $\mathfrak{D}$ with boundary. Assume that $\mathfrak{D}$ has $k \geq 0$ disjoint holes, so that $\partial\mathfrak{D} = \Gamma_0 \cup_{i=1}^k \Gamma_i$, with Γ_0 being the outer boundary and Γ_i, $i = 1, \ldots, k$ being the boundaries of each of the holes. We will need an analogue of the Biot-Savart law, that is, a means of writing the velocity in terms of its curl. In domains with non-trivial topology, in order to recover velocity from vorticity, it is necessary to assign the *circulation* around each hole, γ_i:

$$\gamma_i \equiv \int_{\Gamma_i} \mathbf{u}\cdot\mathbf{n}^\perp \mathrm{d}S, \quad i = 1, \ldots, k. \tag{2-9}$$

If the vector field $\mathbf{u}$ is divergence-free and its curl is a bounded Radon measure then the circulation is well-defined. In fact, it is enough that $\operatorname{div}\mathbf{u} \in \mathfrak{BM}(\mathfrak{D})$ and $\operatorname{curl}\mathbf{u} \in \mathfrak{BM}(\mathfrak{D})$ for the entire tangential component $\mathbf{u}\cdot\mathbf{n}^\perp$ to be a well-defined distribution on $\partial\mathfrak{D}$, see Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [2017]. An analogous fact had already been noted previously by Chen and Frid [2003].

We will also make use of the *harmonic measures* $\mathbf{w}_i$, solutions of the following boundary-value problem:

$$\begin{cases} \Delta\mathbf{w}_i = 0, & \text{in } \mathfrak{D}, \\ \mathbf{w}_i = \delta_{i\ell}, & \text{on } \Gamma_\ell,\ \ell = 1, \ldots, k. \\ \mathbf{w}_i = 0, & \text{on } \Gamma_0. \end{cases}$$

Lastly, let $\mathbf{H}_i$ denote a basis of *harmonic vector fields*, so that each $\mathbf{H}_i$ is divergence-free, curl-free, and the circulation of $\mathbf{H}_i$ around Γ_ℓ is $\delta_{i\ell}$.

With this notation we can express velocity in terms of vorticity and circulations.

Proposition 2.4. (See Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [2017].) *Let $\omega \in \mathfrak{BM}(\mathfrak{D})$ and fix γ_i, $i = 1, \ldots, k$. If $\mathbf{u} \in L^2(\mathfrak{D})$ is a divergence-free vector field such that* $\operatorname{curl}\mathbf{u} = \omega$ *and for which the circulations of* $\mathbf{u}$ *around Γ_i are γ_i, then*

$$\mathbf{u} = K^{\mathfrak{D}}[\omega] + \sum_{i=1}^k \left(\gamma_i + \int_{\mathfrak{D}} \mathbf{w}_i\omega\,\mathrm{d}\mathbf{x} \right) \mathbf{H}_i.$$

It is possible to express the weak velocity formulation in Definition 2.1 in terms of vorticity, using the symmetrization technique. To this end we observe that, up to a constant, there is a one-to-one correspondence between scalar functions $\varphi \in C_c^\infty(\mathbb{R}_+ \times \overline{\mathfrak{D}})$, constant *in a neighborhood* of $\partial\mathfrak{D}$, with possibly *different* constants *on each* Γ_i, and test vector fields $\Phi \in C_c^\infty(\mathbb{R}_+ \times \mathfrak{D})$, divergence-free, given through the map $\varphi \mapsto \Phi = \nabla^\perp \varphi$.

We introduce:

$$H_\varphi^{\mathfrak{D}} = H_\varphi^{\mathfrak{D}}(t, \mathbf{x}, \mathbf{y}) = \frac{1}{2}\Big(K^{\mathfrak{D}}(\mathbf{x}, \mathbf{y}) \cdot \nabla\varphi(t, \mathbf{x}) + K^{\mathfrak{D}}(\mathbf{y}, \mathbf{x}) \cdot \nabla\varphi(t, \mathbf{y})\Big).$$

Let $\mathfrak{Y} \equiv \{\varphi \in C_c^\infty(\mathbb{R}_+ \times \overline{\mathfrak{D}}) \,|\, \varphi$ is constant *in a neighborhood* of each $\Gamma_i\}$, as in Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [2017].

Proposition 2.5. (See Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [ibid.].) *For all $\varphi \in \mathfrak{Y}$, $H_\varphi^{\mathfrak{D}}$ is bounded on $\mathbb{R}_+ \times \overline{\mathfrak{D}} \times \overline{\mathfrak{D}}$, continuous if $\mathbf{x} \neq \mathbf{y}$ and vanishes on $\mathbb{R}_+ \times \partial(\mathfrak{D} \times \mathfrak{D})$, $\mathbf{x} \neq \mathbf{y}$.*

The main result in Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [ibid.] is:

Theorem 2.6. *The vector field $\mathbf{u} \in L^\infty_{\mathrm{loc}}(\mathbb{R}_+; L^2(\mathfrak{D}))$, such that*

$$\operatorname{curl} \mathbf{u} = \omega \in L^\infty(\mathbb{R}_+; \mathcal{BM}(\mathfrak{D}))$$

and whose circulations around Γ_i are γ_i, $i = 0, \ldots, k$, is a weak solution of the Euler equations, with initial data $\mathbf{u}_0 \in L^2(\mathfrak{D})$, if and only if the following identity holds, for all $\varphi \in \mathfrak{Y}$:

$$\begin{aligned}
\text{(2-10)} \quad & \int_0^\infty \int_{\mathfrak{D}} \partial_t\varphi\, \omega \,\mathrm{d}\mathbf{x}\mathrm{d}t - \int_0^\infty \gamma_0(t)\partial_t\varphi(t,\cdot)\big|_{\Gamma_0} \mathrm{d}t \\
& + \sum_{i=1}^k \int_0^\infty \gamma_i(t)\partial_t\varphi(t,\cdot)\big|_{\Gamma_i} \mathrm{d}t + \int_{\mathfrak{D}} \varphi(0,\cdot)\omega_0 \,\mathrm{d}\mathbf{x} - \gamma_0(0)\varphi(0,\cdot)\big|_{\Gamma_0} \\
& + \sum_{i=1}^k \gamma_i(0)\varphi(0,\cdot)\big|_{\Gamma_i} + \int_0^\infty \int_{\mathfrak{D}} \int_{\mathfrak{D}} H_\varphi^{\mathfrak{D}}(t,\mathbf{x},\mathbf{y})\omega(t,\mathbf{x})\omega(t,\mathbf{y}) \,\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\mathrm{d}t \\
& + \sum_{i=1}^k \int_0^\infty \left(\gamma_i + \int_{\mathfrak{D}} \mathbf{w}_i(\mathbf{y})\omega(t,\mathbf{y}) \,\mathrm{d}\mathbf{y}\right) \int_{\mathfrak{D}} \mathbf{H}_i(\mathbf{x}) \cdot \nabla\varphi(t,\mathbf{x})\omega(t,\mathbf{x}) \,\mathrm{d}\mathbf{x}\mathrm{d}t = 0.
\end{aligned}$$

If $\mathbf{u}_0 \in L^2(\mathfrak{D})$, $\operatorname{curl} \mathbf{u}_0 = \omega_0 \in \mathcal{BM}_+(\mathfrak{D})$ then it follows immediately from Delort's theorem and Theorem 2.6 that identity (2-10) holds true for all $\varphi \in \mathfrak{Y}$.

Now, for smooth flows, circulation around material curves – curves which are transported by the velocity field – is a conserved quantity; this is known as Kelvin's circulation theorem. This includes circulation around the boundaries of the holes, Γ_i. Theorem 2.6 deals with non-smooth flows and highlights that, at this level of regularity, it is possible that circulation *may not be conserved*.

Identity (2-10) is the *weak vorticity formulation* in a bounded domain. It is equivalent to the weak formulation in Definition 2.1, yet it *explicitly* incorporates the possibility of violation of Kelvin's circulation theorem, something not apparent in the weak velocity formulation (2-4).

We introduce a stronger notion of weak solution, adapted from Lopes Filho, Nussenzveig Lopes, and Xin [2001]. Let $\overline{\mathfrak{Y}} = \{\varphi \in C_c^\infty(\mathbb{R}_+ \times \overline{\mathfrak{D}}) \mid \varphi \text{ is constant on each } \Gamma_i\}$.

Definition 2.7. Let $\omega_0 \in \mathfrak{BM}(\mathfrak{D}) \cap H^{-1}(\mathfrak{D})$ and $\gamma_{i,0} \in \mathbb{R}$, $i = 1, \dots, k$. We say that $(\omega, \gamma_1, \dots, \gamma_k)$, $\omega \in L^\infty(\mathbb{R}_+; \mathfrak{BM}(\mathfrak{D}) \cap H^{-1}(\mathfrak{D}))$, $\gamma_i \in L^\infty(\mathbb{R}_+)$, $\gamma_0(\cdot) = \int_{\mathfrak{D}} \omega + \sum_{i=1}^k \gamma_i(\cdot)$, is a *boundary-coupled weak solution* of the Euler equations in $\mathfrak{D}$, with initial data ω_0, $\mathbf{u}_0 = K^{\mathfrak{D}}[\omega_0] + \sum_{i=1}^k \left(\gamma_{i,0} + \int_{\mathfrak{D}} \mathbf{w}_i \omega_0\right) \mathbf{H}_i$, if, for every $\varphi \in \overline{\mathfrak{Y}}$, the identity (2-10) holds true.

Assume that $\mathbf{u}$ is a weak solution which is a weak-$*$ limit, in $L^\infty(\mathbb{R}_+; L^2(\mathfrak{D}))$, of a sequence $\{\mathbf{u}^n\}$ of *exact* smooth solutions with initial data $\{\mathbf{u}_0^n\}$ tending to $\mathbf{u}_0$.

Theorem 2.8. (See Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [2017].) *If* $\omega = \operatorname{curl} \mathbf{u} \in \mathfrak{BM}_{c,+}(\mathfrak{D}) \cap H^{-1}(\mathfrak{D})$ *then*

1. $\gamma_i(t) \geq \gamma_{i,0}$;
2. *If* γ_i *is conserved, for all* $i = 1, \dots, k$, *then the solution is boundary-coupled.*

If $\omega_0 \in L^1(\mathfrak{D}) \cap H^{-1}(\mathfrak{D})$ *then* γ_i *is conserved,* $i = 1, \dots, k$ *and the solution is boundary-coupled.*

The role of the conservation of circulation is to ensure that there is no *vorticity concentration* on $\partial\mathfrak{D}$, that is, that ω^n does not attribute mass to the boundary, uniformly in n, see (2-7).

Boundary-coupled weak solutions have additional interesting properties. The net force exerted by the fluid on the boundary is defined, for smooth solutions, as

$$\int_{\partial\mathfrak{D}} p\mathbf{n} \, \mathrm{d}S,$$

where $p = p(t, \mathbf{x})$ is the scalar pressure. Weak solutions, however, lack sufficient smoothness to have a well-defined net force. A weak formulation of the net force, consistent with the definition for smooth flows, can be shown to be a well-defined object *if and only if* the weak solution is boundary-coupled, see Iftimie, Lopes Filho, Nussenzveig Lopes, and Sueur [ibid.]. The same is true of the net torque.

A weak solution is said to satisfy the weak-strong uniqueness property if, given a strong solution, any weak solution with the same initial data must coincide with it. It is noted in Wiedemann [2017] that boundary-coupled weak solutions who satisfy an energy inequality also satisfy the weak-strong uniqueness property. Moreover, such boundary-coupled weak solutions are, in fact, *dissipative weak solutions*, a notion introduced by P.-L. Lions [1996].

We close this section by observing that, while it is clearly desirable to produce boundary-coupled weak solutions, there is no analogue of (2-8), which was key to prove the existence of boundary-coupled weak solutions in the half-plane, for bounded domains. It remains an open problem whether boundary-coupled weak solutions exist in bounded domains, for vortex sheet initial data, even with a distinguished sign.

3 Vanishing viscosity limit and convergence criteria

In this section we return to the problem of vanishing viscosity, or the infinite Reynolds number limit, in a bounded domain with rigid boundary. Here we are concerned with the *mathematical problem* of whether a vector field which is a limit of vanishing viscosity is, in some sense, a solution of the inviscid equations. Hindsight gathered from our previous discussion suggests that, in view of the mismatch between the no slip (1-2) and non-penetration (1-3) boundary conditions, the key issue is to control the production of vorticity at the boundary as $\nu \to 0$.

Let us assume hereafter that $\mathfrak{D}$ is a bounded, connected and simply connected domain with smooth boundary. We will restrict our discussion to two dimensional fluid flow since, from the point-of-view of rigorous mathematical analysis, the Euler and Navier-Stokes equations are better understood in $2D$. We will point out which, among the results we will discuss, have extensions to $3D$.

We begin by recalling the Lighthill principle, which relates the flux of vorticity through the boundary to the tangential derivative of pressure. To see this we assume that the ν-Navier-Stokes equations are valid *up to* the boundary of the domain and we use the no slip condition to deduce, formally, that

$$0 = \nabla p^\nu + \nu \Delta \mathbf{u}^\nu \text{ on } \mathbb{R}_+ \times \partial\mathfrak{D}.$$

Next note that $\Delta \mathbf{u}^\nu = \nabla^\perp \omega^\nu$ and take the inner product with $\mathbf{n}^\perp$ to deduce that

$$\frac{\partial \omega^\nu}{\partial \mathbf{n}} = -\frac{1}{\nu}\frac{\partial p^\nu}{\partial \mathbf{n}^\perp}, \text{ on } (0, +\infty) \times \partial\mathfrak{D}. \tag{3-1}$$

The vorticity formulation of the ν-Navier-Stokes equations is

$$(3\text{-}2)\qquad \begin{cases} \partial_t \omega^\nu + (\mathbf{u}^\nu \cdot \nabla)\omega^\nu = -\nabla p^\nu + \nu\Delta\omega^\nu, & \text{in } (0,+\infty)\times\mathfrak{D};\\ \operatorname{div}\mathbf{u}^\nu = 0, & \text{in } [0,+\infty)\times\mathfrak{D};\\ \operatorname{curl}\mathbf{u}^\nu = \omega^\nu, & \text{in } [0,+\infty)\times\mathfrak{D}, \end{cases}$$

subject to the boundary condition (3-1), and given an initial data.

The Lighthill principle (3-1) was derived *formally* and it does not appear to be particularly useful in general, since estimates on the tangential derivative of the pressure are not usually available.

For certain flows with symmetry, however, (3-1) proves to be sufficient to establish the vanishing viscosity limit. For two-dimensional flows with circular symmetry, see Matsui [1994], Bona and J. Wu [2002], Lopes Filho, Mazzucato, and Nussenzveig Lopes [2008], and Lopes Filho, Mazzucato, Nussenzveig Lopes, and Taylor [2008]. Flows with plane-parallel symmetry were studied in Mazzucato and Taylor [2008], Mazzucato, Niu, and X. Wang [2011], and Gie, Kelliher, Lopes Filho, Mazzucato, and Nussenzveig Lopes [2017] and paralell-pipe flows were discussed in Mazzucato and Taylor [2011], Han, Mazzucato, Niu, and X. Wang [2012], and Gie, Kelliher, Lopes Filho, Mazzucato, and Nussenzveig Lopes [2017].

In general, if there is no mismatch, that is, if the Euler velocity happens to vanish at the boundary at all times $t > 0$, then trivial energy estimates yield convergence of the Navier-Stokes solutions to the inviscid solution as viscosity vanishes.

Let us discuss the general problem; we are interested in *criteria* for the vanishing viscosity limit to hold, that is, conditions under which the limit of solutions of ν-Navier-Stokes, $\nu \to 0$, are solutions, in some sense, of the Euler equations. The baseline result of this nature is known as the *Kato condition*, which we state below.

Fix $T > 0$.

Theorem 3.1. (Kato [1984].) *Let $\mathbf{u}^\nu \in L^\infty((0,T);L^2_\sigma(\mathfrak{D})) \cap L^2((0,T);H^1_0(\mathfrak{D}))$ be a Leray-Hopf solution of the ν-Navier-Stokes equations in $\mathfrak{D} \subset \mathbb{R}^d$, $d = 2, 3$, $\nu > 0$, with initial data $\mathbf{u}_0 \in L^2_\sigma(\mathfrak{D})$, $\mathbf{u}_0 \cdot \mathbf{n} = 0$ on $\partial\mathfrak{D}$. Assume that there exists a smooth solution $\mathbf{u}^0$ of the Euler equations in $\mathfrak{D}$, satisfying the non-penetration boundary condition, with initial data $\mathbf{u}_0$. Then $\mathbf{u}^\nu \to \mathbf{u}^0$ strongly in $L^\infty((0,T);L^2(\mathfrak{D}))$ if and only if*

$$(3\text{-}3)\qquad \nu\int_0^T\int_{\Gamma^\nu} |D\mathbf{u}^\nu|^2 \,d\mathbf{x} \to 0 \text{ as } \nu \to 0,$$

where Γ^ν is a region near the boundary of thickness $\mathcal{O}(\nu)$.

Remark 3.2. Kato proved the equivalence between vanishing viscosity and several other statements. Among these, he showed that the Kato condition (3-3) holds if and only if

$\mathbf{u}^\nu \rightharpoonup \mathbf{u}^0$ weak-$*$ in $L^\infty_{\text{loc}}(\mathbb{R}_+; L^2(\mathfrak{D}))$, paving the way towards understanding the vanishing viscosity limit as a weak limit.

The proof of the Kato criterion is by energy methods, with the necessity of (3-3) deriving immediately from the energy inequality for Leray-Hopf solutions of Navier-Stokes. To show that (3-3) is sufficient Kato introduced the *Kato correctors*, which are certain cut-off functions near the boundary. He then uses these correctors on the Euler solution, "correcting it" at a distance δ, from the boundary. The error term is estimated using the smoothness of the Euler solution. The choice $\delta = \mathcal{O}(\nu)$ allows to estimate terms involving $D\mathbf{u}^\nu$, which are not *a priori* bounded, provided (3-3) holds true.

The Kato criterion has been revisited by several authors. For instance, in Temam and X. Wang [1997] and X. Wang [2001] the full gradient is replaced by the tangential derivative along the boundary; in Kelliher [2007] the gradient is substituted by vorticity. It should be noted that the "Kato layer" is not a physical boundary layer, only a mathematical device.

The key issue in what follows is the fact that the Kato criterion assumes the underlying Euler flow to be smooth. It has been the main point of these notes that this is not what is expected, typically, in the vanishing viscosity limit. We have argued that, due to the mismatch between no slip and non-penetration boundary conditions, vortex sheets arise naturally in the infinite Reynolds number limit. These structures are idealizations of thin shear layers near the boundary. Experiments and the Prandtl asymptotic boundary layer model suggest that these thin layers are unstable and may detach from the boundary, entraining the bulk of the fluid. This justifies the study of the inviscid problem with vortex sheet initial data. Now, solutions of the Euler equations with vortex sheet initial data are far from smooth. Yet these nonsmooth solutions are precisely what is expected at the vanishing viscosity limit! We conclude this discussion with a different criterion for the vanishing viscosity limit to hold, in the two-dimensional case – one which allows the limiting flow to have vortex sheet regularity.

Let $\mathfrak{D} \subset \mathbb{R}^2$ be a smooth, connected and simply connected, bounded domain.

Theorem 3.3. (See Constantin, Lopes Filho, Nussenzveig Lopes, and Vicol [2017].) *Fix $T > 0$. Let $\{\nu_n\}$ be a sequence of positive real numbers such that $\nu_n \to 0$ and choose $\mathbf{u}_0 \in L^2_\sigma(\mathfrak{D})$, $\mathbf{u}_0 \cdot \mathbf{n} = 0$ on $\partial\mathfrak{D}$.*

Let $\mathbf{u}^n \in L^\infty(0, T; L^2(\mathfrak{D})) \cap L^2(0, T; H^1_0(\mathfrak{D}))$ be a Leray-Hopf solution of the ν_n-Navier-Stokes equations, subject to the no slip boundary condition, and with initial data $\mathbf{u}_0$.

Set $\omega^n = \omega^n(t, \cdot) = \operatorname{curl} \mathbf{u}^n \equiv \nabla^\perp \cdot \mathbf{u}^n(t, \cdot)$.

Suppose, additionally, that:

1. *$\mathbf{u}^n \rightharpoonup \mathbf{u}^\infty$ weak-$*$ in $L^\infty(0, T; L^2(\mathfrak{D}))$;*
2. *$\{\omega^n\}$ is uniformly bounded in $L^\infty(0, T; L^1_{\text{loc}}(\mathfrak{D}))$;*

3. For any $\mathfrak{K} \subset\subset \mathfrak{D}$ we have

$$\sup_n \int_0^T \left(\sup_{\mathbf{x}\in\mathfrak{K}} \int_{B(\mathbf{x};r)} |\omega^n(\mathbf{y},t)| \mathrm{d}\mathbf{y} \right) \mathrm{d}t \to 0 \textit{ as } r \to 0.$$

Then $\mathbf{u}^\infty$ is a weak solution of the Euler equations in $\mathfrak{D}$ with initial data $\mathbf{u}_0$.

Theorem 3.3 is inspired on a result contained in Constantin and Vicol [2017]. Its proof is adapted from Schochet's proof of Delort's theorem, see Schochet [1995].

Assumptions 2. and 3. encode the expected behavior of vortex sheets, yet there is no proof that they hold true along viscous approximations in general. In light of our previous discussion, however, these are very natural hypotheses, in contrast with what is assumed in the Kato criterion.

Items 2. and 3. are strictly *local* hypotheses; nothing is assumed about the behavior near the boundary. Surprisingly, the limit flow $\mathbf{u}^\infty$ does satisfy the non-penetration boundary condition, but only in the trace sense in $L^\infty((0,T); H^{-1/2}(\partial\mathfrak{D}))$; this highlights just how unsatisfactory is the weak formulation. Of course, it is hopeless to obtain a boundary-coupled weak solution in this way.

Acknowledgments. The authors wish to thank Dragoş Iftimie and Vlad Vicol for helpful comments.

References

D. J. Acheson (1990). *Elementary fluid dynamics*. Oxford Applied Mathematics and Computing Science Series. The Clarendon Press, Oxford University Press, New York, pp. x+397. MR: 1069557 (cit. on p. 2538).

R. Alexandre, Y.-G. Wang, C.-J. Xu, and T. Yang (2015). "Well-posedness of the Prandtl equation in Sobolev spaces". *J. Amer. Math. Soc.* 28.3, pp. 745–784. MR: 3327535 (cit. on p. 2541).

Garrett Birkhoff (1962). "Helmholtz and Taylor instability". In: *Proc. Sympos. Appl. Math., Vol. XIII*. American Mathematical Society, Providence, R.I., pp. 55–76. MR: 0137423 (cit. on p. 2544).

Jerry L. Bona and Jiahong Wu (2002). "The zero-viscosity limit of the 2D Navier-Stokes equations". *Stud. Appl. Math.* 109.4, pp. 265–278. MR: 1934651 (cit. on p. 2553).

Russel E. Caflisch and Oscar F. Orellana (1986). "Long time existence for a slightly perturbed vortex sheet". *Comm. Pure Appl. Math.* 39.6, pp. 807–838. MR: 859274 (cit. on p. 2544).

– (1989). "Singular solutions and ill-posedness for the evolution of vortex sheets". *SIAM J. Math. Anal.* 20.2, pp. 293–307. MR: 982661 (cit. on p. 2544).

Gui-Qiang Chen and Hermano Frid (2003). "Extended divergence-measure fields and the Euler equations for gas dynamics". *Comm. Math. Phys.* 236.2, pp. 251–280. MR: 1981992 (cit. on p. 2549).

Peter Constantin, Igor Kukavica, and Vlad Vicol (2015). "On the inviscid limit of the Navier-Stokes equations". *Proc. Amer. Math. Soc.* 143.7, pp. 3075–3090. MR: 3336632 (cit. on p. 2541).

Peter Constantin, Milton C. Lopes Filho, Helena J. Nussenzveig Lopes, and V. Vicol (2017). In preparation (cit. on p. 2554).

Peter Constantin and Vlad Vicol (Aug. 2017). "Remarks on high Reynolds numbers hydrodynamics and the inviscid limit". arXiv: 1708.03225 (cit. on p. 2555).

Jean-Marc Delort (1991). "Existence de nappes de tourbillon en dimension deux". *J. Amer. Math. Soc.* 4.3, pp. 553–586. MR: 1102579 (cit. on pp. 2545, 2547).

Ronald J. DiPerna and Andrew J. Majda (1987a). "Concentrations in regularizations for 2-D incompressible flow". *Comm. Pure Appl. Math.* 40.3, pp. 301–345. MR: 882068 (cit. on p. 2545).

– (1987b). "Oscillations and concentrations in weak solutions of the incompressible fluid equations". *Comm. Math. Phys.* 108.4, pp. 667–689. MR: 877643 (cit. on p. 2545).

– (1988). "Reduced Hausdorff dimension and concentration-cancellation for two-dimensional incompressible flow". *J. Amer. Math. Soc.* 1.1, pp. 59–95. MR: 924702 (cit. on p. 2545).

Jean Duchon and Raoul Robert (1988). "Global vortex sheet solutions of Euler equations in the plane". *J. Differential Equations* 73.2, pp. 215–224. MR: 943940 (cit. on p. 2544).

Weinan E and Bjorn Engquist (1997). "Blowup of solutions of the unsteady Prandtl's equation". *Comm. Pure Appl. Math.* 50.12, pp. 1287–1293. MR: 1476316 (cit. on p. 2542).

L. C. Evans and S. Müller (1994). "Hardy spaces and the two-dimensional Euler equations with nonnegative vorticity". *J. Amer. Math. Soc.* 7.1, pp. 199–219. MR: 1220787 (cit. on p. 2547).

David Gérard-Varet and Emmanuel Dormy (2010). "On the ill-posedness of the Prandtl equation". *J. Amer. Math. Soc.* 23.2, pp. 591–609. MR: 2601044 (cit. on p. 2542).

Gung-Min Gie, James P. Kelliher, Milton C. Lopes Filho, Anna L. Mazzucato, and Helena J. Nussenzveig Lopes (June 2017). "The Vanishing viscosity limit for some symmetric flows". arXiv: 1706.06039 (cit. on p. 2553).

Emmanuel Grenier, Yan Guo, and Toan T. Nguyen (2015). "Spectral stability of Prandtl boundary layers: an overview". *Analysis (Berlin)* 35.4, pp. 343–355. MR: 3420316 (cit. on pp. 2541, 2542).

– (2016). "Spectral instability of characteristic boundary layer flows". *Duke Math. J.* 165.16, pp. 3085–3146. MR: 3566199 (cit. on p. 2542).

Daozhi Han, Anna L. Mazzucato, Dongjuan Niu, and Xiaoming Wang (2012). "Boundary layer for a class of nonlinear pipe flow". *J. Differential Equations* 252.12, pp. 6387–6413. MR: 2911839 (cit. on p. 2553).

Dragoş Iftimie, Milton C. Lopes Filho, Helena J. Nussenzveig Lopes, and Franck Sueur (2017). "Weak vorticity formulation of the incompressible 2d euler equations in bounded domains" (cit. on pp. 2549–2551).

Tosio Kato (1984). "Remarks on zero viscosity limit for nonstationary Navier-Stokes flows with boundary". In: *Seminar on nonlinear partial differential equations (Berkeley, Calif., 1983)*. Vol. 2. Math. Sci. Res. Inst. Publ. Springer, New York, pp. 85–98. MR: 765230 (cit. on p. 2553).

James P. Kelliher (2007). "On Kato's conditions for vanishing viscosity". *Indiana Univ. Math. J.* 56.4, pp. 1711–1721. MR: 2354697 (cit. on p. 2554).

– (2008). "Vanishing viscosity and the accumulation of vorticity on the boundary". *Commun. Math. Sci.* 6.4, pp. 869–880. MR: 2511697 (cit. on p. 2542).

– (2017). "Observations on the vanishing viscosity limit". *Trans. Amer. Math. Soc.* 369.3, pp. 2003–2027. MR: 3581225 (cit. on pp. 2541, 2542).

Igor Kukavica, Nader Masmoudi, Vlad Vicol, and Tak Kwong Wong (2014). "On the local well-posedness of the Prandtl and hydrostatic Euler equations with multiple monotonicity regions". *SIAM J. Math. Anal.* 46.6, pp. 3865–3890. MR: 3284569 (cit. on p. 2541).

Igor Kukavica and Vlad Vicol (2013). "On the local existence of analytic solutions to the Prandtl boundary layer equations". *Commun. Math. Sci.* 11.1, pp. 269–292. MR: 2975371 (cit. on p. 2541).

Igor Kukavica, Vlad Vicol, and Fei Wang (2017). "The van Dommelen and Shen singularity in the Prandtl equations". *Adv. Math.* 307, pp. 288–311. MR: 3590519 (cit. on p. 2542).

Gilles Lebeau (2002). "Régularité du problème de Kelvin-Helmholtz pour l'équation d'Euler 2d". *ESAIM Control Optim. Calc. Var.* 8. A tribute to J. L. Lions, pp. 801–825. MR: 1932974 (cit. on p. 2544).

Pierre-Louis Lions (1996). *Mathematical topics in fluid mechanics. Vol. 1*. Vol. 3. Oxford Lecture Series in Mathematics and its Applications. Incompressible models, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, pp. xiv+237. MR: 1422251 (cit. on pp. 2547, 2552).

Jian-Guo Liu and Zhou Ping Xin (1995). "Convergence of vortex methods for weak solutions to the 2-D Euler equations with vortex sheet data". *Comm. Pure Appl. Math.* 48.6, pp. 611–628. MR: 1338471 (cit. on p. 2547).

Maria Carmela Lombardo, Marco Cannone, and Marco Sammartino (2003). "Well-posedness of the boundary layer equations". *SIAM J. Math. Anal.* 35.4, pp. 987–1004. MR: 2049030 (cit. on p. 2541).

Milton C. Lopes Filho, Anna L. Mazzucato, and Helena J. Nussenzveig Lopes (2008). "Vanishing viscosity limit for incompressible flow inside a rotating circle". *Phys. D* 237.10-12, pp. 1324–1333. MR: 2454590 (cit. on p. 2553).

Milton C. Lopes Filho, Anna L. Mazzucato, Helena J. Nussenzveig Lopes, and Michael Taylor (2008). "Vanishing viscosity limits and boundary layers for circularly symmetric 2D flows". *Bull. Braz. Math. Soc. (N.S.)* 39.4, pp. 471–513. MR: 2465261 (cit. on p. 2553).

Milton C. Lopes Filho, Helena J. Nussenzveig Lopes, and Steven Schochet (2007). "A criterion for the equivalence of the Birkhoff-Rott and Euler descriptions of vortex sheet evolution". *Trans. Amer. Math. Soc.* 359.9, pp. 4125–4142. MR: 2309179 (cit. on p. 2547).

Milton C. Lopes Filho, Helena J. Nussenzveig Lopes, and Eitan Tadmor (2000). "Approximate solutions of the incompressible Euler equations with no concentrations". *Ann. Inst. H. Poincaré Anal. Non Linéaire* 17.3, pp. 371–412. MR: 1771138 (cit. on p. 2547).

Milton C. Lopes Filho, Helena J. Nussenzveig Lopes, and Zhou Ping Xin (2001). "Existence of vortex sheets with reflection symmetry in two space dimensions". *Arch. Ration. Mech. Anal.* 158.3, pp. 235–257. MR: 1842346 (cit. on pp. 2548, 2551).

Yasunori Maekawa (2013). "Solution formula for the vorticity equations in the half plane with application to high vorticity creation at zero viscosity limit". *Adv. Differential Equations* 18.1-2, pp. 101–146. MR: 3052712 (cit. on p. 2541).

– (2014). "On the inviscid limit problem of the vorticity equations for viscous incompressible flows in the half-plane". *Comm. Pure Appl. Math.* 67.7, pp. 1045–1128. MR: 3207194 (cit. on p. 2541).

Andrew J. Majda (1993). "Remarks on weak solutions for vortex sheets with a distinguished sign". *Indiana Univ. Math. J.* 42.3, pp. 921–939. MR: 1254126 (cit. on p. 2547).

Carlo Marchioro and Mario Pulvirenti (1994). *Mathematical theory of incompressible nonviscous fluids*. Vol. 96. Applied Mathematical Sciences. Springer-Verlag, New York, pp. xii+283. MR: 1245492 (cit. on p. 2544).

Nader Masmoudi and Tak Kwong Wong (2015). "Local-in-time existence and uniqueness of solutions to the Prandtl equations by energy methods". *Comm. Pure Appl. Math.* 68.10, pp. 1683–1741. MR: 3385340 (cit. on p. 2541).

Shin'ya Matsui (1994). "Example of zero viscosity limit for two-dimensional nonstationary Navier-Stokes flows with boundary". *Japan J. Indust. Appl. Math.* 11.1, pp. 155–170. MR: 1266526 (cit. on p. 2553).

Anna L. Mazzucato, Dongjuan Niu, and Xiaoming Wang (2011). "Boundary layer associated with a class of 3D nonlinear plane parallel channel flows". *Indiana Univ. Math. J.* 60.4, pp. 1113–1136. MR: 2975337 (cit. on p. 2553).

Anna L. Mazzucato and Michael Taylor (2008). "Vanishing viscosity plane parallel channel flow and related singular perturbation problems". *Anal. PDE* 1.1, pp. 35–93. MR: 2431354 (cit. on p. 2553).

– (2011). "Vanishing viscosity limits for a class of circular pipe flows". *Comm. Partial Differential Equations* 36.2, pp. 328–361. MR: 2763344 (cit. on p. 2553).

D. W. Moore (1978). "The equation of motion of a vortex layer of small thickness". *Studies in Appl. Math.* 58.2, pp. 119–140. MR: 0475268 (cit. on p. 2544).

– (1979). "The spontaneous appearance of a singularity in the shape of an evolving vortex sheet". *Proc. Roy. Soc. London Ser. A* 365.1720, pp. 105–119. MR: 527594 (cit. on p. 2544).

O. A. Oleĭnik (1963). "On the system of Prandtl equations in boundary-layer theory". *Dokl. Akad. Nauk SSSR* 150, pp. 28–31. MR: 0153979 (cit. on p. 2541).

– (1966). "On the mathematical theory of boundary layer for an unsteady flow of incompressible fluid". *J. Appl. Math. Mech.* 30, 951–974 (1967). MR: 0223134 (cit. on p. 2541).

O. A. Oleĭnik and V. N. Samokhin (1999). *Mathematical models in boundary layer theory*. Vol. 15. Applied Mathematics and Mathematical Computation. Chapman & Hall/CRC, Boca Raton, FL, pp. x+516. MR: 1697762 (cit. on p. 2541).

Ludwig Prandtl (1905). "Über flüssigkeitsbewegung bei sehr kleiner reibung". In: *Intern. Math. Kongr. Heidelberg 1904*. Vol. III. Leipzig: Teubner, pp. 484–491 (cit. on pp. 2538, 2543).

Nicholas Rott (1956). "Diffraction of a weak shock with vortex generation". *J. Fluid Mech.* 1, 111–128 (1 plate). MR: 0082327 (cit. on p. 2544).

Marco Sammartino and Russel E. Caflisch (1998a). "Zero viscosity limit for analytic solutions of the Navier-Stokes equation on a half-space. II. Construction of the Navier-Stokes solution". *Comm. Math. Phys.* 192.2, pp. 463–491. MR: 1617538 (cit. on p. 2541).

– (1998b). "Zero viscosity limit for analytic solutions, of the Navier-Stokes equation on a half-space. I. Existence for Euler and Prandtl equations". *Comm. Math. Phys.* 192.2, pp. 433–461. MR: 1617542 (cit. on p. 2541).

Hermann Schlichting (1960). *Boundary layer theory*. Translated by J. Kestin. 4th ed. McGraw-Hill Series in Mechanical Engineering. McGraw-Hill Book Co., Inc., New York-Toronto-London; Verlag G. Braun, Karlsruhe, pp. xx+647. MR: 0122222 (cit. on p. 2538).

Steven Schochet (1995). "The weak vorticity formulation of the 2-D Euler equations and concentration-cancellation". *Comm. Partial Diff. Equations* 20.5-6, pp. 1077–1104. MR: 1326916 (cit. on pp. 2545, 2555).

C. Sulem, P.-L. Sulem, C. Bardos, and U. Frisch (1981). "Finite time analyticity for the two- and three-dimensional Kelvin-Helmholtz instability". *Comm. Math. Phys.* 80.4, pp. 485–516. MR: 628507 (cit. on p. 2544).

Roger Temam and Xiaoming Wang (1997). "On the behavior of the solutions of the Navier-Stokes equations at vanishing viscosity". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 25.3-4. Dedicated to Ennio De Giorgi, 807–828 (1998). MR: 1655543 (cit. on p. 2554).

Italo Vecchi and Sijue Wu (1993). "On L^1-vorticity for 2-D incompressible flow". *Manuscripta Math.* 78.4, pp. 403–412. MR: 1208650 (cit. on p. 2547).

Xiaoming Wang (2001). "A Kato type theorem on zero viscosity limit of Navier-Stokes flows". *Indiana Univ. Math. J.* 50.Special Issue. Dedicated to Professors Ciprian Foias and Roger Temam (Bloomington, IN, 2000), pp. 223–241. MR: 1855670 (cit. on p. 2554).

Emil Wiedemann (May 2017). "Weak-strong uniqueness in fluid dynamics". arXiv: 1705.04220 (cit. on p. 2552).

Sijue Wu (2002). "Recent progress in mathematical analysis of vortex sheets". In: *Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002)*. Higher Ed. Press, Beijing, pp. 233–242. MR: 1957535 (cit. on p. 2544).

– (2006). "Mathematical analysis of vortex sheets". *Comm. Pure Appl. Math.* 59.8, pp. 1065–1206. MR: 2230845 (cit. on p. 2544).

Received 2017-11-30.

Helena J. Nussenzveig Lopes
Instituto de Matemática
Universidade Federal do Rio de Janeiro
Caixa Postal 68530
21941-909, Rio de Janeiro, RJ Brazil
hlopes@im.ufrj.br

Milton C. Lopes Filho
Instituto de Matemática
Universidade Federal do Rio de Janeiro
Caixa Postal 68530
21941-909, Rio de Janeiro, RJ Brazil
mlopes@im.ufrj.br

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2559–2584)

THE TEICHMÜLLER TQFT

Jørgen Ellegaard Andersen and Rinat Kashaev

Abstract

We review our construction of the Teichmüller TQFT. We recall our volume conjecture for this TQFT and the examples for which this conjecture has been established. We end the paper with a brief review of our new formulation of the Teichmüller TQFT together with some anticipated future developments.

1 Introduction

Topological Quantum Field Theories (TQFT's) were discovered and axiomatised by Atiyah [1988], Segal [1988] and Witten [1988]. Following Witten's suggestions in Witten [ibid.], the first examples in $2+1$ dimensions were constructed by Reshetikhin and Turaev [1990, 1991] and Turaev [1994] based on the representation theory of quantum groups at roots of unity. The resulting *Witten–Reshetikhin–Turaev* TQFT (WRT-TQFT) has also been formulated in pure topological terms in Blanchet, Habegger, Masbaum, and Vogel [1992, 1995] and it was conjectured in Witten [1988] to be related to quantum conformal field theory and geometric quantization of moduli spaces. It has been further developed in Tsuchiya, Ueno, and Yamada [1989], Axelrod, Della Pietra, and Witten [1991], and Hitchin [1990] and in a series of papers including the work of Laszlo who proved that the Hitchin and the TUY connections agree in the closed surface case in Laszlo [1998]. The equivalence of the geometric and combinatorial constructions has been finally verified in Andersen and Ueno [2007a,b, 2012, 2015] by the first author of this paper jointly with Ueno and exploited in Andersen [2013, 2006, 2008, 2010], Andersen and Himpel [2012], Andersen and Jørgensen [2015], and Andersen, Himpel, Jørgensen, Martens, and McLellan [2017] to establish some strong properties of the WRT-TQFT.

In parallel to the surgery based construction of the WRT-TQFT by Reshetikhin and Turaev, there is the Turaev–Viro construction of the TV-TQFT in Turaev and Viro [1992],

Supported in part by the center of excellence grant "Center for quantum geometry of Moduli Spaces" DNRF95, from the Danish National Research Foundation, Swiss National Science Foundation and the ESF-ITGP network.
MSC2010: primary 57M27; secondary 81T45.

which is also a combinatorial construction, but it uses triangulations instead, and where Reshetikhin and Turaev had to prove the invariance under the surgery presentation of their construction, Turaev and Viro proved invariance under the Pachner 2-3, 3-2, 4-1 and 1-4 moves of theirs. It turned out that the TV-TQFT was the Hermitian endomorphism theory of the WRT-TQFT.

Witten further proposed that quantum Chern–Simon theory for non-compact groups should also exist as generalised TQFT's with underlying infinite dimensional state vector spaces in Witten [1991]. A series of papers on this subject have subsequently emerged in the physics literature, including Bar-Natan and Witten [1991], Dimofte [2013], Dimofte, Gaiotto, and Gukov [2014], Dimofte, Gukov, Lenells, and Zagier [2009], Dimofte and Gukov [2013], Gukov and Murakami [2008], Gukov [2005], Hikami [2001, 2007], and Witten [2011]. However, a mathematical definition of these theories has been lacking for a long time.

In a series of papers Andersen and Kashaev [2014a,c, 2015, 2013, 2014b], the authors of this paper, have provided a rigorous construction of such a TQFT, known as the *Teichmüller TQFT*. Our construction uses combinatorics of Δ-complexes with fixed number of vertices which we call triangulations and it builds on quantum Teichmüller theory, as developed by Kashaev [1998], and Fock and Chekhov [1999], which produces unitary representations of centrally extended mappings class groups of punctured surfaces in infinite-dimensional Hilbert spaces. In this paper we shall first review our original formulation presented in Andersen and Kashaev [2014a,c, 2015].

The central ingredients in quantum Teichmüller theory are, on the one hand, Penner's coordinates of the decorated Teichmüller space and the Ptolemy groupoid introduced in Penner [1987] with applications summarised in Penner [2012] and, on the other hand, Faddeev's quantum dilogarithm presented in Faddeev [1995] which finds its origins and applications in quantum integrable systems Faddeev, Kashaev, and Volkov [2001], Bazhanov, Mangazeev, and Sergeev [2007, 2008], and Teschner [2007]. Faddeev's quantum dilogarithm has already been used in formal state-integral constructions of perturbative invariants of three manifolds in the works Hikami [2001, 2007], Dimofte, Gukov, Lenells, and Zagier [2009], Dijkgraaf, Fuji, and Manabe [2011], and Dimofte [2013], but without addressing the important questions of convergence or triangulation independence.

There are further ingredients which we had to introduce in Andersen and Kashaev [2014a] in order to lift quantum Teichmüller theory to a TQFT. The important one is the weight function for tetrahedra, whose edges are labeled by dihedral angles of hyperbolic ideal tetrahedra. It is not immediately clear what are the topological invariance properties of our TQFT which depends on those dihedral angles. It turns out, however, the partition function of a given triangulation is invariant under certain Hamiltonian gauge group action in the space of angles so that the corresponding symplectically reduced space is determined by the total dihedral angles around edges and the first cohomology group of the (cusp) boundary. As a consequence, under the condition that the triangulation in question is such that the second homology group of the complement of the vertices is trivial, the partition function descends to a well defined function on an open convex subset of this reduced angle space (corresponding to strictly positive angles). Furthermore, if we have two triangulations admitting angle structures (which correspond to balanced edges with total dihedral angles equal to 2π) and related by a Pachner 2-3 or 3-2 move, then the two convex subsets intersect non-trivially and the two partitions functions agree on the overlap. The additional fact that the partition functions depend analytically on the dihedral angles implies that their common restriction to the overlap completely determines both of them and it is in this sense that our TQFT is topologically invariant.

The partition functions of our TQFT take their values in the vector spaces of tempered distributions over euclidian spaces which do not form a category, since it is not always possible to multiply and push forward tempered distributions. Instead, they form what we call a categroid that is the same as a category, except that we are allowed to compose not all morphisms which are composable in the categorical sense, but only a subset thereof (which we review in Section 3). Symmetrically, the domain of our TQFT, the set of oriented triangulated pseudo 3-manifolds, also forms only a categroid, due to the above mentioned homological condition on triangulations.

We shall further review a version of the volume conjecture for the Teichmüller TQFT, which states that the partition function *decays* exponentially fast in Planck's constant with the rate given by the hyperbolic volume of the manifold.

Interestingly, due to subsequent developments, we have now at least two formulations of the Teichmüller TQFT. The original formulation, which is defined only for admissible pseudo 3-manifolds (see Definition 3 below), and the new formulation, which does not impose any restrictions on the topology of pseudo 3-manifolds. We will briefly discuss the new formulation in the end of this paper, together with a number of future developments which we anticipate.

The paper is organised as follows. In Section 2, we review the domain categroid, on which our original formulation of the Teichmüller TQFT is defined, while the target categroid is reviewed in Section 3. In Section 4, we review our TQFT functor between these two categroids and state the main Theorem 3, proved in Andersen and Kashaev [ibid.], which establishes the well definedness of the functor. In Section 5, we formulate the volume conjecture for the Teichmüller TQFT and describe a couple of examples for which that conjecture has already been established. In the final Section 6, we briefly describe the new formulation of the Teichmüller TQFT and anticipate a number of future developments.

Acknowledgments. We would like to thank Gregor Masbaum, Nikolai Reshetikhin, and Vladimir Turaev for valuable discussions. Our special thanks go to Feng Luo for explaining to us the topological significance of non-negative angle structures and Anton Mellit for rising important questions about analytically continued partition functions.

2 The topological domain categroid

In this section we set up the topological domain categroid on which our Teichmüller TQFT is defined. Since, in its original formulation, this TQFT is distributional in nature, we cannot use the full 3-dimensional bordism category of triangulated pseudo 3-manifolds (with extra structures), and we consider a suitable sub-categroid of it as follows.

Let X be a triangulated pseudo 3-manifold, that is a CW-complex obtained by gluing finitely many tetrahedra with ordered vertices along codimension one faces with respect to order preserving simplicial maps.

For $i \in \{0,1,2,3\}$, we will denote by $\Delta_i(X)$ the set of i-dimensional cells in X. For any $i > j$, we also denote

$$\Delta_i^j(X) := \{(a,b) \mid a \in \Delta_i(X),\ b \in \Delta_j(a)\},$$

where the cell a, when considered as a CW-complex, is taken in the form of the standard simplex without identifications on its boundary induced by gluings.

A *shape structure* on X is an assignment to each edge of each tetrahedron of X a positive number, $\alpha_X : \Delta_3^1(X) \to \mathbb{R}_{>0}$, called a *dihedral angle* such that the sum of the angles at any three edges sharing a vertex of a tetrahedron is π. It is straightforward to see that the dihedral angles at opposite edges of all tetrahedra are equal, so that each tetrahedron acquires three dihedral angles associated to the three pairs of opposite edges which sum up to π, see Figure 1. In other words, a shape structure provides each tetrahedron with the geometric structure of an ideal hyperbolic tetrahedron. An oriented triangulated pseudo 3-manifold with a shape structure is called a *shaped pseudo 3-manifold*. We denote the set of shape structures on X by $S(X)$.

An edge is called *balanced* if it is internal and the sum of the dihedral angles around it is 2π. An edge which is not balanced is called *unbalanced*. An *angle structure* on a closed triangulated pseudo 3-manifold, introduced by Cassson [n.d.], Rivin [2003], and Lackenby [2000], is a shape structure where all edges are balanced.

We will also consider the situation, where we are given a one dimensional sub-complex Γ of $\Delta_1(X)$, such that all univalent vertices of Γ are on the boundary of X (such a sub-complex we will call an allowed one dimensional sub-complex $\Gamma \subset X$).

We extend the shape structure by a real parameter called the *level*. This is an analog of the framing in the context of the WRT-TQFT. Thus, a *levelled shaped pseudo 3-manifold*

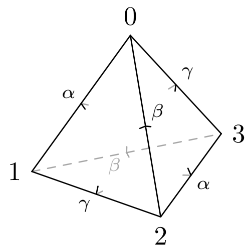


Figure 1: A tetrahedron with ordered vertices and dihedral angles.

is a pair (X, ℓ_X) consisting of a shaped pseudo 3-manifold X and a level $\ell_X \in \mathbb{R}$, and we denote by $LS(X)$ the set of all levelled shaped structures on X.

Two levelled shaped pseudo 3-manifolds (X, ℓ_X) and (Y, ℓ_Y) are called *gauge equivalent* if there exists an isomorphism $h\colon X \to Y$ of the underlying cellular structures and a function

$$g\colon \Delta_1(X) \to \mathbb{R}, \quad g|_{\Delta_1(\partial X)} = 0,$$

such that the shape structures of X and Y and levels ℓ_X, ℓ_Y are related by the formulae stated in Definition 2 of Andersen and Kashaev [2014a]. As is explained in Section 2 of Andersen and Kashaev [ibid.], this equivalence is induced by a Hamiltonian group action corresponding to the Neumann–Zagier symplectic structure. In the particular case $X = Y$ and the identity isomorphism, we get the notion of *based gauge equivalence* of levelled shaped pseudo 3-manifolds. The set of based gauge equivalence classes of levelled shape structures on X is denoted $LS_r(X)$ and we denote by $S_r(X)$ the corresponding set of based gauge equivalence classes of just shape structures (obtained by forgetting the level).

By removing the positivity condition in the definition of a shape structure, we define a *generalised shape* structure on X, and we denote by $\tilde{S}(X)$ the set of generalised shape structures. Levelled generalised shaped structures as well as their gauge equivalence are defined analogously. The space of based gauge equivalence classes of generalised shape structures (respectively levelled generalised shaped structures) is denoted $\tilde{S}_r(X)$ (respectively $\widetilde{LS}_r(X)$). Remark that $S_r(X)$ is an open convex subset of $\tilde{S}_r(X)$.

Let $\tilde{\Omega}_X : \tilde{S}(X) \to \mathbb{R}^{\Delta_1(X)}$ be the map which associates to an edge e the sum of the dihedral angles around e. The values of $\tilde{\Omega}_X$ will be called *(edge) weights*. Due to gauge invariance, $\tilde{\Omega}_X$ induces a unique map $\tilde{\Omega}_{X,r} : \tilde{S}_r(X) \to \mathbb{R}^{\Delta_1(X)}$.

Let $N_0(X)$ be a sufficiently small tubular neighbourhood of $\Delta_0(X)$. The boundary $\partial N_0(X)$ is a two dimensional surface, which is possibly disconnected and possibly with boundary, if $\partial X \neq \varnothing$. Theorem 1 of Andersen and Kashaev [ibid.] states that the map $\tilde{\Omega}_{X,r}$ is an affine $H^1(\partial N_0(X), \mathbb{R})$-bundle. The space $\tilde{S}_r(X)$ carries a Poisson structure whose symplectic leaves are the fibers of $\tilde{\Omega}_{X,r}$ and which is identical to the Poisson structure induced by the $H^1(\partial N_0(X), \mathbb{R})$-bundle structure. The natural projection map from $\widetilde{LS}_r(X)$ to $\tilde{S}_r(X)$ is an affine $\mathbb{R}$-bundle which restricts to the affine $\mathbb{R}$-bundle $LS_r(X)$ over $S_r(X)$.

If $h : X \to Y$ is an isomorphism of cellular structures, then we get an induced Poisson isomorphism $h^* : \tilde{S}_r(Y) \to \tilde{S}_r(X)$ which is an affine bundle isomorphism with respect to the induced group homomorphism

$$h^* : H^1(\partial N_0(Y), \mathbb{R}) \to H^1(\partial N_0(X), \mathbb{R})$$

and which maps $S_r(Y)$ to $S_r(X)$. Furthermore, h induces an isomorphism

$$h^* \colon \widetilde{LS}_r(Y) \to \widetilde{LS}_r(X)$$

of affine $\mathbb{R}$-bundles covering the map $h^* \colon \tilde{S}_r(Y) \to \tilde{S}_r(X)$ and which also maps $LS_r(Y)$ to $LS_r(X)$.

Let us now consider the 3-2 Pachner move illustrated in Figure 2. Let e be a balanced edge of a shaped pseudo 3-manifold X and assume that e is shared by exactly three distinct tetrahedra t_1, t_2, t_3. Let S be a shaped pseudo 3-submanifold of X composed of the tetrahedra t_1, t_2, t_3. Note that S has e as its only internal and balanced edge. There exists another triangulation S_e of the topological space underlying S such that the triangulation of ∂S coincides with that of ∂S_e, but which consists of only two tetrahedra t_4, t_5. We see that this change has the effect of removing the edge e so that $\Delta_1(S_e) = \Delta_1(S) \setminus \{e\}$. Moreover, there exists a unique shaped structure on S_e which induces the same edge weights as the shape structure of S. For shape variables $(\alpha_i, \beta_i, \gamma_i)$ for t_i (where α_i are the angles at e), the explicit map is given by

$$\begin{array}{ll} \alpha_4 = \beta_2 + \gamma_1 & \alpha_5 = \beta_1 + \gamma_2 \\ \beta_4 = \beta_1 + \gamma_3 & \beta_5 = \beta_3 + \gamma_1 \\ \gamma_4 = \beta_3 + \gamma_2 & \gamma_5 = \beta_2 + \gamma_3. \end{array} \tag{1}$$

We observe that the equation $\alpha_1 + \alpha_2 + \alpha_3 = 2\pi$ implies that the angles for t_4 and t_5 sum up to π. Moreover the positivity of the angles for t_1, t_2, t_3 implies that the angles for t_4 and t_5 are also positive. On the other hand, it is not automatic that we can solve for positive angles for t_1, t_2, t_3 given the positive angles for t_4 and t_5. However if we have two positive solutions for the angles for t_1, t_2, t_3 for the same t_4, t_5, then they are gauge equivalent and satisfy the equality $\alpha_1 + \alpha_2 + \alpha_3 = 2\pi$.

Definition 1. *We say that a shaped pseudo 3-manifold Y is obtained from X by a* shaped 3-2 Pachner move *along the edge e if Y is obtained from X by replacing S by S_e, and we write $Y = X_e$.*

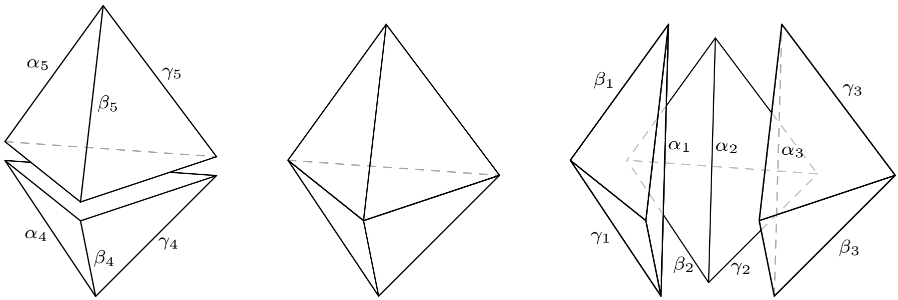


Figure 2: The 3-2 Pachner move

We observe from the above that there is a canonical map $P^e : S(X) \to S(Y)$, which naturally extends to a map $\tilde{P}^e : \tilde{S}(X) \to \tilde{S}(Y)$. We get the following commutative diagram

(2)
$$\begin{tikzcd}
\tilde{\Omega}_X(e)^{-1}(2\pi) \arrow[r, "\tilde{P}^e"] \arrow[d] & \tilde{S}(Y) \arrow[d] \\
\tilde{\Omega}_{X,r}(e)^{-1}(2\pi) \arrow[r, "\tilde{P}^e_r"] \arrow[d, "\mathrm{proj}\ \circ\ \tilde{\Omega}_{X,r}"] & \tilde{S}_r(Y) \arrow[d, "\tilde{\Omega}_{Y,r}"] \\
\mathbb{R}^{\Delta_1(X)\setminus\{e\}} \arrow[r, "="] & \mathbb{R}^{\Delta_1(Y)}
\end{tikzcd}$$

Moreover
$$\tilde{P}^e_r(\tilde{\Omega}_{X,r}(e)^{-1}(2\pi) \cap S_r(X)) \subset S_r(Y).$$
In particular, we observe that if $\tilde{\Omega}_{X,r}(e)^{-1}(2\pi) \cap S_r(X) \neq \varnothing$ then $S_r(Y) \neq \varnothing$, but the converse is not necessarily true. The following theorem is proved in Section 2 of Andersen and Kashaev [2014a].

Theorem 1. *Suppose that a shaped pseudo 3-manifold Y is obtained from a shaped pseudo 3-manifold X by a levelled shaped 3-2 Pachner move along the edge e. Then the map $\tilde{P}^e_r$ is a Poisson isomorphism, which is covered by an affine $\mathbb{R}$-bundle isomorphism from $\widetilde{LS}_r(X)|_{\tilde{\Omega}_{X,r}(e)^{-1}(2\pi)}$ to $\widetilde{LS}_r(Y)$.*

We also say that a levelled shaped pseudo 3-manifold (Y, ℓ_Y) is obtained from a levelled shaped pseudo 3-manifold (X, ℓ_X) by a *levelled shaped 3-2 Pachner move* if there exists $e \in \Delta_1(X)$ such that $Y = X_e$ and the levels are related by the formula stated just above Definition 9 in Andersen and Kashaev [ibid.].

Definition 2. *A (levelled) shaped pseudo 3-manifold X is called a* Pachner refinement *of a (levelled) shaped pseudo 3-manifold Y if there exists a finite sequence of (levelled) shaped pseudo 3-manifolds $X = X_1, X_2, \dots, X_n = Y$ such that for any $i \in \{1, \dots, n-1\}$, X_{i+1} is obtained from X_i by a (levelled) shaped 3-2 Pachner move. Two (levelled) shaped pseudo 3-manifolds X and Y are called* equivalent *if there exist gauge equivalent (levelled) shaped pseudo 3-manifolds X' and Y' which are respective Pachner refinements of X and Y.*

In its original formulation, our Teichmüler TQFT is not defined on all levelled shaped pseudo 3-manifolds. It is only guaranteed to be well defined on $S_r(X)$ (since we need the positivity of the angles to make certain integrals absolutely convergent) and when $H_2(X - \Delta_0(X), \mathbb{Z}) = 0$. The latter condition guarantees that we can multiply the distributions for all the tetrahedra and peform the necessary push forward of this product. We therefore need the following definition.

Definition 3. *An oriented triangulated pseudo 3-manifold is called* admissible *if $S_r(X) \neq \varnothing$ and $H_2(X - \Delta_0(X), \mathbb{Z}) = 0$.*

The equivalence of admissible levelled shaped pseudo 3-manifolds also needs to be such that all involved pseudo 3-manifolds are admissible, hence we introduce a stronger notion of *admissibly equivalence*.

Definition 4. *Two admissible (levelled) shaped pseudo 3-manifolds X and Y are called* admissibly equivalent *if there exists a gauge equivalence $h\colon X' \to Y'$ of (levelled) shaped pseudo 3-manifolds X' and Y' which are respective Pachner refinements of X and Y, such that*

$$\Delta_1(X') = \Delta_1(X) \cup D_X,\ \Delta_1(Y') = \Delta_1(Y) \cup D_Y$$

and

$$h(S_r(X') \cap \tilde{\Omega}_{X',r}(D_X)^{-1}(2\pi)) \cap \tilde{\Omega}_{Y',r}(D_Y)^{-1}(2\pi) \neq \varnothing.$$

The corresponding equivalence classes are called admissible equivalence classes.

Theorem 2 (Andersen and Kashaev [2014a]). *Suppose two (levelled) shaped pseudo 3-manifolds X and Y are equivalent. Then there exist $D \subset \Delta_1(X)$ and $D' \subset \Delta_1(Y)$, a bijection*

$$i\colon \Delta_1(X) \setminus D \to \Delta_1(Y) \setminus D',$$

and a Poisson isomorphism

$$R\colon \tilde{\Omega}_{X,r}(D)^{-1}(2\pi) \to \tilde{\Omega}_{Y,r}(D')^{-1}(2\pi),$$

covered by an affine $\mathbb{R}$-bundle isomorphism

$$\tilde{R}\colon \widetilde{LS}_r(X)|_{\tilde{\Omega}_{X,r}(D)^{-1}(2\pi)} \to \widetilde{LS}_r(Y)|_{\tilde{\Omega}_{Y,r}(D')^{-1}(2\pi)},$$

such that the following diagram is commutative

$$\begin{tikzcd}
\tilde{\Omega}_{X,r}(D)^{-1}(2\pi) \arrow[r, "R"] \arrow[d, "\mathrm{proj}\,\circ\,\tilde{\Omega}_{X,r}"] & \tilde{\Omega}_{Y,r}(D')^{-1}(2\pi) \arrow[d, "\mathrm{proj}\,\circ\,\tilde{\Omega}_{Y,r}"] \\
\mathbb{R}^{\Delta_1(X)\setminus D} \arrow[r, "i^*"] & \mathbb{R}^{\Delta_1(Y)\setminus D'}.
\end{tikzcd}$$

Moreover, if X and Y are admissible and admissibly equivalent, then the isomorphism R takes an open non-empty convex subset U of $S_r(X) \cap \tilde{\Omega}_{X,r}(D)^{-1}(2\pi)$ onto an open non-empty convex subset U' of $S_r(Y) \cap \tilde{\Omega}_{Y,r}(D)^{-1}(2\pi)$.

We observe that in the notation of Definition 4

$$D = \Delta_1(X) \cap h^{-1}(D_Y), \; D' = \Delta_1(Y) \cap h(D_X).$$

Let us now recall the categroid of admissible levelled shaped pseudo 3-manifolds. To this end we first need to recall the underlying category $\mathcal{B}$ where equivalence classes of levelled shaped pseudo 3-manifolds form morphisms, the objects are triangulated surfaces, and composition is given by gluings along the relevant parts of the boundaries by edge orientation preserving and face orientation reversing CW-homeomorphisms with the obvious composition of dihedral angles and addition of levels. Depending on the way we split the boundary into a source and a target, one and the same levelled shaped pseudo 3-manifold can be interpreted as different morphisms in $\mathcal{B}$. Nonetheless, there is one canonical choice defined as follows.

For a tetrahedron $T = [v_0, v_1, v_2, v_3]$ in $\mathbb{R}^3$ with ordered vertices v_0, v_1, v_2, v_3, we define its sign by

$$\operatorname{sign}(T) = \operatorname{sign}(\det(v_1 - v_0, v_2 - v_0, v_3 - v_0)),$$

as well as the signs of its faces

$$\operatorname{sign}(\partial_i T) = (-1)^i \operatorname{sign}(T), \quad i \in \{0, \ldots, 3\}.$$

For a pseudo 3-manifold X, the signs of the faces of the tetrahedra of X induce a sign function on the faces of the boundary of X,

$$\operatorname{sign}_X \colon \Delta_2(\partial X) \to \{\pm 1\},$$

which permits to split the boundary of X into two subsets,

$$\partial X = \partial_+ X \cup \partial_- X, \quad \Delta_2(\partial_\pm X) = \operatorname{sign}_X^{-1}(\pm 1),$$

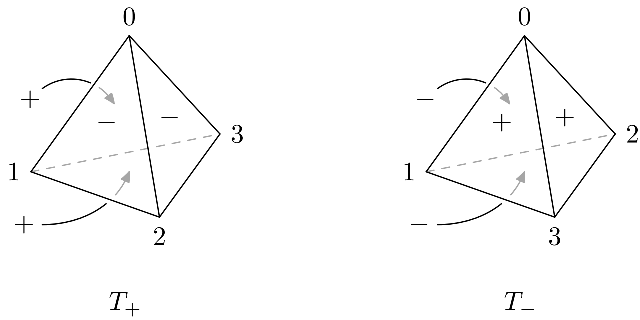


Figure 3: Face orientations

composed of equal numbers of triangles. For example, in the case of a tetrahedron T with $\mathrm{sign}(T) = 1$, we have $\Delta_2(\partial_+ T) = \{\partial_0 T, \partial_2 T\}$, and $\Delta_2(\partial_- T) = \{\partial_1 T, \partial_3 T\}$. In what follows, unless specified otherwise, (the equivalence class of) a levelled shaped pseudo 3-manifold X will always be thought of as a $\mathcal{B}$-morphism between the objects $\partial_- X$ and $\partial_+ X$, i.e.

$$X \in \mathrm{Hom}_{\mathcal{B}}(\partial_- X, \partial_+ X).$$

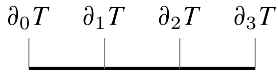
We will also consider more general morphisms to be included in $\mathcal{B}$, namely morphisms as above, but where we add an allowed one dimensional sub-complexes as defined above. This means that we allow objects where we have special marked vertices on the boundaries and when we compose such morphisms, we assume that all univalent vertices of the sub-complexes, which are contained in the surfaces we glue on, match up pairwise, thus resulting in an allowed one dimensional sub-complex in the morphism obtained by gluing. In the equivalence relation, this one dimensional sub-complex should be carried all the way through the equivalences specified in Definition 2 and 4, but in such a way that we never perform any 3-2 Pachner moves on edges, which are part of the one-dimensional sub-complex. In the rest of this paper we use the term *levelled shaped pseudo* 3*-manifold* to mean any morphism of $\mathcal{B}$ (including also the morphisms we just added to $\mathcal{B}$).

Our TQFT is not defined on the full category $\mathcal{B}$, but only on the sub-categroid of admissible equivalence classes of admissible morphisms.

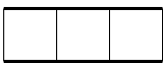
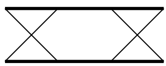
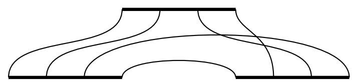
Definition 5. *The categroid $\mathcal{B}_a$ of admissible levelled shaped pseudo* 3*-manifolds is the sub-categroid of the category of levelled shaped pseudo* 3*-manifolds whose morphisms consist of admissible equivalence classes of admissible levelled shaped pseudo* 3*-manifolds.*

Gluing in this sub-categroid is the one induced from the category $\mathcal{B}$ and it is only defined for those pairs of admissible morphisms for which the glued morphism in $\mathcal{B}$ is also admissible.

There is a concise graphical presentation of pseudo 3-manifolds introduced in Andersen and Kashaev [2014a]. To each tetrahedron T, it associates the graph

(3) $T =$ 

where each of the four codimension one faces of T corresponds to a vertical half edge. We connect the half edges according to the face identifications of the tetrahedra in a given triangulated pseudo 3-manifold (in the case of non-empty boundary, the resulting graph will also have open half edges). For example, the graphs for the pseudo 3-manifolds representing the complements of the trefoil knot 3_1, the figure eight knot 4_1 and the 5_2-knot are as follows

 ,  and 

All these examples correspond to 3-manifolds with one cusp, i.e. they are 1-vertex triangulations with the vertex having a neighborhood homeomorphic to the cone over the torus. As it will be seen below, our TQFT functor, up to overall orientation, can be written down just based on such graphical presentation.

3 The target categroid

The target categroid for the Teichmüller TQFT is given by tempered distributions. They form only a categroid since the kind of composition of distributions we have in mind is not defined for all tempered distributions.

Recall that the space of (complex) tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ is the space of continuous linear functionals on the (complex) Schwartz space $\mathcal{S}(\mathbb{R}^n)$. By the Schwartz presentation theorem (see e.g. Theorem V.10 p. 139 Reed and Simon [1972]), any tempered distribution can be represented by a finite derivative of a continuous function with polynomial growth, hence we may informally think of tempered distributions as functions defined on $\mathbb{R}^n$. The integral formula

$$\varphi(f) = \int_{\mathbb{R}^n} \varphi(x) f(x) dx.$$

exhibits the inclusion $\mathcal{S}(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$.

Definition 6. *The categroid $\mathfrak{D}$ has as objects finite sets and for two finite sets n, m the set of morphisms from n to m is*

$$\mathrm{Hom}_{\mathfrak{D}}(n, m) = \mathcal{S}'(\mathbb{R}^{n \sqcup m}).$$

Denoting by $\mathcal{L}(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^m))$ the space of continuous linear maps from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^m)$, we remark that we have an isomorphism

$$\tilde{\ }: \mathcal{L}(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^m)) \to \mathcal{S}'(\mathbb{R}^{n\sqcup m})$$

determined by the formula

$$\varphi(f)(g) = \tilde{\varphi}(f \otimes g)$$

for all $\varphi \in \mathcal{L}(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^m))$, $f \in \mathcal{S}(\mathbb{R}^n)$, and $g \in \mathcal{S}(\mathbb{R}^m)$. This is the content of the Nuclear theorem, see e.g. Reed and Simon [1972], Theorem V.12, p. 141. The reason why we get a categroid rather than a category is because we cannot compose all composable (in the usual categorical sense) morphisms, but only a subset thereof. The partially defined composition in this categroid is defined as follows.

Let n, m, l be three finite sets, $A \in \mathrm{Hom}_{\mathfrak{D}}(n, m)$ and $B \in \mathrm{Hom}_{\mathfrak{D}}(m, l)$. According to the tempered distribution analog of Theorem 6.1.2. in Hörmander [1969], associated to the projections

$$\pi_{n,m}\colon \mathbb{R}^{n\sqcup m\sqcup l} \to \mathbb{R}^{n\sqcup m}, \quad \pi_{m,l}\colon \mathbb{R}^{n\sqcup m\sqcup l} \to \mathbb{R}^{m\sqcup l},$$

we have the pull back maps

$$\pi_{n,m}^*\colon \mathcal{S}'(\mathbb{R}^{n\sqcup m}) \to \mathcal{S}'(\mathbb{R}^{n\sqcup m\sqcup l}) \text{ and } \pi_{m,l}^*\colon \mathcal{S}'(\mathbb{R}^{m\sqcup l}) \to \mathcal{S}'(\mathbb{R}^{n\sqcup m\sqcup l}).$$

By theorem IX.45 in Reed and Simon [1975] (see also Appendix B in Andersen and Kashaev [2014a]), the product

$$\pi_{n,m}^*(A)\pi_{m,l}^*(B) \in \mathcal{S}'(\mathbb{R}^{n\sqcup m\sqcup l})$$

is well defined provided the wave front sets of $\pi_{n,m}^*(A)$ and $\pi_{m,l}^*(B)$ satisfy the following transversality condition

$$(\mathrm{WF}(\pi_{n,m}^*(A)) \oplus \mathrm{WF}(\pi_{m,l}^*(B))) \cap Z_{n\sqcup m\sqcup l} = \varnothing, \tag{4}$$

where $Z_{n\sqcup m\sqcup l}$ is the zero section of $T^*(\mathbb{R}^{n\sqcup m\sqcup l})$. If we now further assume that $\pi_{n,m}^*(A)\pi_{m,l}^*(B)$ continuously extends to $\mathcal{S}(\mathbb{R}^{n\sqcup m\sqcup l})_m$ as is defined in Appendix B of Andersen and Kashaev [ibid.], then we obtain a well defined element

$$(\pi_{n,l})_*(\pi_{n,m}^*(A)\pi_{m,l}^*(B)) \in \mathcal{S}'(\mathbb{R}^{n\sqcup l}).$$

Definition 7. *For $A \in \mathrm{Hom}_{\mathfrak{D}}(n, m)$ and $B \in \mathrm{Hom}_{\mathfrak{D}}(m, l)$ satisfying condition (4) and such that $\pi_{n,m}^*(A)\pi_{m,l}^*(B)$ continuously extends to a well defined element of the dual of $S(\mathbb{R}^{n\sqcup m\sqcup l})_m$, we define*

$$AB = (\pi_{n,l})_*(\pi_{n,m}^*(A)\pi_{m,l}^*(B)) \in \mathrm{Hom}_{\mathfrak{D}}(n, l).$$

For any $A \in \mathfrak{L}(\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^m))$, we have unique adjoint $A^* \in \mathfrak{L}(\mathcal{S}(\mathbb{R}^m), \mathcal{S}'(\mathbb{R}^n))$ defined by the formula

$$A^*(f)(g) = \overline{A(\bar{g})(\bar{f})}$$

for all $f \in \mathcal{S}(\mathbb{R}^m)$ and $g \in \mathcal{S}(\mathbb{R}^n)$.

4 The TQFT functor

We shall describe the Teichmüller TQFT functor $F_\hbar$ from Andersen and Kashaev [ibid.]. First we recall the definition of a $*$-functor in our context.

Definition 8. *A functor $F : \mathfrak{B}_a \to \mathfrak{D}$ is said to be a $*$-*functor *if*

$$F(X^*) = F(X)^*,$$

where X^* is X with opposite orientation, and $F(X)^*$ is the adjoint of $F(X)$.

On the level of objects we define

$$F_\hbar(\Sigma) = \Delta_2(\Sigma), \quad \forall \Sigma \in \operatorname{Ob} \mathfrak{B}_a.$$

In order to define $F_\hbar$ on morphisms, we need a special function called Faddeev's quantum dilogarithm defined in Faddeev [1995].

Definition 9. Faddeev's quantum dilogarithm *is the function of two complex arguments z and b defined for $|\operatorname{Im} z| < \frac{1}{2}|\mathsf{b} + \mathsf{b}^{-1}|$ by the formula*

$$\Phi_\mathsf{b}(z) := \exp\left(\int_C \frac{e^{-2\mathsf{i}zw}\, dw}{4\sinh(w\mathsf{b})\sinh(w/\mathsf{b})w}\right),$$

where the contour C runs along the real axis deviating into the upper half plane in the vicinity of the origin, and extended by the functional equation

$$\Phi_\mathsf{b}(z - \mathsf{i}\mathsf{b}^{\pm 1}/2) = (1 + e^{2\pi \mathsf{b}^{\pm 1} z})\Phi_\mathsf{b}(z + \mathsf{i}\mathsf{b}^{\pm 1}/2)$$

to a meromorphic function for $z \in \mathbb{C}$.

It is easily seen that $\Phi_\mathsf{b}(z)$ depends on b only through the combination $\hbar$ defined by the formula

$$\hbar := \left(\mathsf{b} + \mathsf{b}^{-1}\right)^{-2}.$$

In what follows, we assume that the complex parameter b is such that $\hbar \in \mathbb{R}_{>0}$. This assumption guarantees we get a *unitary* TQFT, but, in case of need, one can easily go to arbitrary $\mathsf{b} \in \mathbb{C} \setminus \mathsf{i}\mathbb{R}$ by analytic continuation.

The value of $F_\hbar$ on a morphism (X, Γ) of $\mathcal{B}_a$ is given by the formula which singles out the dependence on the level

$$F_\hbar(X, \Gamma) = e^{\mathsf{i}\pi \frac{\ell_X}{4\hbar}} Z_\hbar(X, \Gamma) \in \mathcal{S}'\left(\mathbb{R}^{\Delta_2(\partial X)}\right),$$

where $Z_\hbar(X, \Gamma)$ is the level independent part.

The value of $Z_\hbar$ on the morphism of $\mathcal{B}_a$ determined by a single tetrahedron T with $\operatorname{sign}(T) = 1$ is an element

$$Z_\hbar(T, \alpha_T) \in \mathcal{S}'(\mathbb{R}^{\Delta_2(T)})$$

given by the explicit formula

$$Z_\hbar(T, \alpha_T)(x_0, x_1, x_2, x_3) = \delta(x_0 - x_1 + x_2) \frac{e^{2\pi\mathsf{i}(x_3 - x_2)\left(x_0 + \frac{\alpha_3}{2\mathsf{i}\sqrt{\hbar}}\right) + \pi\mathsf{i}\frac{\varphi_T}{4\hbar}}}{\Phi_\mathsf{b}\left(x_3 - x_2 + \frac{1-\alpha_1}{2\mathsf{i}\sqrt{\hbar}}\right)} \tag{5}$$

where δ is Dirac's delta-function supported at $0 \in \mathbb{R}$,

$$\varphi_T := \alpha_1\alpha_3 + \frac{\alpha_1 - \alpha_3}{3} - \frac{2\hbar + 1}{6}, \quad \alpha_i := \frac{1}{\pi}\alpha_T(\partial_0\partial_i T), \quad i \in \{1, 2, 3\},$$

and

$$x_i := x(\partial_i T), \quad x \colon \Delta_2(\partial T) \to \mathbb{R}.$$

For a negative tetrahedron $\bar{T}$ with $\operatorname{sign}(\bar{T}) = -1$ we set

$$Z_\hbar(\bar{T}) = Z_\hbar(T)^*.$$

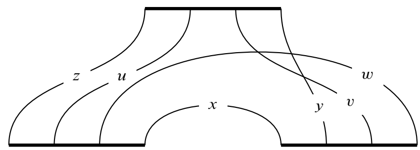
It is not hard to check that these assignments give tempered distributions provided $\alpha_i > 0$, $i = 1, 2, 3$.

The value of $Z_\hbar$ on arbitrary morphism (X, Γ) in $\mathcal{B}_a$ is given by composing all the distributions $Z_\hbar(T)$, where T runs over $\Delta_3(X)$, according to the face identifications which build X out of the disjoint union

$$\tilde{X} := \bigsqcup_{T \in \Delta_3(X)} T.$$

By using the graphical presentation of X described above, with the additional information on the orientation of X, the prescription is as follows. One should label the thin edges of the graph with variables x_i, where $i = 1, \ldots, |\Delta_2(X)|$, then take the product over all tetrahedra of the expression (5) or its complex conjugate adapted to each tetrahedron in accordance with the variables attached to its four faces and the dihedral angles, and integrate over all real values of x_i.

Let us illustrate this with the example of the complement X of the knot 5_2 which is represented by the diagram



We denote T_1, T_2, T_3 the left, right, and top tetrahedra respectively with their dihedral angles $\alpha_{T_i} = 2\pi(a_i, b_i, c_i)$, such that $a_i + b_i + c_i = \frac{1}{2}$, $i = 1, 2, 3$. We choose the orientation so that all tetrahedra are positive, and we impose the conditions that all edges are balanced ($\Gamma = \varnothing$) which correspond to two equations

$$2a_3 = a_1 + c_2, \quad b_3 = c_1 + b_2.$$

We thus get by the definition of our TQFT that

$$Z_\hbar(X) = \int_{\mathbb{R}^6} Z_\hbar(T_1, \alpha_{T_1})(z, u, w, x) Z_\hbar(T_2, \alpha_{T_2})(x, y, v, w) \\ \times Z_\hbar(T_3, \alpha_{T_3})(y, v, u, z)\, \mathrm{d}^6(x, y, z, u, v, w), \tag{6}$$

where we observe that the integrand indeed extends to the dual of $S(\mathbb{R}^6)_6$, thus it can be pushed forward to a point, which is the precise meaning of the integral in (6). The calculation in Section 11.6 of Andersen and Kashaev [2014a] gives

$$Z_\hbar(X) = \nu_{c_1,b_1}\nu_{b_2,a_2}\nu_{c_3,b_3} e^{\mathsf{i}\pi c_{\mathsf{b}}^2(1-2a_1)(1-2c_2)} \int_{2c_{\mathsf{b}}(a_1-a_3)+\mathbb{R}} \chi_{5_2}(x, \lambda)\, \mathrm{d}x,$$

where $\lambda := a_1 - c_1 + b_2 - a_3$,

$$\nu_{a,b} := e^{4\pi \mathsf{i} c_{\mathsf{b}}^2 a(a+b)} e^{-\pi \mathsf{i} c_{\mathsf{b}}^2(4(a-b)+1)/6}, \quad c_{\mathsf{b}} := \frac{\mathsf{i}}{2\sqrt{\hbar}} = \frac{\mathsf{i}}{2}(\mathsf{b} + \mathsf{b}^{-1}),$$

and

$$\chi_{5_2}(x, \lambda) := \chi_{5_2}(x) e^{4\pi \mathsf{i} c_{\mathsf{b}} x \lambda}, \\ \chi_{5_2}(x) := e^{-\mathsf{i}\pi/3} \int_{\mathbb{R}-\mathsf{i}0} \mathrm{d}z\, \frac{e^{\mathsf{i}\pi(z-x)(z+x)}}{\Phi_{\mathsf{b}}(z+x)\Phi_{\mathsf{b}}(z-x)\Phi_{\mathsf{b}}(z)}. \tag{7}$$

Returning now back to the case of a general X, we need to know that all the compositions of the $Z_\hbar(T, \alpha_T)$'s are allowed in $\mathfrak{D}$. This is precisely the content of Theorem 9 in

Andersen and Kashaev [2014a], which establishes that for admissible X, the wave front sets of the distributions $\pi_T^* Z_\hbar(T)$, where $\pi_T : \mathbb{R}^{\Delta_2(X)} \to \mathbb{R}^{\Delta_2(T)}$ is the natural projection for each $T \in \Delta_3(X)$, are transverse and hence they can be multiplied and their product can be pulled back to $\mathbb{R}^{\Delta_2(X)}$ and pushed forward along the projection from $\mathbb{R}^{\Delta_2(X)}$ to $\mathbb{R}^{\Delta_2(\partial X)}$.

We emphasize that for an admissible pseudo 3-manifold X together with an allowed sub-complex Γ of $\Delta_1(X)$, our TQFT functor provides us with the following well defined *real analytic* function

$$F_\hbar(X,\Gamma) : LS_r(X) \cap \tilde{\Omega}_{X,r}(E_\Gamma)^{-1}(2\pi) \to \mathcal{S}'(\mathbb{R}^{\partial X}),$$

where E_Γ is the set of internal edges of X which are not in Γ. We note that if (X,Γ) is admissibly equivalent to (X',Γ'), then Theorem 2 provides an explicit affine map from a non-empty open convex subset of $LS_r(X) \cap \tilde{\Omega}_{X,r}(E_\Gamma)^{-1}(2\pi)$ to an open convex subset of $LS_r(X') \cap \tilde{\Omega}_{X',r}(E_{\Gamma'})^{-1}(2\pi)$ and under this map the restrictions of $F_\hbar(X,\Gamma)$ and $F_\hbar(X',\Gamma')$ to these two non-empty convex open subsets agree. Since a real analytic map defined on an open convex set of some Euclidian space is uniquely determined by its restriction to any smaller non-empty open convex subset, we see that $F_\hbar(X,\Gamma)$ and $F_\hbar(X',\Gamma')$ uniquely determine each other on their domains of definition. It is in this sense that our Teichmüller TQFT is well-defined on the set of equivalence classes of admissible levelled shaped pseudo 3-manifolds.

For the case $\partial X = \varnothing$, we have $\mathcal{S}'(\mathbb{R}^{\partial X}) = \mathbb{C}$ and so, in this case, we simply get a complex valued function on $LS_r(X) \cap \tilde{\Omega}_{X,r}(E_\Gamma)^{-1}(2\pi)$. In particular, the value of the functor $F_\hbar$ on any fully balanced admissible levelled shaped 3-manifold is a complex number, which is a topological invariant, in the sense that if two fully balanced admissible levelled shaped 3-manifolds are admissibly equivalent, then $F_\hbar$ assigns one and the same complex number to them. We recall that fully balanced means that all edges are balanced.

Our main Theorem 4 of Andersen and Kashaev [ibid.] now guarantees that this assignment, in fact, gives a well-defined functor.

Theorem 3. *For any $\hbar \in \mathbb{R}_{>0}$, the above assignment defines a $*$-functor*

$$F_\hbar : \mathcal{B}_a \to \mathfrak{D}$$

which we call the Teichmüller TQFT.

5 The volume conjecture for the Teichmüller TQFT

In this subsection we recall our conjecture from Andersen and Kashaev [ibid.] concerning our Teichmüller TQFT $F_\hbar$, which, among other things, provides a relation to the hyperbolic volume in the asymptotic limit $\hbar \to 0$.

Conjecture 1. *Let M be a closed oriented compact 3-manifold. For any hyperbolic knot $K \subset M$, there exists a smooth function $J_{M,K}(\hbar, x)$ on $\mathbb{R}_{>0} \times \mathbb{R}$ which has the following properties.*

1. *For any fully balanced shaped ideal triangulation X of the complement of K in M, there exist a gauge invariant real linear combination of dihedral angles λ and a (gauge non-invariant) real quadratic polynomial of dihedral angles ϕ such that*
$$Z_\hbar(X) = e^{\mathsf{i}\frac{\phi}{\hbar}} \int_{\mathbb{R}} J_{M,K}(\hbar, x) e^{-\frac{x\lambda}{\sqrt{\hbar}}} \,\mathrm{d}x$$
2. *The hyperbolic volume of the complement of K in M is recovered as the following limit*
$$\lim_{\hbar \to 0} 2\pi\hbar \log |J_{M,K}(\hbar, 0)| = -\operatorname{Vol}(M \setminus K).$$

Remark 1. *It is very important to notice that we in part (2) of this conjecture have a negative sign on the right hand side, which differs from the volume conjecture of Kashaev [1997]. In this case, the invariant exponentially decays (rather than grows) with the rate being given by the hyperbolic volume.*

In Andersen and Kashaev [2014a], we checked this conjecture for the first two hyperbolic knots.

Theorem 4. *Conjecture 1 is true for the pairs $(S^3, 4_1)$ and $(S^3, 5_2)$ with*
$$J_{S^3,4_1}(\hbar, x) = \chi_{4_1}(x), \quad J_{S^3,5_2}(\hbar, x) = \chi_{5_2}(x),$$
where the functions $\chi_{4_1}(x)$ and $\chi_{5_2}(x)$ are given by
$$\chi_{4_1}(x) = \int_{\mathbb{R}-i0} \frac{\Phi_{\mathsf{b}}(x-y)}{\Phi_{\mathsf{b}}(y)} e^{2\pi i x(2y-x)} \,\mathrm{d}y$$
and $\chi_{5_2}(x)$ is given in (7) above.

See also Andersen and Marzioni [2017] for a precise statement of the generalisation of the above conjecture to the higher level generalisation of the Teichmüller TQFT and more examples in Andersen and Nissen [2017]. Further, in Andersen and Malusà [2017] we have presented the precise formulation of the AJ-conjecture for the Teichmüller TQFT.

6 Future perspectives

In the paper Andersen and Kashaev [2013] we have presented a new formulation of the Teichmüller TQFT and a further higher level generalization of the theory, which we think of as a version of the complex quantum Chern-Simons theory Andersen and Kashaev [2014b] (see also Andersen and Marzioni [2017]). Let us here briefly recall this new formulation and state some predictions for the further perspectives for the Teichmüller TQFT.

The main player behind the new formulation of the Teichmüller TQFT is the edge-face tranform using the Weil–Gel'fand–Zak (WGZ) transformation, which we now recall.

We consider the following multiplier construction for a line bundle over the two torus. We have the natural translation action of $\mathbb{Z}^2$ on $\mathbb{R}^2$ with the quotient $\Pi := \mathbb{S}^2$ where $\mathbb{S} := \mathbb{R}/\mathbb{Z}$. Consider the following multipliers

$$\varphi\colon \mathbb{Z}^2 \times \mathbb{R}^2 \to U(1), \quad \varphi((m,n),(x,y)) = (-1)^{mn} e^{\pi \mathsf{i}(nx-my)},$$

which induce an action of $\mathbb{Z}^2$ on the trivial bundle $\mathbb{R}^2 \times \mathbb{C}$ and we define

$$L = (\mathbb{R}^2 \times \mathbb{C})/\mathbb{Z}^2$$

as a complex line bundle over Π.

We define the Weil–Gel'fand–Zak (WGZ) transformation

$$W : \mathcal{S}(\mathbb{R}) \to C^\infty(\Pi, L)$$

by the formula

$$(Wf)(x,y) = e^{\pi \mathsf{i} xy} \sum_{m\in\mathbb{Z}} f(x+m) e^{2\pi \mathsf{i} my}.$$

Proposition 1. *The WGZ-transformation*

$$W : \mathcal{S}(\mathbb{R}) \to C^\infty(\Pi, L)$$

is an isomorphism of Fréchet spaces with the inverse explicitly given by the formula

$$(W^{-1}g)(x) = \int_0^1 g(x,y) e^{-\pi \mathsf{i} xy}\, \mathrm{d}y$$

for any $g \in C^\infty(\Pi, L)$.

The same statement also holds for

$$\overline{W} : \mathcal{S}(\mathbb{R}) \to C^\infty(\Pi, L^*)$$

defined by

$$\overline{W}(f)(x,y) = W(f)(x,-y).$$

Guided by Equation (5), we define for any $f \in \mathcal{S}(\mathbb{R})$ the associated tempered distribution $H(f) \in \mathcal{S}'(\mathbb{R}^4)$ as follows

$$H(f)(x_0, x_1, x_2, x_3) = \delta(x_0 + x_2 - x_1) f(x_3 - x_2) e^{2\pi \mathsf{i} x_0 (x_3 - x_2)},$$

which we consider as a continuous linear map

$$H(f) : \mathcal{S}(\mathbb{R}^2) \to \mathcal{S}'(\mathbb{R}^2)$$

via the formula

$$H(f)(g)(h) = H(f)(\hat{\pi}^*(g)\check{\pi}^*(h)).$$

Here $\hat{\pi}, \check{\pi} : \mathbb{R}^4 \to \mathbb{R}^2$ are given by $\hat{\pi}(x_0, x_1, x_2, x_3) = (x_1, x_3)$ and $\check{\pi}(x_0, x_1, x_2, x_3) = (x_0, x_2)$. We now consider the following tensor extension of the WGZ-transform

$$W \otimes W : \mathcal{S}(\mathbb{R}^2) \to C^\infty(\Pi \times \Pi, L \boxtimes L)$$

defined by

$$W \otimes W(h)(s, t, x, y) = \\ e^{\pi i (st + xy)} \sum_{m_1, m_2 \in \mathbb{Z}} h(s + m_1, x + m_2) e^{2\pi i (m_1 t + m_2 y)}.$$

and similarly

$$\overline{W} \otimes \overline{W} : \mathcal{S}(\mathbb{R}^2) \to C^\infty(\Pi \times \Pi, L^* \boxtimes L^*),$$

given by

$$\overline{W} \otimes \overline{W}(h)(s, t, ; x, y) = \\ e^{-\pi i (st + xy)} \sum_{m_1, m_2 \in \mathbb{Z}} h(s + m_1, x + m_2) e^{-2\pi i (m_1 t + m_2 y)}.$$

Consider now the maps

$$F : \mathbb{S}^5 \to \Pi, \quad \tilde{\pi}_i : \mathbb{S}^5 \to \Pi^2, \quad i = 1, 2,$$

given by

$$F(u, s, t, x, y) = (u, s + t + u - y),$$

$$\tilde{\pi}_1(u, s, t, x, y) = (s + x, t + u, x + u, y - t - u)$$

and $\tilde{\pi}_2$ is the map which projects away the first factor and onto the last four factors. We then get that there exists a natural isomorphism

$$F^* L \otimes \tilde{\pi}_1^*(L \boxtimes L) \cong \tilde{\pi}_2^*(L \boxtimes L).$$

There is an obvious embedding

$$C^\infty(\Pi^2, L \boxtimes L) \subset C^\infty(\Pi^2, L^* \boxtimes L^*)^*,$$

obtained by pointwise evaluation followed by integration over Π^2. Thus we see that

$$(\tilde{\pi}_2)_*(F^* W(f)\tilde{\pi}_1^*(W \otimes W)(g)) \in C^\infty(\Pi^2, L^* \boxtimes L^*)^*$$

for any $g \in \mathcal{S}(\mathbb{R}^2)$.

Let $T : \mathbb{S}^3 \to \Pi$ be given by $T(x, y, z) = (y - x, z - y)$ and further $E : \mathbb{S}^6 \to \Pi$ by

$$E(x_{01}, x_{02}, x_{03}, x_{12}, x_{13}, x_{23}) = (x_{02} + x_{13} - x_{03} - x_{21}, x_{02} + x_{13} - x_{01} - x_{23}).$$

We also introduce the following two maps $\pi_i : \mathbb{S}^6 \to \Pi^2$, $i = 1, 2$, given by

$$\pi_1(x_{01}, x_{02}, x_{03}, x_{12}, x_{13}, x_{23}) = T \times T(x_{23}, x_{03}, x_{02}, x_{12}, x_{02}.x_{01})$$

and

$$\pi_2(x_{01}, x_{02}, x_{03}, x_{12}, x_{13}, x_{23}) = T \times T(x_{23}, x_{13}, x_{12}, x_{13}, x_{03}, x_{01}).$$

Finally we further need the map $P : \mathbb{S}^6 \to \mathbb{S}^5$ given by

$$\begin{aligned} &P(x_{01}, x_{02}, x_{03}, x_{12}, x_{13}, x_{23}) \\ &\qquad = (x_{02} + x_{13} - x_{03} - x_{12}, x_{13} - x_{23}, x_{12} - x_{13}, x_{03} - x_{13}, x_{01} - x_{03}) \end{aligned}$$

It is elementary to verify that

$$E = F \circ P \text{ and } \pi_i = \tilde{\pi}_i \circ P.$$

We now arrive at the important "edge-face transformation" as established in Andersen and Kashaev [2013].

Theorem 5. *The distribution*

$$H(f) \in \mathcal{S}'(\mathbb{R}^4)$$

and the section

$$E^* W(f) \in C^\infty(\mathbb{S}^6, E^* L)$$

are related by the formula

$$H(f)(\hat{\pi}^*(g)\check{\pi}^*(h)) = \pi_*^{(6)}(E^* W(f)\pi_1^*(W \otimes W)(g)\pi_2^*(\overline{W} \otimes \overline{W})(h))$$

for all $g, h \in S(\mathbb{R}^2)$.

This theorem is the corner stone in understanding how to transform our original Teichmüller TQFT, which is build from the fundamental distribution

$$F_{\hbar}(T) \in \mathcal{S}'(\mathbb{R}^{\Delta_2(T)}),$$

to our new formulation $\mathfrak{F}_{\hbar}$ of the Teichmüller TQFT, where the state variables live on the edges instead and the fundamental object associated to a tetrahedron is the section

$$\mathfrak{F}_{\hbar}(T) \in C^{\infty}(\mathbb{S}^{\Delta_1(T)}, E^* L).$$

Here we implicitly use the following notation

$$x_{ij} := x(v_T(i)v_T(j))$$

for any $x \in \mathbb{S}^{\Delta_1(T)}$. Now, Theorem 5 simply tells us that the two distributions $F_{\hbar}(T)$ and $\mathfrak{F}_{\hbar}(T)$ are related via the tensor square of the WGZ-transform W.

Let us recall the formula for $\mathfrak{F}_{\hbar}(T)$ from Andersen and Kashaev [ibid.]

$$\mathfrak{F}_{\hbar}(T) = E^*(g_{\alpha_0,\alpha_2}),$$

where, for two positive real numbers a and c satisfying $a + c < 1/2$, we let

$$g_{a,c} = W(\tilde{\psi}'_{a,c}),$$

$$\tilde{\psi}'_{a,c}(s) := e^{-\pi \mathsf{i} s^2} \tilde{\psi}_{a,c}(s),$$

$$\tilde{\psi}_{a,c}(s) := \int_{\mathbb{R}} \psi_{a,c}(t) e^{-2\pi \mathsf{i} s t} dt,$$

and

$$\psi_{a,c}(t) := \bar{\Phi}_{\mathsf{b}}(t - 2c_{\mathsf{b}}(a+c)) e^{-4\pi \mathsf{i} c_{\mathsf{b}} a (t - c_{\mathsf{b}}(a+c))} e^{-\pi \mathsf{i} c_{\mathsf{b}}^2 (4(a-c)+1)/6},$$

where we use the notation $\bar{\Phi}_{\mathsf{b}}(x) := 1/\Phi_{\mathsf{b}}(x)$.

As it is described in Andersen and Kashaev [ibid.] we get a new formulation $\mathfrak{F}_{\hbar}$ of the Teichmüller TQFT following the same lines as discussed above determining $\mathfrak{F}_{\hbar}$ on all objects of $\mathcal{B}_a$ by a similar gluing construction.

However, this new formulation has several advantages. It allows us to extend the partition function to complex dihedral angles and as such it depends meromorphically on these complexified angles. This allows us to actually establish that we do not need the condition of admissibility and that this new formulation $\mathfrak{F}_{\hbar}$ is well-defined on the full bordism category $\mathcal{B}$ consisting of equivalence classes of levelled shaped pseudo 3-manifolds. We stress that this means that the functor $\mathfrak{F}_{\hbar}$ is in fact invariant under all 2-3 and 3-2 Pachner moves.

By modifying the construction of the functor $\mathfrak{F}_\hbar$, we can further extend this functor to a version which depends on a first cohomology class with coefficient in $\mathbb{S}$ (see Andersen and Kashaev [2013]) and this generalised new version can be related to the original functor $F_\hbar$ by integration over this first cohomology group.

We can use similar ideas (see also Andersen and Kashaev [ibid.]) to produce a meromorphic extension of $F_\hbar$ to complex angles and as such we can establish that this theory is also invariant under all 2-3 and 3-2 Pachner moves as argued in Andersen and Kashaev [ibid.]. Furthermore, via a certain gauge fixing technique also described in Andersen and Kashaev [ibid.], we can extend the original theory $F_\hbar$ to be defined on the full bordism category $\mathfrak{B}$, and we describe the precise relation between this extension of the original formulation and the new one in Andersen and Kashaev [ibid.].

References

J. E. Andersen (2006). "Asymptotic faithfulness of the quantum SU(n) representations of the mapping class groups". *Ann. of Math. (2)* 163.1, pp. 347–368. MR: 2195137 (cit. on p. 2559).

– (2008). "The Nielsen-Thurston classification of mapping classes is determined by TQFT". *J. Math. Kyoto Univ.* 48.2, pp. 323–338. MR: 2436739 (cit. on p. 2559).

– (2010). "Asymptotics of the Hilbert-Schmidt norm of curve operators in TQFT". *Lett. Math. Phys.* 91.3, pp. 205–214. MR: 2595923 (cit. on p. 2559).

– (2013). "The Witten-Reshetikhin-Turaev invariants of finite order mapping tori I". *J. Reine Angew. Math.* 681, pp. 1–38. MR: 3181488 (cit. on p. 2559).

J. E. Andersen and B. Himpel (2012). "The Witten-Reshetikhin-Turaev invariants of finite order mapping tori II". *Quantum Topol.* 3.3-4, pp. 377–421. MR: 2928090 (cit. on p. 2559).

J. E. Andersen, B. Himpel, S. F. Jørgensen, J. Martens, and B. McLellan (2017). "The Witten-Reshetikhin-Turaev invariant for links in finite order mapping tori I". *Adv. Math.* 304, pp. 131–178. MR: 3558207 (cit. on p. 2559).

J. E. Andersen and S. F. Jørgensen (2015). "On the Witten-Reshetikhin-Turaev invariants of torus bundles". *J. Knot Theory Ramifications* 24.11, pp. 1550055, 48. MR: 3413187 (cit. on p. 2559).

J. E. Andersen and R. M. Kashaev (May 2013). "A new formulation of the Teichmüller TQFT". arXiv: 1305.4291 (cit. on pp. 2560, 2575, 2578–2580).

– (2014a). "A TQFT from Quantum Teichmüller theory". *Comm. Math. Phys.* 330.3, pp. 887–934. MR: 3227503 (cit. on pp. 2560, 2561, 2563, 2565, 2566, 2569–2571, 2573–2575).

– (Sept. 2014b). "Complex Quantum Chern-Simons". arXiv: 1409.1208 (cit. on pp. 2560, 2576).
– (2014c). "Quantum Teichmüller theory and TQFT". In: *XVIIth International Congress on Mathematical Physics*. World Sci. Publ., Hackensack, NJ, pp. 684–692. MR: 3204520 (cit. on p. 2560).
– (2015). "Faddeev's quantum dilogarithm and state-integrals on shaped triangulations". In: *Mathematical aspects of quantum field theories*. Math. Phys. Stud. Springer, Cham, pp. 133–152. MR: 3330241 (cit. on p. 2560).
J. E. Andersen and A. Malusà (Nov. 2017). "The AJ-conjecture for the Teichmüller TQFT". arXiv: 1711.11522 (cit. on p. 2575).
J. E. Andersen and S. Marzioni (2017). "Level N Teichmüller TQFT and complex Chern-Simons theory". In: *Quantum geometry of moduli spaces with applications to TQFT and RNA folding*. Vol. 25. Trav. Math. Fac. Sci. Technol. Commun. Univ. Luxemb., Luxembourg, pp. 97–146. MR: 3700061 (cit. on pp. 2575, 2576).
J. E. Andersen and J. J. K. Nissen (2017). "Asymptotic aspects of the Teichmüller TQFT". In: *Quantum geometry of moduli spaces with applications to TQFT and RNA folding*. Vol. 25. Trav. Math. Fac. Sci. Technol. Commun. Univ. Luxemb., Luxembourg, pp. 41–95. MR: 3700060 (cit. on p. 2575).
J. E. Andersen and K. Ueno (2007a). "Abelian conformal field theory and determinant bundles". *Internat. J. Math.* 18.8, pp. 919–993. MR: 2339577 (cit. on p. 2559).
– (2007b). "Geometric construction of modular functors from conformal field theory". *J. Knot Theory Ramifications* 16.2, pp. 127–202. MR: 2306213 (cit. on p. 2559).
– (2012). "Modular functors are determined by their genus zero data". *Quantum Topol.* 3.3-4, pp. 255–291. MR: 2928086 (cit. on p. 2559).
– (2015). "Construction of the Witten-Reshetikhin-Turaev TQFT from conformal field theory". *Invent. Math.* 201.2, pp. 519–559. MR: 3370620 (cit. on p. 2559).
M. Atiyah (1988). "Topological quantum field theories". *Inst. Hautes Études Sci. Publ. Math.* 68, 175–186 (1989). MR: 1001453 (cit. on p. 2559).
S. Axelrod, S. Della Pietra, and E. Witten (1991). "Geometric quantization of Chern-Simons gauge theory". *J. Differential Geom.* 33.3, pp. 787–902. MR: 1100212 (cit. on p. 2559).
D. Bar-Natan and E. Witten (1991). "Perturbative expansion of Chern-Simons theory with noncompact gauge group". *Comm. Math. Phys.* 141.2, pp. 423–440. MR: 1133274 (cit. on p. 2560).
V. V. Bazhanov, V. V. Mangazeev, and S. M. Sergeev (2007). "Faddeev-Volkov solution of the Yang-Baxter equation and discrete conformal symmetry". *Nuclear Phys. B* 784.3, pp. 234–258. MR: 2353414 (cit. on p. 2560).
– (2008). "Quantum geometry of 3-dimensional lattices". *J. Stat. Mech.* 0807, P07004. arXiv: 0801.0129 (cit. on p. 2560).

C. Blanchet, N. Habegger, G. Masbaum, and P. Vogel (1992). "Three-manifold invariants derived from the Kauffman bracket". *Topology* 31.4, pp. 685–699. MR: 1191373 (cit. on p. 2559).

– (1995). "Topological quantum field theories derived from the Kauffman bracket". *Topology* 34.4, pp. 883–927. MR: 1362791 (cit. on p. 2559).

A. Cassson (n.d.). Private communication (cit. on p. 2562).

R. Dijkgraaf, H. Fuji, and M. Manabe (2011). "The volume conjecture, perturbative knot invariants, and recursion relations for topological strings". *Nuclear Phys. B* 849.1, pp. 166–211. MR: 2795276 (cit. on p. 2560).

T. Dimofte (2013). "Quantum Riemann surfaces in Chern-Simons theory". *Adv. Theor. Math. Phys.* 17.3, pp. 479–599. MR: 3250765 (cit. on p. 2560).

T. Dimofte, D. Gaiotto, and S. Gukov (2014). "Gauge theories labelled by three-manifolds". *Comm. Math. Phys.* 325.2, pp. 367–419. MR: 3148093 (cit. on p. 2560).

T. Dimofte and S. Gukov (2013). "Chern-Simons theory and S-duality". *J. High Energy Phys.* 5, 109, front matter+65. MR: 3080552 (cit. on p. 2560).

T. Dimofte, S. Gukov, J. Lenells, and D. Zagier (2009). "Exact results for perturbative Chern-Simons theory with complex gauge group". *Commun. Number Theory Phys.* 3.2, pp. 363–443. MR: 2551896 (cit. on p. 2560).

L. D. Faddeev (1995). "Discrete Heisenberg-Weyl group and modular group". *Lett. Math. Phys.* 34.3, pp. 249–254. MR: 1345554 (cit. on pp. 2560, 2571).

L. D. Faddeev, R. M. Kashaev, and A. Y. Volkov (2001). "Strongly coupled quantum discrete Liouville theory. I. Algebraic approach and duality". *Comm. Math. Phys.* 219.1, pp. 199–219. MR: 1828812 (cit. on p. 2560).

V. V. Fock and L. O. Chekhov (1999). "Quantum Teichmüller spaces". *Teoret. Mat. Fiz.* 120.3, pp. 511–528. MR: 1737362 (cit. on p. 2560).

N. Geer, R. M. Kashaev, and V. G. Turaev (2012). "Tetrahedral forms in monoidal categories and 3-manifold invariants". *J. Reine Angew. Math.* 673, pp. 69–123. MR: 2999129.

S. Gukov (2005). "Three-dimensional quantum gravity, Chern-Simons theory, and the A-polynomial". *Comm. Math. Phys.* 255.3, pp. 577–627. MR: 2134725 (cit. on p. 2560).

S. Gukov and H. Murakami (2008). "$\mathrm{SL}(2,\mathbb{C})$ Chern-Simons theory and the asymptotic behavior of the colored Jones polynomial". *Lett. Math. Phys.* 86.2-3, pp. 79–98. MR: 2465747 (cit. on p. 2560).

K. Hikami (2001). "Hyperbolicity of partition function and quantum gravity". *Nuclear Phys. B* 616.3, pp. 537–548. MR: 1865275 (cit. on p. 2560).

– (2007). "Generalized volume conjecture and the A-polynomials: the Neumann-Zagier potential function as a classical limit of the partition function". *J. Geom. Phys.* 57.9, pp. 1895–1940. MR: 2330673 (cit. on p. 2560).

N. J. Hitchin (1990). "Flat connections and geometric quantization". *Comm. Math. Phys.* 131.2, pp. 347–380. MR: 1065677 (cit. on p. 2559).

L. Hörmander (1969). *Linear partial differential operators*. Third revised printing. Die Grundlehren der mathematischen Wissenschaften, Band 116. Springer-Verlag New York Inc., New York, pp. vii+288. MR: 0248435 (cit. on p. 2570).

– (2003). *The analysis of linear partial differential operators. I*. Classics in Mathematics. Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)]. Springer-Verlag, Berlin, pp. x+440. MR: 1996773.

R. M. Kashaev (1994). "Quantum dilogarithm as a $6j$-symbol". *Modern Phys. Lett. A* 9.40, pp. 3757–3768. MR: 1317945.

– (1997). "The hyperbolic volume of knots from the quantum dilogarithm". *Lett. Math. Phys.* 39.3, pp. 269–275. MR: 1434238 (cit. on p. 2575).

– (1998). "Quantization of Teichmüller spaces and the quantum dilogarithm". *Lett. Math. Phys.* 43.2, pp. 105–115. MR: 1607296 (cit. on p. 2560).

– (1999). "The Liouville central charge in quantum Teichmüller theory". *Tr. Mat. Inst. Steklova* 226.Mat. Fiz. Probl. Kvantovoĭ Teor. Polya, pp. 72–81. MR: 1782553.

– (2001). "On the spectrum of Dehn twists in quantum Teichmüller theory". In: *Physics and combinatorics, 2000 (Nagoya)*. World Sci. Publ., River Edge, NJ, pp. 63–81. MR: 1872252.

– (2012). "Discrete Liouville equation and Teichmüller theory". In: *Handbook of Teichmüller theory. Volume III*. Vol. 17. IRMA Lect. Math. Theor. Phys. Eur. Math. Soc., Zürich, pp. 821–851. MR: 2952777.

M. Lackenby (2000). "Word hyperbolic Dehn surgery". *Invent. Math.* 140.2, pp. 243–282. MR: 1756996 (cit. on p. 2562).

Y. Laszlo (1998). "Hitchin's and WZW connections are the same". *J. Differential Geom.* 49.3, pp. 547–576. MR: 1669720 (cit. on p. 2559).

R. C. Penner (1987). "The decorated Teichmüller space of punctured surfaces". *Comm. Math. Phys.* 113.2, pp. 299–339. MR: 919235 (cit. on p. 2560).

– (2012). *Decorated Teichmüller theory*. QGM Master Class Series. With a foreword by Yuri I. Manin. European Mathematical Society (EMS), Zürich, pp. xviii+360. MR: 3052157 (cit. on p. 2560).

M. Reed and B. Simon (1972). *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, New York-London, pp. xvii+325. MR: 0493419 (cit. on pp. 2569, 2570).

– (1975). *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, pp. xv+361. MR: 0493420 (cit. on p. 2570).

N. Y. Reshetikhin and V. G. Turaev (1990). "Ribbon graphs and their invariants derived from quantum groups". *Comm. Math. Phys.* 127.1, pp. 1–26. MR: 1036112 (cit. on p. 2559).

– (1991). "Invariants of 3-manifolds via link polynomials and quantum groups". *Invent. Math.* 103.3, pp. 547–597. MR: 1091619 (cit. on p. 2559).

I. Rivin (2003). "Combinatorial optimization in geometry". *Adv. in Appl. Math.* 31.1, pp. 242–271. MR: 1985831 (cit. on p. 2562).

G. B. Segal (1988). "The definition of conformal field theory". In: *Differential geometrical methods in theoretical physics (Como, 1987)*. Vol. 250. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Kluwer Acad. Publ., Dordrecht, pp. 165–171. MR: 981378 (cit. on p. 2559).

J. Teschner (2007). "From Liouville theory to the quantum geometry of Riemann surfaces". In: *Prospects in mathematical physics*. Vol. 437. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 231–246. MR: 2354665 (cit. on p. 2560).

A. Tsuchiya, K. Ueno, and Y. Yamada (1989). "Conformal field theory on universal family of stable curves with gauge symmetries". In: *Integrable systems in quantum field theory and statistical mechanics*. Vol. 19. Adv. Stud. Pure Math. Academic Press, Boston, MA, pp. 459–566. MR: 1048605 (cit. on p. 2559).

V. G. Turaev (1994). *Quantum invariants of knots and 3-manifolds*. Vol. 18. De Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, pp. x+588. MR: 1292673 (cit. on p. 2559).

V. G. Turaev and O. Y. Viro (1992). "State sum invariants of 3-manifolds and quantum $6j$-symbols". *Topology* 31.4, pp. 865–902. MR: 1191386 (cit. on p. 2559).

A. Y. Volkov (2005). "Noncommutative hypergeometry". *Comm. Math. Phys.* 258.2, pp. 257–273. MR: 2171695.

E. Witten (1988). "Topological quantum field theory". *Comm. Math. Phys.* 117.3, pp. 353–386. MR: 953828 (cit. on p. 2559).

– (1991). "Quantization of Chern-Simons gauge theory with complex gauge group". *Comm. Math. Phys.* 137.1, pp. 29–66. MR: 1099255 (cit. on p. 2560).

– (2011). "Analytic continuation of Chern-Simons theory". In: *Chern-Simons gauge theory: 20 years after*. Vol. 50. AMS/IP Stud. Adv. Math. Amer. Math. Soc., Providence, RI, pp. 347–446. MR: 2809462 (cit. on p. 2560).

Received 2017-12-08.

Jørgen Ellegaard Andersen
Center for Quantum Geometry of Moduli Spaces
Department of Mathematics
University of Aarhus
DK-8000
Denmark

andersen@qgm.au.dk

Rinat Kashaev
University of Geneva
2-4 rue du Lièvre, Case postale 64
1211 Genève 4
Switzerland
Rinat.Kashaev@unige.ch

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2585–2598)

SPECIAL GEOMETRY ON CALABI–YAU MODULI SPACES AND Q-INVARIANT MILNOR RINGS

Alexander Belavin

Abstract

The moduli spaces of Calabi–Yau (CY) manifolds are the special Kähler manifolds. The special Kähler geometry determines the low-energy effective theory which arises in Superstring theory after the compactification on a CY manifold. For the cases, where the CY manifold is given as a hypersurface in the weighted projective space, a new procedure for computing the Kähler potential of the moduli space has been proposed by Konstantin Aleshkin and myself. The method is based on the fact that the moduli space of CY manifolds is a marginal subspace of the Frobenius manifold which arises on the deformation space of the corresponding Landau–Ginzburg superpotential. I review this approach and demonstrate its efficiency by computing the Special geometry of the 101-dimensional moduli space of the quintic threefold around the orbifold point.

1 Introduction

To compute the low-energy Lagrangian of the string theory compactified on a CY manifold Candelas, Horowitz, Strominger, and Witten [1985], one needs to know the Special geometry of the corresponding CY moduli space Candelas, Green, and Hübsch [1989], Strominger [1990], Candelas, Green, and Hübsch [1990], and Candelas and de la Ossa [1991].

More precisely, the effective Lagrangian of the vector multiplets in the superspace contains $h^{2,1}$ supermultiplets. Scalars from these multiplets take value in the target space $\mathcal{M}$, which is a moduli space of complex structures on a CY manifold and is a special Kähler manifold. Metric $G_{a\bar{b}}$ and Yukawa couplings κ_{abc} on this space are given by the following

MSC2010: 32Q25.

formulae in the special coordinates z^a :

$$G_{a\bar{b}} = \partial_a \partial_{\bar{b}} K, \quad e^{-K} = -i \int_X \Omega \wedge \bar{\Omega},$$
$$\kappa_{abc} = \int_X \Omega \wedge \partial_a \partial_b \partial_c \Omega = \frac{\partial^3 F}{\partial z^a \partial z^b \partial z^c},$$

where

$$z^a = \int_{A_a} \Omega, \ \frac{\partial F}{\partial z^a} = \int_{B^a} \Omega$$

are the period integrals of the holomorphic volume form Ω on X. Here A_a and B^a form the symplectic basis in $H_3(X, \mathbb{Z})$.

We can rewrite the expression (1) for the Kähler potential using the periods as

$$e^{-K} = -i \Pi \Sigma \Pi^{\dagger}, \ \Pi = (\partial F, \ z),$$

where matrix $(\Sigma)^{-1}$ is an intersection matrix of cycles $A_a, \ B^a$ equal to the symplectic unit.

The computation of periods in the symplectic basis appears to be very non-trivial. It was firstly performed for the case of the quintic CY manifold in the distinguished paper Candelas, de la Ossa, Green, and Parkes [1991].
Here I present an alternative approach to the computation of Kähler potential for the case where CY manifold is given by a hypersurface $W(x, \phi) = 0$ in a weighted projective space. The approach is based on the connection of CY manifold with a Frobenius ring which arises on the deformations of the singularity defined by the superpotential $W_0(x)$ Lerche, Vafa, and N. P. Warner [1989], Martinec [1989], and Vafa and N. Warner [1989].

Let a CY manifold X be given as a solution of an equation

$$W(x, \phi) = W_0(x) + \sum_{s=1}^{h^{2,1}} \phi_s e_s(x) = 0$$

in some weighted projective space, where $W_0(x)$ is a quasihomogeneous function in $\mathbb{C}^5$ of degree d that defines an isolated singularity at $x = 0$. The monomials $e_s(x)$ also have degree d and are in a correspondence to deformations of the complex structure of X.

Polynomial $W_0(x)$ defines a Milnor ring R_0. Inside R_0 there exists a subring R_0^Q which is invariant under the action of the so-called quantum symmetry group Q that acts on $\mathbb{C}^5$ diagonally, and preserves $W(x, \phi)$. In many cases $\dim R_0^Q = \dim H^3(X)$ and the ring itself has a Hodge structure $R_0^Q = (R_0^Q)^0 \oplus (R_0^Q)^1 \oplus (R_0^Q)^2 \oplus (R_0^Q)^3$ in correspondence with the elements of degrees $0, d, 2d, 3d$.

Another important group is the subgroup of phase symmetries G, which acts diagonally on $\mathbb{C}^5$, commutes with the quantum symmetry Q and preserves $W_0(x)$. It acts naturally on the invariant ring R_0^Q, and this action respects the Hodge decomposition of R_0^Q. This allows to choose a basis $e_\mu(x)$ in each of the Hodge decomposition components of R_0^Q to be eigenvectors for the G group action.

On the ring R_0^Q we introduce the invariant pairing η. The pairing turns the ring to a Frobenius algebra Dubrovin [1992]. The pairing η plays an important for our construction of the explicit expression for the volume of the Calabi-Yau manifold.

Using the invariant ring R_0^Q and differentials $D_\pm = \mathrm{d} \pm \mathrm{d}W_0 \wedge$ we construct two $Q-$invariant cohomology groups $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$. These groups inherit the Hodge structure from R_0^Q. We can choose in $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$ the eigenbasises $e_\mu(x)\,\mathrm{d}^5x$ which are also invariant under the phase symmetry action.

As shown in Candelas [1988], elements of these cohomology groups are in correspondence with the harmonic forms of $H^3(X)$. This isomorphism allows to define the antilinear involution $*$ on the invariant cohomology $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$ that corresponds to the complex conjugation on the space of the harmonic forms of $H^3(X)$.

It turns out, that in the basis $e_\mu(x)$ it reads

$$*e_\mu(x)\,\mathrm{d}^5x = M_\mu^\nu e_\nu(x)\,\mathrm{d}^5x,\; M_\mu^\nu = \delta_{e_\mu \cdot e_\nu, e_\rho} A^\mu$$

where $e_\rho(x)$ is the unique element of degree $3d$ in R_0^Q, and $\delta_{e_\mu \cdot e_\nu, e_\rho}$ is 1 if $e_\mu \cdot e_\nu = e_\rho$ and 0 otherwise.

Having $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$ we define the relative invariant homology subgroups $\mathcal{H}_5^{\pm,inv} := H_5(\mathbb{C}^5, W_0 = L,\ \mathrm{Re}L \to \pm\infty)_{inv}$ inside the relative homology groups $H_5(\mathbb{C}^5, W_0 = L,\ \mathrm{Re}L \to \pm\infty)$. To do this we will use the oscillatory integrals and their pairing with elements of $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$. Using this pairing we define a cycle $\Gamma_\mu^\pm$ in the basis of relative invariant homology to be dual to $e_\mu(x)\,\mathrm{d}^5x$.

At last we define periods $\sigma_\mu^\pm(\phi)$ to be oscillatory integrals over the basis of cycles $\Gamma_\mu^\pm$. They are equal to periods of the holomorphic volume form Ω on X in a special basis of cycles of $H_3(X, \mathbb{C})$ with complex coefficients.

It follows from the phase symmetry invariance that in the chosen basis of cycles $\Gamma_\mu^\pm$ the formula for Kähler potential has the diagonal form:

$$e^{-K(\phi)} = \sum_\mu (-1)^{|\nu|} \sigma_\mu^+(\phi) A^\mu \overline{\sigma_\mu^-(\phi)}.$$

On the other hand, as shown in Aleshkin and A. Belavin [n.d.(a)], matrix $A = \text{diag}\{A^{\mu}\}$ is equal to the product of the matrix of the invariant pairing η in the Frobenius algebra R_0^Q and the real structure matrix M such that

$$e^{-K(\phi)} = \sum_{\mu,\nu} \sigma^{+}_{\mu}(\phi)\eta^{\mu\lambda}\, M^{\nu}_{\lambda}\overline{\sigma^{-}_{\mu}(\phi)}.$$

The real structure matrix is nothing but matrix M from (1). Using this we are able to explicitly compute the diagonal matrix elements A^{μ} and to obtain the explicit expression for the whole e^{-K}.

2 The special geometry on the CY moduli space

It was shown in in Strominger [1990], Candelas, Green, and Hübsch [1989, 1990], and Candelas and de la Ossa [1991] that the moduli space $\mathcal{M}$ of complex (or Kähler) structures of a given CY manifold is a special Kähler manifold.
Namely on $\mathcal{M}$ there exist so-called special (projective) coordinates $z^1 \cdots z^{n+1}$ and a holomorphic homogeneous function $F(z)$ of degree 2 in z, called a prepotential, such that the Kähler potential $K(z)$ of the moduli space metric is given by

$$e^{-K(z)} = \int_X \Omega \wedge \bar{\Omega} = z^a \cdot \frac{\partial \bar{F}}{\partial \bar{z}^{\bar{a}}} - \bar{z}^{\bar{a}} \cdot \frac{\partial F}{\partial z^a}$$

To obtain this formula, we choose Poincare dual symplectic basises $\alpha_a, \beta^b \in H^3(X,\mathbb{Z})$ and $A^a, B_b \in H_3(X,\mathbb{Z})$ and define the periods as

$$z^a = \int_{A^a} \Omega,\ F_b = \int_{B_b} \Omega.$$

Then using the Kodaira Lemma

$$\partial_a \Omega = k_a \Omega + \chi_a,$$

we can show that

$$F_a(z) = \frac{1}{2}\partial_a(F(z)),$$

where $F(z) = 1/2 z^b F_b(z)$.

Therefore, according to the definition (2) metric $G_{a\bar{b}} = \partial_a \bar{\partial}_{\bar{b}} K(z)$ is a special Kähler metric with prepotential $F(z)$ and with the special coordinates given by the period vector

$$\Pi = \left(F_\alpha, z^b\right)$$

we write the expression for the Kähler potential as

$$e^{-K(z)} = \Pi_\mu \Sigma^{\mu\nu} \bar{\Pi}_\nu,$$

where Σ is a symplectic unit, which is an inverse intersection matrix for cycles A^a and B_b.

Using formula (2), we can rewrite this expression in a basis of periods defined as integrals over arbitarary basisis of cycles $q_\mu \in H_3(X, \mathbb{Z})$

$$\omega_\mu = \int_{q_\mu} \Omega\,.$$

Such that

$$e^{-K} = \omega_\mu C^{\mu\nu} \bar{\omega}_\nu,$$

where $C^{\mu\nu}$ is the inverse marix of the intersection of the cycles q_μ.

So to find the Kähler potential, we must compute the periods over a basis of cycles on CY manifold and find their intersection matrix.

3 Hodge structure on the middle cohomology of the quintic

Now let us specialize to the case where X is a quintic threefold:

$$X = \{(x_1 : \cdots : x_5) \in \mathbb{P}^4 \mid W(x, \phi) = 0\},$$

and

$$W(x, \phi) = W_0(x) + \sum_{t=0}^{100} \phi_t e_t(x), \;\; W_0(x) = x_1^5 + x_2^5 + x_3^5 + x_4^5 + x_5^5$$

and $e_t(x)$ are the degree 5 monomials such that each variable has the power that is a non-negative integer less then four. Let us denote monomials $e_t(x) = x_1^{t_1} x_2^{t_2} x_3^{t_3} x_4^{t_4} x_5^{t_5}$ by its degree vector $t = (t_1, \cdots, t_5)$. Then there are precisely 101 of such monomials, which can be divided into 5 sets in respect to the permutation group S_5: $(1,1,1,1,1)$, $(2,1,1,1,0)$, $(2,2,1,0,0)$, $(3,1,1,0,0)$, $(3,2,0,0,0)$. In these groups there are correspondingly 1, 20, 30, 30, 20 different monomials. We denote $e_0(x) := e_{(1,1,1,1,1)}(x) = x_1x_2x_3x_4x_5$ to be the so-called fundamental monomial, which will be somewhat distinguished in our picture.

For this CY dim $H_3(X) = 204$ and period integrals have the form

$$\omega_\mu(x) = \int_{q_\mu} \frac{x_5\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3}{\partial W(x,\phi)/\partial x_4} = \int_{Q_\mu} \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_5}{W(x,\phi)},$$

where $q_\mu \in H_3(X,\mathbb{Z})$ and the corresponding cycles $Q_\mu \in H_5(\mathbb{C}^5 \backslash (W(x,\phi) = 0), \mathbb{Z})$.

Cohomology groups of the Kähler manifold X possess a Hodge structure $H^3(X) = H^{3,0}(X) \oplus H^{2,1}(X) \oplus H^{1,2}(X) \oplus H^{0,3}(X)$. Period integrals measure variation of the Hodge structure on $H^3(X)$ as the complex structure on X varies with ϕ.

This Hodge structure variation is in correspondence with a Frobenius ring which we will now describe.

4 Hodge structure on the invariant Milnor ring

Now we will consider $W_0(x)$ as an isolated singularity in $\mathbb{C}^5$ and the associated with it Milnor ring

$$R_0 = \frac{\mathbb{C}[x_1, \cdots, x_5]}{\langle \partial_i W_0 \rangle}.$$

We can choose its elements to be unique smallest degree polynomial representatives. For the quintic threefold X its Milnor ring R_0 is generated as a vector space by monomials where each variable has degree less than four, and dim $R_0 = 1024$.

Since the polynomial $W_0(x)$ is homogeneous one of the fifth degree it follows that $W_0(\alpha x_1, \ldots, \alpha x_5) = W_0(x_1, \ldots, x_5)$ for $\alpha^5 = 1$. This action preserves $W_0(x)$ and is trivial in the corresponding projective space and on X. Such a group with this action is called a *quantum symmetry* Q, in our case $Q \simeq \mathbb{Z}_5$. Q obviously acts on the Milnor ring R_0.

We define a subring R_0^Q to be a Q-invariant part of the Milnor ring

$$R_0^Q := \{e_\mu(x) \in R_0 \mid e_\mu(\alpha x) = e_\mu(x)\},\ \alpha^5 = 1.$$

R_0^Q is multiplicatively generated by 101 fifth-degree monomials $e_t(x)$ from (3) and consists of elements of degree $0, 5, 10$ and 15. The dimensions of the corresponding subspaces are $1, 101, 101$ and 1.

This degree filtration defines a Hodge structure on R_0^Q. Actually, R_0^Q is isomorphic to $H^3(X)$ and this isomorphism sends the degree filtration on R_0^Q to the Hodge filtration on $H^3(X)$ Candelas [1988].

Let us denote $\chi^i_{\bar{j}} = g^{i\bar{k}}\,\chi_{\bar{k}\bar{j}}$ as an extrinsic curvature tensor and $g_{i\bar{k}}$ is a metric for the hypersurface $W(x,\phi) = 0$ in $\mathbb{P}^4$. Then the isomorphism above can be written as a map from R_0^Q to closed differential forms in $H^3(X)$:

$$
\begin{aligned}
&1 \to \Omega_{ijk} \in H^{3,0}(X),\\
&e_\mu(x) \to e_\mu(x(y))\,\chi^l_{\bar{i}}\,\Omega_{ljk} \in H^{2,1}(X) \text{ if } |\mu| = 5,\\
&e_\mu(x) \to e_\mu(x(y))\,\chi^l_{\bar{i}}\,\chi^m_{\bar{j}}\,\Omega_{lmk} \in H^{1,2}(X) \text{ if } |\mu| = 10,\\
&e_\rho(x) = x_1^3x_2^3x_3^3x_4^3x_5^3 \to \chi^l_{\bar{i}}\,\chi^m_{\bar{j}}\,\chi^p_{\bar{k}}\,\Omega_{lmp} = \kappa\bar{\Omega} \in H^{0,3}(X)
\end{aligned}
$$

The details of this map can be found in Candelas [ibid.]. We also introduce the notation $e_\mu(x)$ for elements of the monomial basis of R_0^Q, where $\mu = (\mu_1, \cdots, \mu_5)$, $\mu_i \in \mathbb{Z}_+^5$, $e_\mu(x) = \prod_i x_i^{\mu_i}$ and the degree of $e_\mu(x)$ $\mu = \sum \mu_i$ is equal to zero module 5. In particular, $\rho = (3,3,3,3,3)$, that is $e_\rho(x)$ is the unique degree 15 element of R_0^Q.

The phase symmetry group $\mathbb{Z}_5^5$ acts diagonally on $\mathbb{C}^5$: $\alpha \cdot (x_1, \cdots, x_5) = (\alpha_1 x_1, \cdots, \alpha_5 x_5)$, $\alpha_i^5 = 1$. This action preserves $W_0 = \sum_i x_i^5$. The mentioned above quantum symmetry Q is a diagonal subgroup of the phase symmetries. Basis $\{e_\mu(x)\}$ consits of the eigenvectors of the phase symmetry and each $e_\mu(x)$ has a unique weight. Note that the action of the phase symmetry preserves the Hodge decomposition.

Another important fact is that on the invariant ring R_0^Q there exists a natural invariant pairing turning it into a Frobenius algebra Dubrovin [1992]:

$$\eta_{\mu\nu} = \operatorname{Res}\frac{e_\mu(x)\,e_\nu(x)}{\prod_i \partial_i W_0(x)}.$$

Up to an irrelevant constant for the monomial basis it is $\eta_{\mu\nu} = \delta_{\mu+\nu,\rho}$. This pairing plays a crucial role in our construction.

Let us introduce a couple of Saito differentials as in Aleshkin and A. Belavin [n.d.(a)] on differential forms on $\mathbb{C}^5$: $D_\pm = \mathrm{d} \pm \mathrm{d}W_0(x)\wedge$. They define two cohomology groups $H^*_{D_\pm}(\mathbb{C}^5)$. The cohomologies are only nontrivial in the top dimension $H^5_{D_\pm}(\mathbb{C}^5) \overset{J}{\simeq} R_0$.

The isomorphism J has an explicit description

$$J(e_\mu(x)) = e_\mu(x)\,\mathrm{d}^5x,\ e_\mu(x) \in R_0.$$

We see, that $Q = \mathbb{Z}_5$ naturally acts on $H^5_{D_\pm}(\mathbb{C}^5)$ and J sends the elements of Q-invariant ring R_0^Q to Q-invariant subspace $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$. Therefore, the latter space obtains the Hodge structure as well. Actually, this Hodge structure naturally corresponds to the Hodge structure on $H^3(X)$.

The complex conjugation acts on $H^3(X)$ so that $\overline{H^{p,q}(X)} = H^{q,p}(X)$, in particular $\overline{H^{2,1}(X)} = H^{1,2}(X)$. Through the isomorphism between R_0^Q and $H^3(X)$ the complex conjugation acts also on the elements of the ring R_0^Q as $*e_\mu(x) = p_\mu e_{\rho-\mu}(x)$, where $p_\mu p_{\rho-\mu} = 1$ and p_μ is a constant to be determined. In particular, differential form built from the linear combinations $e_\mu(x) + p_\mu e_{\rho-\mu}(x) \in H^3(X,\mathbb{R})$ is real.

5 Oscillatory representation and computation periods $\sigma_\mu(\phi)$

Relative homology groups $H_5(\mathbb{C}^5, W_0 = L,\ \mathrm{Re}L \to \pm\infty)$ have a natural pairing with Q-invariant cohomology groups $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$ defined as

$$\langle e_\mu(x)\mathrm{d}^5x, \Gamma^\pm\rangle = \int_{\Gamma^\pm} e_\mu(x)e^{\mp W_0(x)}\mathrm{d}^5x,\ H_5(\mathbb{C}^5, W_0 = L,\ \mathrm{Re}L \to \pm\infty).$$

Using this we introduce two Q-invariant homology groups[1] $\mathcal{H}_5^{\pm,inv}$ as quotient of $H_5(\mathbb{C}^5, W_0 = L,\ \mathrm{Re}L \to \pm\infty)$ with respect to the subgroups orthogonal to $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$. Now we introduce basises $\Gamma^\pm_\mu$ in the homology groups $\mathcal{H}_5^{\pm,inv}$ using the duality with the basises in $H^5_{D_\pm}(\mathbb{C}^5)_{inv}$:

$$\int_{\Gamma^\pm_\mu} e_\nu(x)e^{\mp W_0(x)}\mathrm{d}^5x = \delta_{\mu\nu}$$

and the corresponding periods

$$\sigma^\pm_{\alpha\mu}(\phi) := \int_{\Gamma^\pm_\mu} e_\alpha(x)e^{\mp W(x,\phi)}\mathrm{d}^5x,$$
$$\sigma^\pm_\mu(\phi) := \sigma^\pm_{0\mu}(\phi)$$

[1] We are grateful to V. Vasiliev for explaining to us the details about these homology groups and their connection with the middle homology of X.

which are understood as series expansions in ϕ around zero.
The periods $\sigma^{\pm}_{\mu}(\phi)$ satisfy the same differential equation as periods $\omega_{\mu}(\phi)$ of the holomorphic volume form on X. Moreover, these sets of periods span same subspaces as functions of ϕ. Therefore we can define cycles $Q^{\pm}_{\mu} \in \mathcal{H}^{\pm,inv}_{5}$ such that

$$\int_{Q^{\pm}_{\mu}} e^{\mp W(x,\phi)} \mathrm{d}^5 x = \int_{q_{\mu}} \Omega = \int_{Q_{\mu}} \frac{\mathrm{d}^5 x}{W(x,\phi)}.$$

So the periods $\omega^{\pm}_{\alpha\mu}(\phi)$ are given by the integrals over these cycles analogous to (5).

With these notations the idea of computation of periods A. Belavin and V. Belavin [2016]

$$\sigma^{\pm}_{\mu}(\phi) = \int_{\Gamma^{\pm}_{\mu}} e^{\mp W(x,\phi)}\, \mathrm{d}^5 x$$

can be stated as follows.
To explicitly compute $\sigma^{\pm}_{\mu}(\phi)$, first we expand the exponent in the integral (5) in ϕ representing $W(x,\phi) = W_0(x) + \sum_s \phi_s e_s(x)$

$$\sigma^{\pm}_{\mu}(\phi) = \sum_m \left(\prod_s \frac{(\pm\phi_s)^{m_s}}{m_s!}\right) \int_{\Gamma^{\pm}_{\mu}} \prod_s e_s(x)^{m_s}\, e^{\mp W_0(x)}\, \mathrm{d}^5 x.$$

We note, that $\sigma^{-}_{\mu}(\phi) = (-1)^{|\mu|}\sigma^{+}_{\mu}(\phi)$, so we focus on $\sigma_{\mu}(\phi) := \sigma^{+}_{\mu}(\phi)$.
For each of the summands in (5) the form $\prod_s e_s(x)^{m_s}\, \mathrm{d}^5 x$ belongs to $H^5_{D_{\pm}}(\mathbb{C}^5)_{inv}$, because it is $Q-$invariant. Therefore, we can expand it in the basis $e_{\mu}(x)\, \mathrm{d}^5 x \in H^5_{D_{\pm}}(\mathbb{C}^5)_{inv}$. Namely we can find such a polynomial $4-$form U, that

$$\prod_s e_s(x)^{m_s}\, \mathrm{d}^5 x = \sum_{\nu} C_{\nu}(m)\, e_{\nu}(x)\, \mathrm{d}^5 x + D_+ U.$$

In result we obtain for the integral in (5)

$$\int_{\Gamma^{\pm}_{\mu}} \prod_s e_s(x)^{m_s}\, e^{\mp W_0(x)}\, \mathrm{d}^5 x = C_{\mu}(m).$$

So from (5) we have

$$\sigma_{\mu}(\phi) = \sum_m \left(\prod_s \frac{\phi_s^{m_s}}{m_s!}\right) \int_{\Gamma^{+}_{\mu}} \prod_{s,i} x_i^{\sum_s m_s s_i}\, e^{-W_0(x)}\, \mathrm{d}^5 x.$$

We can rewrite the sum in the exponent of x_i as $\sum_s m_s s_i = 5n_i + \nu_i, \ \nu_i < 5$. Therefore we need to compute the coefficients c^m_ν in the equations

$$\prod x_i^{5n_i+\nu_i}\, \mathrm{d}^5x = \sum_\nu c^m_\nu\, e_\nu(x)\, \mathrm{d}^5x + D_+U.$$

Note that

$$D_+\left(\frac{1}{5}x_1^{5n+k-4}\, f(x_2,\cdots,x_5)\, \mathrm{d}x_2\wedge\cdots\wedge\mathrm{d}x_5\right) = \\ = \left[x_1^{5n+k} + \left(n+\frac{k-4}{5}\right)x_1^{5(n-1)+k}\right] f(x_2,\cdots,x_5)\, \mathrm{d}^5x$$

Therefore in D_+ cohomology we have

$$\prod_i x_i^{5n_i+\nu_i}\, \mathrm{d}^5x = -\left(n_1+\frac{\nu_1-4}{5}\right)x_1^{5(n_1-1)+\nu_1}\prod_{i=2}^{5} x_i^{5n_i+\nu_i}\, \mathrm{d}^5x, \ \nu_i < 5.$$

By induction we obtain

$$\prod_i x_i^{5n_i+\nu_i}\, \mathrm{d}^5x = (-1)^{\sum_i n_i}\prod_i\left(\frac{\nu_i+1}{5}\right)_{n_i}\prod_i x_i^{\nu_i}\, \mathrm{d}^5x, \ \nu_i < 5.$$

where $(a)_n = \Gamma(a+n)/\Gamma(a)$.
Using (5) once again, we see that if any $\nu_i = 4$ then the differential form is trivial and the integral is zero. Hence, rhs of (5) is proportional to $e_\nu(x)$ and gives the desired expression. Plugging (5) into (5) and integrating over Γ^+_μ we obtain the answer

$$\sigma_\mu(\phi) = \sigma^+_\mu(\phi) = \sum_{n_i\geq 0}\prod_i\left(\frac{\mu_i+1}{5}\right)_{n_i}\sum_{m\in\Sigma_n}\prod_s\frac{\phi_s^{m_s}}{m_s!},$$

where

$$\Sigma_n = \{m \mid \sum_s m_s s_i = 5n_i + \mu_i\}$$

Further we will also use the periods with slightly different normalization, which turn out to be convenient

$$\hat{\sigma}_\mu(\phi) = \prod_i \Gamma\left(\frac{\mu_i + 1}{5}\right) \sigma_\mu(\phi) = \sum_{n_i \geq 0} \prod_i \Gamma\left(n_i + \frac{\mu_i + 1}{5}\right) \sum_{m \in \Sigma_n} \prod_s \frac{\phi_s^{m_s}}{m_s!}.$$

6 Computation of the Kähler potential

Pick any basis $Q_\mu^\pm$ of cycles with integer or real coefficients as in (5). Then for the Kähler potential we have the formula

$$e^{-K} = \omega_\mu^+(\phi) C^{\mu\nu} \overline{\omega_\nu^-(\phi)}$$

in which the matrix $C^{\mu\nu}$ is related with the Frobenius pairing η as

$$\eta_{\alpha\beta} = \omega_{\alpha\mu}^+(0) C^{\mu\nu} \omega_{\beta\nu}^-(0).$$

The derivation of the last relation is given in Cecotti and Vafa [1991] and Chiodo, Iritani, and Ruan [2014].

Let also $T^\pm$ be the matrix that connects the cycles $Q_\mu^\pm$ and $\Gamma_\nu^\pm$. That is

$$Q_\mu^\pm = (T^\pm)_\mu^\nu \Gamma_\nu^\pm$$

. Then $M = (T^-)^{-1} \overline{T^-}$ is a real structure matrix, that is $M \bar{M} = 1$ and by construction M doesn't depend on the choice of basis $Q_\mu^\pm$. M is only defined by our choice of $\Gamma_\mu^\pm$.

In Aleshkin and A. Belavin [n.d.(a)] we deduced from (6) and (6) the formula

$$e^{-K(\phi)} = \sigma_\mu^+(\phi) \eta^{\mu\lambda} M_\lambda^\nu \overline{\sigma_\nu^-(\phi)} = \sigma_\mu A^{\mu\nu} \overline{\sigma_\nu},$$

where $\eta^{\mu\nu} = \eta_{\mu\nu} = \delta_{\mu,\rho-\nu}$.

Now we show that the matrix $A^{\mu\nu}$ in (6) is diagonal. To do this we extend the action of the phase symmetry group to the action $\mathcal{A}$ on the parameter space $\{\phi_s\}$ such that $W = W_0 + \sum_s \phi_s e_s(x)$ is invariant under this new action. It easy to see that each $e_s(x)$ has an unique weight under this group action. Action $\mathcal{A}$ can be compensated using the coordinate tranformation and therefore is trivial on the moduli space of the quintic (implying that point $W = W_0$ is an orbifold point of the moduli space).

In particular, $e^{-K} = \int_X \Omega \wedge \bar{\Omega}$ is $\mathcal{A}$ invariant. Consider

$$e^{-K} = \sigma_\mu A^{\mu\nu} \overline{\sigma_\nu}$$

as a series in ϕ_s and $\overline{\phi_t}$. Each monomial has a certain weight under $\mathcal{A}$. For the series to be invariant, each monomial must have weight 0. But weight of $\sigma_\mu\overline{\sigma_\nu}$ equals to $\mu-\nu$ and due to non-degeneracy of weights of σ_μ only the ones with $\mu=\nu$ have weight zero. Thus, (6) becomes

$$e^{-K}=\sum_{\mu} A^{\mu}|\sigma_\mu(\phi)|^2.$$

Moreover, the matrix A should be real and, because $A=\eta\cdot M$, $M\bar{M}=1$ and $\eta_{\mu\nu}=\delta_{\mu+\nu,\rho}$, we have

$$A^{\mu}\,A^{\rho-\mu}=1.$$

Monodromy considerations. To fix finally the real numbers A^{μ} we use monodromy invariance of e^{-K} around $\phi_0=\infty$. Pick some $t=(t_1,t_2,t_3,t_4,t_5)$ with $|t|=5$ and let $\phi_s|_{s\neq t,0}=0,$. We will consider only the first order in ϕ_t.
Then the condition that period $\sigma_\mu(\phi)$ contains only non-zero summands of the form $\phi_0^{m_0}\,\phi_t$ implies that $\mu=t+const\cdot(1,1,1,1,1)$ mod 5. For each t from the table below the only such possibilities are $\mu=t$ and $\mu=\rho-t'=(3,3,3,3,3)-t'$, where t' denotes a vector obtained from t by permutation (written explicitly in the table below) of its coordinates.

Therefore, in this setting (6) becomes

$$e^{-K}=\sum_{k=0}^{3} a_k|\hat{\sigma}_{(k,k,k,k,k)}|^2+a_t|\hat{\sigma}_t|^2+a_{\rho-t'}|\hat{\sigma}_{\rho-t'}|^2+O(\phi_t^2),$$

here we use periods $\hat{\sigma}$ from (5) and denote $a_t=A^t/\prod_i\Gamma((t_i+1)/5)^2$. And the coefficients a_k, $k=0,1,2,3$ are already known from Candelas, de la Ossa, Green, and Parkes [1991]. This expression has to be monodromy invariant under the transport of ϕ_0 around ∞. From the formula (5) we have

$$F_1=\hat{\sigma}_k(\phi_t,\phi_0)=g_t\phi_k\;F(a,b;a+b\,|\,(\phi_0/5)^5)+O(\phi_t^6),$$
$$F_2=\hat{\sigma}_{\rho-t'}(\phi_t,\phi_0)=g_{\rho-t'}\phi_t\;\phi_0^{1-a-b}\;F(1-a,1-b;2-a-b\,|\,(\phi_0/5)^5)+O(\phi_t^6),$$

where g_t, $g_{\rho-t'}$ are some constants. Explicitly for all different labels t

t	$\rho - t'$	(a, b)
(2,1,1,1,0)	(3,2,2,2,1)	(2/5,2/5)
(2,2,1,0,0)	(3,3,2,1,1)	(1/5,3/5)
(3,1,1,0,0)	(0,3,3,2,2)	(1/5,2/5)
(3,2,0,0,0)	(1,0,3,3,3)	(1/5,1/5)

When ϕ_0 goes around infinity

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = B \cdot \begin{pmatrix} F_1 \\ F_2 \end{pmatrix},$$

where

$$B = \frac{1}{is(a+b)} \begin{pmatrix} c(a-b) - e^{i\pi(a+b)} & 2s(a)s(b) \\ 2e^{2\pi i(a+b)} s(a)s(b) & e^{\pi i(a+b)}[e^{2\pi i a} + e^{2\pi i b} - 2]/2 \end{pmatrix}.$$

Here $c(x) = \cos(\pi x),\ s(x) = \sin(\pi x)$. It is straightforward to show the following

Proposition 1.

$$a_t|\hat{\sigma}_t|^2 + a_{\rho-t'}|\hat{\sigma}_{\rho-t'}|^2 = a_t \prod_i \Gamma\left(\frac{t_i+1}{5}\right)^2 |\sigma_t|^2 + a_{\rho-t'} \prod_i \Gamma\left(\frac{4-t_i}{5}\right)^2 |\sigma_{\rho-t'}|^2$$

is B-invariant iff $a_t = -a_{\rho-t'}$.

Due to symmetry we have $a_{\rho-t'} = a_{\rho-t}$ in each case. From (6) it follows that the product of the coefficients at $|\sigma_\mu|^2$ and $|\sigma_{\rho-\mu}|^2$ in the expression for e^{-K} should be 1:

$$A^{\rho-t'} \cdot A^t = a_{\rho-t'} \cdot a_t \prod_i \Gamma\left(\frac{t_i+1}{5}\right)^2 \Gamma\left(\frac{4-t_i}{5}\right)^2 = 1.$$

Due to reflection formula $a_t = \pm \prod_i \sin(\pi(t_i+1)/5)$ up to a common factor of π. The sign turns out to be minus for Kähler metric to be positive definite in the origin. Therefore

$$A^\mu = (-1)^{\deg(\mu)/5} \prod \gamma\left(\frac{\mu_i+1}{5}\right).$$

Finally the Kähler potential becomes

$$e^{-K(\phi)} = \sum_{\mu=0}^{203} (-1)^{\deg(\mu)/5} \prod \gamma\left(\frac{\mu_i+1}{5}\right) |\sigma_\mu(\phi)|^2,$$

where $\gamma(x) = \frac{\Gamma(x)}{\Gamma(1-x)}$.

7 Real structure on the cycles $\Gamma^{\pm}_{\mu}$

Let cycles $\gamma_\mu \in H_3(X)$ be the images of cycles Γ^{+}_{μ} under the isomorphism $\mathcal{H}_5^{+,inv} \simeq H_3(X)$.
Complex conjugation sends $(2,1)$-forms to $(1,2)$-forms. Similarly it extends to a mapping on the dual homology cycles γ_μ.

Lemma 1. *Conjugation of homology classes has the following form: $*\gamma_\mu = p_\mu \gamma_{\rho-\mu}$, where $\rho = (3,3,3,3,3)$ is a unique maximal degree element in the Milnor ring.*

Proof. We perform a proof for the cohomology classes represented by differential forms. For one-dimensional $H^{3,0}(X)$ and $H^{0,3}(X)$ it is obvious. Let

$$\Omega_{2,1} := e_t(x)\, \chi^l_{\bar{i}}\, \Omega_{ljk} \in H^{2,1}(X).$$

Any element from $H^{1,2}(X)$ is representable by a degree 10 polynomial $P(x)$ as follows from (4) as

$$\overline{\Omega_{2,1}} = \Omega_{1,2} := P(x)\, \chi^l_{\bar{i}}\, \chi^m_{\bar{j}}\, \Omega_{lmk} \in H^{1,2}(X).$$

The group of phase symmetries modulo common factor acts by isomorphisms on X. Therefore, it also acts on the differential forms. Lhs and rhs of the previous equation should have the same weigth under this action, and weight of the lhs is equal $-t$ modulo $(1,1,1,1,1)$. It follows that $P(x) = p_t\, e_{\rho-t}(x)$ with some constant p_t. $\square$

Using this lemma and applying the complex conjugation of cycles to the formula (6) to obtain

$$e^{-K} = \sum_\mu A^\mu |\sigma_\mu|^2 = \sum_\mu p_\mu^2 A^\mu\, |\sigma_{\rho-\mu}|^2,$$

it follows that $A^\mu = \pm 1/p_\mu$. Now formula (6) implies

$$p_\mu = \prod \gamma\left(\frac{4-\mu_i}{5}\right).$$

8 Conclusion

I am grateful first of all to my coathor K. Aleshkin for the interesting collaboration. This talk is based on the joint work with K. Aleshkin. Also I am thankful to M. Bershtein, V. Belavin, S.Galkin, D. Gepner, A. Givental, M. Kontsevich, A. Okounkov, A. Rosly, V. Vasiliev for the useful discussions. The work has been performed for FASO budget project No. 0033-2018-0006.

References

Konstantin Aleshkin and Alexander Belavin (n.d.[a]). "A new approach for computing the geometry of the moduli spaces for a Calabi-Yau manifold". arXiv: 1706.05342 (cit. on pp. 2588, 2591, 2595).

– (n.d.[b]). "Special geometry on the 101 dimesional moduli space of the quintic threefold". arXiv: 1710.11609.

– (2018). "Special geometry on the moduli space for the two-moduli non-Fermat Calabi–Yau". *Phys. Lett. B* 776, pp. 139–144. MR: 3741463.

Alexander Belavin and Vladimir Belavin (2016). "Flat structures on the deformations of Gepner chiral rings". *J. High Energy Phys.* 10, 128, front matter+17. MR: 3578553 (cit. on p. 2593).

P. Candelas (1988). "Yukawa couplings between $(2, 1)$-forms". *Nuclear Phys. B* 298.3, pp. 458–492. MR: 928307 (cit. on pp. 2587, 2591).

P. Candelas, Gary T. Horowitz, Andrew Strominger, and Edward Witten (1985). "Vacuum configurations for superstrings". *Nuclear Phys. B* 258.1, pp. 46–74. MR: 800347 (cit. on p. 2585).

Philip Candelas, Paul S. Green, and Tristan Hübsch (1989). "Finite distance between distinct Calabi-Yau manifolds". *Phys. Rev. Lett.* 62.17, pp. 1956–1959. MR: 991081 (cit. on pp. 2585, 2588).

– (1990). "Rolling among Calabi-Yau vacua". *Nuclear Phys. B* 330.1, pp. 49–102. MR: 1038473 (cit. on pp. 2585, 2588).

Philip Candelas and Xenia C. de la Ossa (1991). "Moduli space of Calabi-Yau manifolds". *Nuclear Phys. B* 355.2, pp. 455–481. MR: 1103074 (cit. on pp. 2585, 2588).

Philip Candelas, Xenia C. de la Ossa, Paul S. Green, and Linda Parkes (1991). "A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory". *Nuclear Phys. B* 359.1, pp. 21–74. MR: 1115626 (cit. on pp. 2586, 2596).

Sergio Cecotti and Cumrun Vafa (1991). "Topological–anti-topological fusion". *Nuclear Phys. B* 367.2, pp. 359–461. MR: 1139739 (cit. on p. 2595).

Alessandro Chiodo, Hiroshi Iritani, and Yongbin Ruan (2014). "Landau-Ginzburg/Calabi-Yau correspondence, global mirror symmetry and Orlov equivalence". *Publ. Math.*

Inst. Hautes Études Sci. 119, pp. 127–216. arXiv: 1201.0813. MR: 3210178 (cit. on p. 2595).

B. Dubrovin (1992). "Integrable systems in topological field theory". *Nuclear Phys. B* 379.3, pp. 627–689. MR: 1175565 (cit. on pp. 2587, 2591).

Wolfgang Lerche, Cumrun Vafa, and Nicholas P. Warner (1989). "Chiral rings in $N = 2$ superconformal theories". *Nuclear Phys. B* 324.2, pp. 427–474. MR: 1025424 (cit. on p. 2586).

Emil J. Martinec (1989). "Algebraic geometry and effective Lagrangians". *Phys. Lett. B* 217.4, pp. 431–437. MR: 981536 (cit. on p. 2586).

Andrew Strominger (1990). "Special geometry". *Comm. Math. Phys.* 133.1, pp. 163–180. MR: 1071240 (cit. on pp. 2585, 2588).

Cumrun Vafa and Nicholas Warner (1989). "Catastrophes and the classification of conformal theories". *Physics Letters B* 218.1, pp. 51–58 (cit. on p. 2586).

Received 2017-12-14.

Alexander Belavin
L.D. Landau Institute for Theoretical Physics
Akademika Semenova av. 1-A
Chernogolovka, 142432 Moscow
Russia
sashabelavin@gmail.com

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2599–2614)

INTEGRABLE COMBINATORICS

PHILIPPE DI FRANCESCO

Abstract

We explore various combinatorial problems mostly borrowed from physics, that share the property of being continuously or discretely integrable, a feature that guarantees the existence of conservation laws that often make the problems exactly solvable. We illustrate this with: random surfaces, lattice models, and structure constants in representation theory.

1 Introduction

In this note we deal with combinatorial *objects*, mostly provided by physical systems or models. These are: random surfaces, lattice models, and structure constants. We will illustrate how to solve the various problems, mostly of exact or asymptotic enumeration, via a panel of techniques borrowed from pure combinatorics as well as statistical physics. The *tools* utilized are: generating functions, transfer matrices, bijections, matrix integrals, determinants, field theory, etc.

We have organized this collection of problems according to some common or analogous properties, essentially related to their underlying symmetries. Among them the most powerful is the notion of *integrability*. The latter appears under many different guises. The first form is continuous: Existence of conservation laws, flat connections, commuting transfer matrices, links to the Yang-Baxter equation, infinite dimensional algebra symmetries. The second form is discrete: Existence of discrete integrals of motion in discrete time.

What kind of results did we obtain? Solving a system completely usually entails a complete understanding of correlation functions within the model. This can be achieved by explicit diagonalization of the transfer matrix or Hamiltonians, explicit computation of generating functions, or derivation of complete systems of equations for averaged quantities. As usual in statistical physics, one also investigates the asymptotic (or thermodynamic) properties of the systems, leading to such results as asymptotic enumeration,

I acknowledge support by the Morris and Gertrude Fine foundation.
MSC2010: primary 05Axx; secondary 82B23, 05E05, 05E10, 82B41.

identification of phases and their separations, identification of underlying field theoretical descriptions of fluctuations.

One of the main features common to all the problems listed above is some kind of connection to discrete *paths* or *trees*, the two simplest and most fundamental combinatorial objects. The constructs of this note place these two main characters in new non-standard contexts which shed some new light on their deep significance. Together they form the basis of the notion of combinatorial integrability, i.e. the properties shared by combinatorial problems that connect them to discrete or continuous integrable systems.

The paper is organized as follows. In Section 2, we explore discretized models of random surfaces, whether Lorentzian in 1+1 dimensions (Section 2.1), or Euclidian in 2 dimensions (Section 2.2). Both type of models display integrability respectively via commuting transfer matrices and discrete integrals of motion, which allows to solve them explicitly.

In Section 3, we first describe the 6 vertex model and its many combinatorial wonders (Sect. 3.1), among which a description of Alternating Sign Matrices (ASM), and links to special types of plane partitions, as well as the geometry of nilpotent matrix varieties.

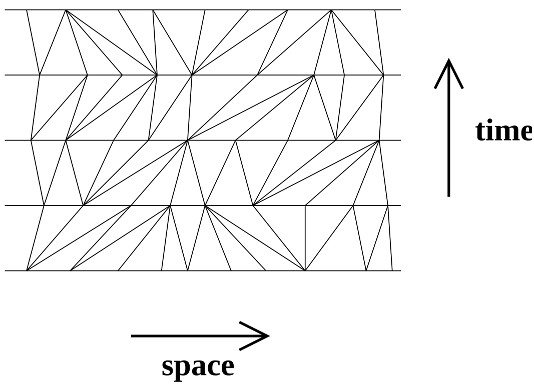
Section 4 focuses on Lie algebraic structures with a description of Whittaker vectors (Section 4.1) using path models, and of graded multiplicities in tensor products occurring in inhomogeneous quantum spin chains with Lie symmetry (Section 4.2). The description of the latter involves a construction of difference operators that generalize the celebrated Macdonald operators, and can be understood within the context of polynomial representations of Double-Affine Hecke Algebras (DAHA), quantum toroidal algebras, and Elliptic Hall Algebras (EHA).

Finally we gather some important open problems in Section 5, which we think should shape the future of integrable combinatorics.

Acknowledgments. Much of the work described in this note is the fruit of collaborations with various authors: R. Behrend, O. Golinelli, E. Guitter, C. Itzykson, R. Kedem, M. Lapa, N. Reshetikhin, R. Soto Garrido, P. Zinn-Justin, J.-B. Zuber. I also benefited from many important discussions with fellow mathematicians and physicists: O. Babelon, A. Borodin, J.-E. Bourgine, I. Cherednik, I. Corwin, S. Fomin, M. Gekhtmann, M. Jimbo, R. Kenyon, M. Kontsevich, Y. Matsuo, G. Musiker, A. Negut, A. Okounkov, V. Pasquier, M. Shapiro, O. Schiffman, V. Retakh.

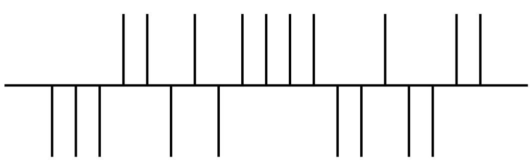
2 Random surfaces

2.1 1+1-dimensional Lorentzian triangulations and (continuous) integrability. Lorentzian triangulations Di Francesco, Guitter, and Kristjansen [2000] are used as a discrete model for quantum gravity in one (space)+1 (time) dimension. Pure gravity deals with fluctuations of such bare space-times, while matter theories include for instance particle systems in interction defined on such space-times. General relativity expresses the relation between those fluctuations and in particular the associated fluctuations of the metric, area and curvature of the space-time and the matter stress tensor. The model for a fluctuating 1+1D space-time is an arrangement of triangles organized into time slices as depicted below:



Fluctuations of space are represented by random arrangements of triangles in each time slice, while the time direction remains regular. These triangulations are best described in the dual picture by considering triangles as vertical half-edges and pairs of triangles that share a time-like (horizontal) edge as vertical edges between two consecutive time-slices. We may now concentrate on the transition between two consecutive time-slices which typically looks like:

(2-1)



with say i half-edges on the bottom and j on the top (here for instance we have $i = 9$ and $j = 10$). To take into account the *area* and *curvature* of space-time, we may introduce a Boltzmann weight g per triangle (i.e. per trivalent vertex in the dual picture) and a weight a per pair of consecutive triangles in a time-slice pointing in the same direction (both up or both down). The total weight of a configuration is the product of all local weights pertaining ot it. It is easy to see that these weights correspond to a transfer operator $\mathcal{T}(g, a)$ which describes the configurations of one time-slice with a total i of up-pointing triangles and j of down-pointing ones. The matrix element between states i and j reads:

$$T(g,a)_{i,j} = (ag)^{i+j} \sum_{k=0}^{\mathrm{Min}(i,j)} \binom{i}{k}\binom{j}{k} a^{-2k} \qquad (i,j \geq 0)$$

Equivalently, the double generating function for matrix elements of $\mathcal{T}(g,a)$ reads:

$$f_{T(g,a)}(z,w) = \sum_{i,j\geq 0} T(g,a)_{i,j} z^i w^j = \frac{1}{1-ga(z+w)-g^2(1-a^2)zw} \tag{2-2}$$

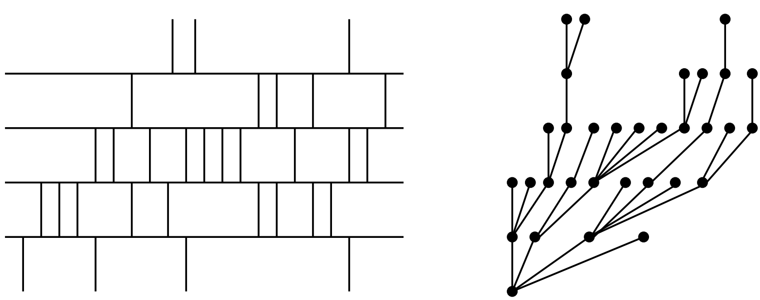
This model turns out to provide one of the simplest examples of quantum integrable system, with an infinite family of commuting transfer matrices. Indeed, we have:

Theorem 2.1 (Di Francesco, Guitter, and Kristjansen [2000])**.** *The transfer matrices $T(g,a)$ and $T(g',a')$ commute if and only if the parameters (g,a,g',a') are such that $\varphi(g,a)=\varphi(g',a')$ where:*

$$\varphi(g,a) = \frac{1-g^2(1-a^2)}{ag}$$

This and the explicit generating function (2-2) were extensively used in Di Francesco, Guitter, and Kristjansen [ibid.] to diagonalize $\mathcal{T}(g,a)$ and to compute correlation functions of boundaries in random Lorentzian triangulations.

For suitable choices of boundary conditions, the dual random Lorentzian triangulations introduced above may be viewed as random plane trees. This is easily realized by gluing all the bottom vertices of successive parallel vertical edges (no interlacing with the neighboring time slices). A typical such example reads:



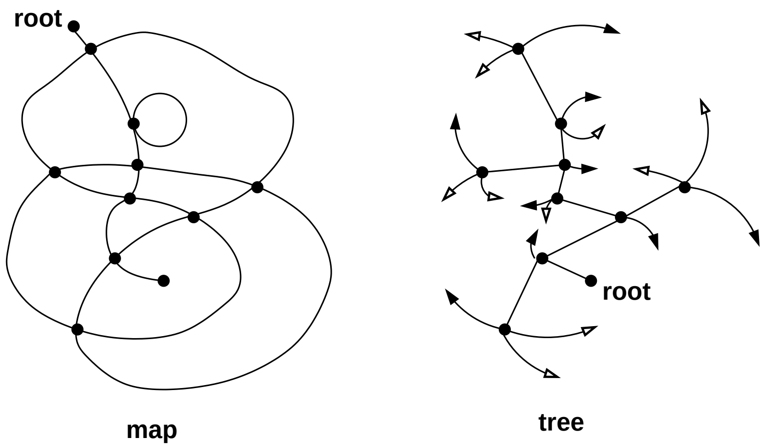
Note that the tree is naturally rooted at its bottom vertex.

To summarize, we have unearthed some integrable structure attached naturally to plane trees, one of the most fundamental objects of combinatorics. Note that in tree language the weights are respectively g^2 per edge, and a per pair of consecutive descendent edges and per pair of consecutive leaves at each vertex (from left to right).

2.2 2-dimensional Euclidian tessellations and (discrete) integrability. As opposed to Lorentzian gravity, the 2D Euclidian theory involves fluctuations of both space and time, allowing for space-times that look like random surfaces of arbitrary genus. Those are discretized by tessellations. A powreful tool for enumerating those maps was provided by matrix integrals, allowing to keep track of the area, as well as the genus via the size N of the matrices (see Ref. Di Francesco, Ginsparg, and J. Zinn-Justin [1995] and references therein). In a parallel way, the field-theoretical descriptions of the (critical) continuum limit of two-dimensional quantum gravity (2DQG) have blossomed into a more complete picture with identification of relevant operators and computation of their correlation functions Di Francesco and Kutasov [1991]. This was finally completed by an understanding in terms of the intersection theory of the moduli space of curves with punctures and fixed genus Kontsevich [1991]. Remarkably, in all these approaches a common integrable structure is always present. It takes the form of commuting flows in parameter space. However, a number of issues were left unadressed by the matrix/field theoretical approaches. What about the intrinsic geometry of the random surfaces? Correlators must be integrated w.r.t. the position of their insertions, leaving us only with topological invariants of the surfaces. But how to keep track for instance of the geodesic distances between two marked points of a surface, while at the same time summing over all surface fluctuations?

Answers to these questions came from a better combinatorial understanding of the structure of the (planar) tessellations involved in the discrete models. And, surprisingly, yet another form of integrability appeared. Following pioneering work of Schaeffer [1997], it was observed that all models of discrete 2DQG with a matrix model solution (at least in genus 0) could be expressed as statistical models of (decorated) trees, and moreover, the decorations allowed to keep track of geodesic distances between some faces of the tessellations. Marked planar tessellations are known as rooted planar maps in combinatorics. They correspond to connected graphs (with vertices, edges, faces) embedded into the Riemann sphere. Such maps are usually represented on a plane with a distinguished face “at infinity”, and a marked edge adjacent to that face. The degree of a vertex is the number of distinct half-edges adjacent to it, the degree of a face is the number of edges forming its boundary.

Consider the example of tetravalent (degree 4) planar maps with 2 univalent (degree 1) vertices, one of which is singled out as the root. The Schaeffer bijection associates to each of these a unique rooted tetravalent (with inner vertices of degree 4) tree called blossom-tree, with two types of leaves (black and white), and such that there is exactly one black leaf attached to each inner vertex:



This is obtained by the following cutting algorithm: travel clockwise along the bordering edges of the face at infinity, starting from the root. For each traversed edge, cut it if and only if after the cut, the new graph remains connected, and replace the two newly formed half-edges by a black and a white leaf respectively in clockwise order. Once the loop is traveled, this has created a larger face at infinity. Repeat the procedure until the graph has only one face left: it is the desired blossom-tree, which we reroot at the other univalent vertex, while the original root is transformed into a white leaf.

This bijection allows to keep track of the geodesic distance between the 2 univalent vertices. Defining $R_n(g)$ to be the generating function for maps with geodesic distance $\leq n$ between the two univalent vertices, we have the following recursion relation Bouttier, Di Francesco, and Guitter [2003]:

(2-3) $$R_n(g) = 1 + gR_n(g)\left(R_{n+1}(g) + R_n(g) + R_{n-1}(g)\right)$$

easily derived by inspecting the environment of the vertex attached to the root of the tree when it exists. It is supplemented by boundary conditions $R_{-1}(g) = 0$ and $\lim_{n\to\infty} R_n(g) = R(g) = \frac{1-\sqrt{1-12g}}{6g}$, the generating function of maps with no geodesic distance constraint.

Equation (2-3), viewed as governing the evolution of the quantity $R_n(g)$ in the discrete time variable n, is a classical *discrete integrable system*. By this we mean that it has a *discrete integral of motion*, expressed as follows. The function $\phi(x, y)$ defined by

(2-4) $$\phi(x, y) = xy(1 - g(x + y)) - x - y$$

is such that for any solution S_n of the recursion relation (2-3), the quantity $\phi(S_n, S_{n+1})$ is independent of n. In other words, the quantity $\phi(S_n, S_{n+1})$ is conserved modulo (2-3).

(This is easily shown by factoring $\phi(S_n, S_{n+1}) - \phi(S_{n-1}, S_n)$.). This conservation law gives in particular a relation of the form:

$$\phi(R_n(g), R_{n+1}(g)) = \lim_{m\to\infty} \phi(R_m(g), R_{m+1}(g)) = \phi(R(g), R(g))$$

. It turns out that we can solve explicitly for $R_n(g)$:

Theorem 2.2 (Bouttier, Di Francesco, and Guitter [ibid.]). *The generating function $R_n(g)$ for rooted tetravalent planar maps with two univalent vertices at geodesic distance at most n from each other reads:*

$$R_n(g) = R(g)\frac{(1-x(g)^{n+1})(1-x(g)^{n+4})}{(1-x(g)^{n+2})(1-x(g)^{n+3})}$$

where $x(g)$ is the unique solution of the equation: $x + \frac{1}{x} + 4 = \frac{1}{gR(g)^2}$ with a power series expansion of the form $x(g) = g + O(g^2)$.

The form of the solution in Theorem 2.2 is that of a discrete soliton with tau-function $\tau_n = 1 - x(g)^n$. Imposing more general boundary conditions on the equation (2-3) leads to elliptic solutions of the same flavor. The solution above and its generalizations to many classes of planar maps Di Francesco [2005] have allowed for a better understanding of the critical behavior of surfaces and their intrinsic geometry. Recent developments include planar three-point correlations, as well as higher genus results.

To summarize, we have seen yet another integrable structure emerge in relation to (decorated) trees. This is of a completely different nature from the one discussed in Section 2.1, where a quantum integrable structure was attached to rooted planar trees. Here we have a discrete classical integrable system, with soliton-like solutions.

3 Lattice models

3.1 The six-vertex model and beyond. The Six Vertex (6V) model is the archetypical example of 2D integrable lattice model. It is defined on domains of the square lattice $\mathbb{Z}^2$, with configurations obtained by orienting all the nearest neigbor edges in such a way that there are exactly to ingoing and two outgoing edges incident to each vertex in the interior of the domain (ice rule). This gives rise to $\binom{4}{2} = 6$ local vertex configurations, to which one usually attaches Boltzmann weights. The integrability of the model becomes manifest if we parametrize these weights with rapidities (spectral parameters) that are derived from the relevant R-matrix solution of the Yang-Baxter equation. This ensures that the system has an infinite set of commuting transfer matrices, similarly to Section 2.1. This property ensures that the transfer matrix is explicitly diagonalizable by means of Bethe Ansatz

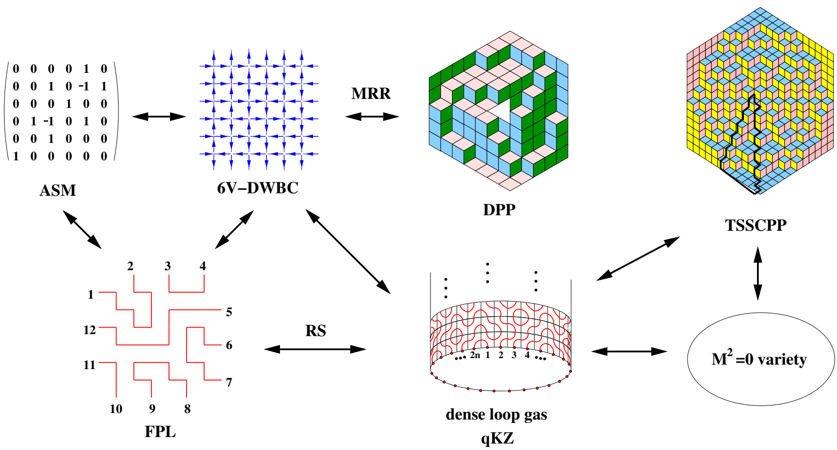


Figure 1: The combinatorial family of ASMs. From left to right: ASM, 6V-DWBC and FPL , all in bijection; dense O(n) loop gas: its groundstate/limiting probability vector satisfies the qKZ equation, the components measure FPL correlations (RS conjecture); DPP: their refined evaluation matches that of ASMs; TSSCPP: their refined enumeration matches a sum rule for qKZ solutions at generic q and $z_i = 1$; Variety $M^2 = 0$: its degree/multidegree matches solutions of qKZ for $q = 1$.

techniques. Note that a certain limit of the transfer matrix yields the Hamiltonian of the anisotropic XXZ spin chain.

A remarkable web of connections between many combinatorial objects relates to the 6V model, as shown in Figure 1. The configurations of the 6V model on a square grid of size n with the so-called Domain Wall Boundary Conditions (DWBC) that all boundary horizontal arrows are pointed towards the grid while all boundary vertical arrows point outward, are in bijection with Alternating Sign Matrices (ASM), namely matrices with entries in $\{-1, 0, 1\}$ with alternance of 1's and -1's along each row and column, and with row and column sums all equal to 1. This observation allowed Kuperberg [1996] to come up with an elegant proof of the ASM conjecture for the number of $n \times n$ ASMs, soon after the combinatorial proof of Zeilberger [1996]. Another bijection related the configurations of the 6V model with DWBC to so-called Fully Packed Loops (FPL) obtained by coloring edges of the square grid in such a way that exactly 2 edges incident to each inner vertex are colored, while every other boundary edge is colored. The colored edges form closed loops or open paths connecting boundary edges by pairs (such a pattern of connections is equivalent to non-crossing partitions or link patterns). The latter remark prompted the celebrated Razumov and Stroganov [2004] conjecture that FPL configurations with prescribed boundary edge connections form the Perron-Frobenius eigenvector of the XXZ spin chain at its *combinatorial point* (when all Boltzmann weights are 1), when expressed in the link pattern basis (in the O(n) model formulation of the spin chain), later proved by Cantini and Sportiello [2011]. Among the many developments around the conjecture, we used the link between the combinatorial problem and solutions of the quantum Knizhnik-Zamolodchikov (qKZ) equation for the O(n) model Di Francesco and P. Zinn-Justin [2005] and P. Zinn-Justin and Di Francesco [2008], which led us to connections with the geometry of the variety of square zero matrices Di Francesco and P. Zinn-Justin [2006]. Beyond bijections other sets of combinatorial objects have the same cardinality A_n. These are the Totally Symmetric Self-Complementary Plane Partitions (TSSCPP) on one hand and the Descending Plane Partitions (DPP) on the other. Both classes of objects can be formulated as the rhombus tilings of particular domains of the triangular lattice with particular symmetries. We found a proof Behrend, Di Francesco, and P. Zinn-Justin [2012] and Behrend, Di Francesco, and P. Zinn-Justin [2013] of the Mills-Robbins-Rumsey refined ASM-DPP conjecture Mills, Robbins, and Rumsey [1983] using generating functions similar to (2-2), however no bijection is known to this day.

It turns out that, among other formulations, the 6V model with DWBC may be expressed as a model of osculating paths, namely non-intersecting paths with unit steps $(1, 0)$ and $(0, 1)$, from the W border edges to the N border edges of the grid with allowed "kissing" or *osculating* vertices visited by two paths that do not cross. The latter path formulation allows to predict the arctic curve phenomenon (i.e. the sharp separation between ordered and disordered phases) for random ASMs Colomo and Sportiello [2016] as well as for random ASMs with a vertical reflection symmetry Di Francesco and Lapa [2018].

4 Lie algebras, quantum spin chains and CFT

In this section, we present combinatorial problems/approaches to algebra representation theory.

4.1 Whittaker vectors and path models. Whittaker vectors Kostant [1978] are fundamental objects in the representation theory of Lie algebras expressed in terms of Chevalley generators $\{e_i, f_i, h_i\}_{i=\epsilon}^r$ and relations (here $\epsilon = 1$ for finite algebras, and $\epsilon = 0$ for affine algebras). They are instrumental in constructing Whittaker functions, which are eigenfunctions for the quantum Toda operators, namely Schroedinger operators with kinetic and potential terms coded by the root system of the algebra. Given a Verma module $V_\lambda = \mathfrak{U}(\{f_i\})|\lambda\rangle$ with highest weight vector $|\lambda\rangle$, a Whittaker vector $v_{\mu,\lambda}$ with parameters μ_i is an element of the completion of V_λ (an infinite series in V_λ) that satisfies the relations $e_i v_{\lambda,\mu} = \mu_i v_i$ for all i. It is unique upon a choice of normalization. In Ref.Di Francesco, Kedem, and Turmunkh [2017] we developed a general approach to the computation of Whittaker vectors by expanding them on the "words" of the form $f_{i_1} f_{i_2} \cdots f_{i_k}|\lambda\rangle$ for arbitrary $i_j \in [\epsilon, r]$ and $k \geq 0$. The latter are of course not linearly independent, but we found some extremely nice and simple expression for their coefficients $c_{i_1,\ldots,i_k}$ in the expression of $v_{\lambda,\mu} = \sum_{k\geq 0} \sum_{i_1,\ldots,i_k\in[\epsilon,r]} c_{i_1,\ldots,i_k} f_{i_1} f_{i_2} \cdots f_{i_k}|\lambda\rangle$ (the normalization is chosen so that the empty word has coefficient $c_{0,0,\ldots,0} = 1$). We made the observation that the set of vectors of the form $f_{i_1} f_{i_2} \cdots f_{i_k}|\lambda\rangle$ is in bijection with the set of paths $p \in \mathcal{P}$ on the positive cone Q_+ of the root lattice, from the origin to some root $\beta = (\beta_i)_{i\in[\epsilon,r]}$ where β_i is the number of occurrences of f_i in the vector (or of the letter i in the word). Indeed, the steps of p are taken successively in the directions $i_k, i_{k-1}, \ldots, i_1$ in $[\epsilon, r]$. We denote by $|p\rangle = f_{i_1} f_{i_2} \cdots f_{i_k}|\lambda\rangle$. We have the following general result.

Theorem 4.1 (Di Francesco, Kedem, and Turmunkh [ibid.]). *For finite or affine Lie algebras, the Whittaker vector $v_{\lambda,\mu}$ is expressed as:*

$$v_{\lambda,\mu} = \sum_{\beta\in Q_+} \prod_i \mu_i^{\beta_i} \sum_{\text{paths } p:0\to\beta} w(p)|p\rangle$$

where the weight $w(p)$ is a product of local weights:

$$w(p) = \prod_{\substack{\gamma\in Q_+^* \\ \gamma \text{ vertex of } p}} \frac{1}{v(\gamma)}, \qquad v(\gamma) = (\lambda+\rho|\gamma) - \frac{1}{2}(\gamma|\gamma)$$

This construction was shown in Di Francesco, Kedem, and Turmunkh [ibid.] to extend to the A type quantum algebras $\mathcal{U}_q(\mathfrak{sl}_{r+1})$ with local weights depending on both the vertex and the direction of the step from the vertex.

This new formulation of Whittaker vectors yields a very simple proof for the fact that the corresponding Whittaker function obeys the quantum Toda equation (classical case), the Lamé-like deformed Toda equation (affine, non-critical case) or the q-difference Toda equation (quantum case).

The q-Whittaker functions are known to be a degenerate limit of Macdonald polynomials, when $t \to 0$ or ∞. This suggests to look for a possible path formulation of Macdonald polynomials.

4.2 Fusion product, Q-system cluster algebra and Macdonald theory.

4.2.1 Graded characters and quantum Q-system. We now turn to the combinatorial problem of finding the fusion coefficients $\mathrm{Mult}_q(\otimes KR_{\alpha,n}^{\otimes n_{\alpha,n}}; V_\lambda)$ for graded tensor product decompositions of so-called Kirillov and Reshetikhin [1987] (KR) modules $KR_{\alpha,n}$ ($\alpha \in [1,r]; n \in \mathbb{N}$) of a Lie algebra into irreducibles. The grading, inherited from the loop algebra Feigin and Loktev [1999] (fusion product) turns out to have many equivalent formulations: as energy of the corresponding crystal, as linearized energy in the Bethe Ansatz solutions of the corresponding inhomogeneous isotropic XXX quantum spin chain (the physical system at the origin of the problem, from which so-called fermionic formulas for graded multiplicities were derived Hatayama, Kuniba, Okado, Takagi, and Yamada [1999]). Recently, we have found yet another interpretation of this grading, as being provided by the canonical quantum deformation of the cluster algebra of the so-called Q-system for the algebra Di Francesco and Kedem [2014].

The latter is a recursive system for scalar variables $Q_{\alpha,n}$ $\alpha = 1, 2..., r$, $n \in \mathbb{Z}$. For the case of A_r it takes the form:

$$Q_{\alpha,n+1}\, Q_{\alpha,n-1} = (Q_{\alpha,n})^2 - Q_{\alpha+1,n}\, Q_{\alpha-1,n}$$

with boundary conditions $Q_{0,n} = Q_{r+1,n} = 1$ for all $n \in \mathbb{Z}$. It is satisfied by the KR characters $Q_{\alpha,n} = \chi_{KR_{\alpha,n}}(\mathbf{x})$. This is a discrete integrable system: there exist r algebraically independent polynomial quantities of the Q's that are conserved modulo the system, which we can view as describing evolution of the variables $Q_{\alpha,n}$ in discrete time n Di Francesco and Kedem [2010, 2018]. Taking advantage of this property, we were able to solve such systems by means of (strongly) non-intersecting lattice paths (the solution involves also a new continued fraction rearrangement theory Di Francesco and Kedem [2010] and Di Francesco [2011a]).

Such systems exist for all finite and affine algebras, and were shown to be particular sets of mutations in some cluster algebras Kedem [2008] and Di Francesco and Kedem [2009]. As such, they admit a natural quantization into a q-deformed, non-commutative Q-system, coined the quantum Q-system. For the case A_r it reads:

$$Q_{\alpha,n}\, Q_{\beta,n+1} = q^{\lambda_{\alpha,\beta}}\, Q_{\beta,n+1}\, Q_{\alpha,n} \tag{4-1}$$

$$q^{\lambda_{\alpha,\alpha}}\, Q_{\alpha,n+1}\, Q_{\alpha,n-1} = (Q_{\alpha,n})^2 - q\, Q_{\alpha+1,n}\, Q_{\alpha-1,n} \tag{4-2}$$

where $\lambda_{\alpha,\beta} = (C^{-1})_{\alpha,\beta}$, C the Cartan matrix of the algebra, and with the boundary conditions $Q_{0,n} = Q_{r+1,0} = 1$, $Q_{r+2,n} = 0$ for all $n \in \mathbb{Z}$. The non-commuting variables $Q_{\alpha,n}$ play the role of quantized KR characters. The path solutions of the classical Q-system admit a non-commutative version using non-commutative continued fractions Di Francesco and Kedem [2011].

For simplicity let us perform a change of variables. Define $A = Q_{r+1,1}$ and the degree operator Δ such that $\Delta\, Q_{\alpha,n} = q^{\alpha n} Q_{\alpha,n}\, \Delta$. Then the new variables $M_{\alpha,n} := q^{-\frac{1}{2}\lambda_{\alpha,\alpha}(n+r+1)}\, Q_{\alpha,n} \Delta^{\frac{\alpha}{r+1}}$ are subject to the new "M-system":

$$M_{\alpha,n}\, M_{\beta,n+1} = q^{\mathrm{Min}(\alpha,\beta)}\, M_{\beta,n+1}\, M_{\alpha,n} \tag{4-3}$$

$$q^{\alpha}\, M_{\alpha,n+1}\, M_{\alpha,n-1} = (M_{\alpha,n})^2 - M_{\alpha+1,n}\, M_{\alpha-1,n} \tag{4-4}$$

with boundary conditions $M_{0,n} = 1$ and $M_{r+1,n} = A^n \Delta$.

Theorem 4.2. *We have the following representation of the M-system via difference operators acting on the ring of symmetric functions of $N = r+1$ variables $(x_1, \ldots, x_N)$:*

$$M_{\alpha,n} = \sum_{\substack{I\subset[1,N]\\ |I|=\alpha}} x_I^n \prod_{\substack{i\in I\\ j\notin I}} \frac{x_i}{x_i - x_j} \Gamma_I \tag{4-5}$$

where $x_I = \prod_{i\in I} x_i$, $\Gamma_I = \prod_{i\in I} \Gamma_i$, and Γ_i is the multiplicative q-shift operator on the i-th variable: $(\Gamma_i\, f)(x_1, x_2, \ldots, x_N) = f(x_1, \ldots, x_{i-1}, q x_i, x_{i+1}, \ldots, x_N)$, and with moreover $A = x_1 x_2 \cdots x_N$, and $\Delta = \Gamma_1 \Gamma_2 \cdots \Gamma_N$.

Let $\chi_{\mathbf{n}}(q;\mathbf{x})$, $\mathbf{n} = \{n_{\alpha,n}\}_{\alpha\in[1,r];n\in[1,k]}$, $\mathbf{x} = (x_1, \ldots, x_N)$, denote the graded character: $\chi_{\mathbf{n}}(q;\mathbf{x}) = \sum_\lambda \mathrm{Mult}_q(\otimes KR_{\alpha,n}^{\otimes n_{\alpha,n}}; V_\lambda)\chi_\lambda(\mathbf{x})$, i.e. the generating function for graded multiplicities, where the irreducible characters $\chi_\lambda(\mathbf{x}) = s_\lambda(\mathbf{x})$ are the Schur functions.

Theorem 4.3 (Di Francesco and Kedem [2018])**.** *The graded character for the tensor product $\otimes KR_{\alpha,n}^{\otimes n_{\alpha,n}}$ reads*

$$\chi_{\mathbf{n}}(q^{-1};\mathbf{x}) = q^{-a(\mathbf{n})} \prod_{j=k}^{1} \prod_{\alpha=1}^{r} (M_{\alpha,j})^{n_{\alpha,j}} \cdot 1$$

with $a(\mathbf{n}) = \frac{1}{2}\sum_{i,j,\alpha,\beta} n_{\alpha,i}\mathrm{Min}(i,j)\mathrm{Min}(\alpha,\beta)n_{\beta,j} - \frac{1}{2}\sum_{i,\alpha} i\alpha\, n_{\alpha,i}$.

The results above were so far only derived for the A case, but they can be extended to B, C, D types Di Francesco and Kedem [n.d.].

4.2.2 From Cluster algebra to quantum toroidal and Elliptic Hall algebras. The form of the difference operator (4-5) is reminiscent of the celebrated Macdonald operators Macdonald [1995], for which the Macdonald polynomials form a complete family of eigenvectors. These were best understood in the context of Double Affine Hecke Algebra Cherednik [2005], in the functional representation. We actually found that a certain action of the natural $SL_2(\mathbb{Z})$ symmetry of DAHA produces the following *generalized Macdonald* difference operators in the functional representation:

$$\mathfrak{M}_{\alpha,n} = \sum_{\substack{I\subset[1,N]\\ |I|=\alpha}} x_I^n \prod_{\substack{i\in I\\ j\notin I}} \frac{tx_i - x_j}{x_i - x_j}\Gamma_I \tag{4-6}$$

We note the relation $M_{\alpha,n} = \lim_{t\to\infty} t^{-\alpha(N-\alpha)}\,\mathfrak{M}_{\alpha,n}$. We may therefore identify the quantum cluster algebra solution of the Q-system with the $t \to \infty$ (so-called dual q-Whittaker) limit of generalized Macdonald operators.

These operators allow to construct a representation of the so-called Ding–Iohara–Miki Ding and Iohara [1997] and Miki [2007] (DIM) or quantum toroidal $\mathfrak{g}l_1$ algebra as follows Di Francesco and Kedem [2017b]. Introduce the currents:

$$e(z) := \frac{q^{1/2}}{1-q}\sum_{n\in\mathbb{Z}} q^{n/2} z^n \mathfrak{M}_{1,n} \quad \text{and} \quad \mathfrak{f}(z) := e(z)\big|_{q\to q^{-1},t\to t^{-1}} \tag{4-7}$$

as well as the series

$$\psi^{\pm}(z) := \prod_{i=1}^{N} \frac{(1-q^{-\frac{1}{2}}t(zx_i)^{\pm1})(1-q^{\frac{1}{2}}t^{-1}(zx_i)^{\pm1})}{(1-q^{-\frac{1}{2}}(zx_i)^{\pm1})(1-q^{\frac{1}{2}}(zx_i)^{\pm1})} \in \mathbb{C}[[z^{\pm1}]] \tag{4-8}$$

Theorem 4.4 (Di Francesco and Kedem [ibid.])**.** *The currents and series* $(e,\mathfrak{f},\psi^{\pm})$ *form a level* $(0,0)$ *representation of the DIM algebra.*

In particular, we have the following exchange relation:

$$g(z,w)\,e(z)\,e(w) + g(w,z)\,e(w)\,e(z) = 0,\;\; g(z,w) = (z-qw)(z-t^{-1}w)(z-q^{-1}tw)$$

In the $t \to \infty$ limit this reduces to the $\mathfrak{U}_{\sqrt{q}}(\widehat{\mathfrak{sl}_2})$ upper Borel subalgebra relation, while the DIM relations go to some interesting degeneration, directly connected to the quantum

Q-system cluster algebra Di Francesco and Kedem [2018, 2017b]. Algebra relations allow in particular to derive a quantum determinant formula for $M_{\alpha,n}$ as a polynomial of the $M_n := M_{1,n}$'s. Let us introduce the currents $m_\alpha(z) := \sum_{n\in\mathbb{Z}} z^n M_{\alpha,n}$, and in particular $m(z) := m_1(z) = \lim_{t\to\infty} t^{1-N} \frac{1-q}{q^{1/2}} e(z)$, then:

Theorem 4.5 (Di Francesco and Kedem [2017a]).

$$m_\alpha(z) = \left\{ \left(\prod_{1\le i<j\le\alpha} \left(1 - q\frac{u_j}{u_i}\right) \right) m(u_1)m(u_2)\cdots m(u_\alpha) \right\} \Bigg|_{(u_1u_2\cdots u_\alpha)^n} \tag{4-9}$$

where the subscript stands for the coefficient of $(u_1u_2\cdots u_\alpha)^n$.

Note also that the function of $\mathbf{u}$ in (4-9) is a multi-current generating series. Let us define $M_{a_1,\dots,a_\alpha}$ to be the coefficient of $u_1^{a_1}\cdots u_\alpha^{a_\alpha}$ in this series. There is a very nice expression for $M_{a_1,\dots,a_\alpha}$ as a quantum determinant, involving a sum over Alternating Sign Matrices. This is because the quantity $\prod_{i<j} v_i + \lambda v_j$ is the λ-determinant $\lambda\det(V_n)$ (as defined by Robbins and Rumsey [1986]) of the Vandermonde matrix $V_n := (v_i^{n-j})_{1\le i,j\le n}$. We denote by ASM_n the set of ASM of size n. The inversion number of an ASM is the quantity $I(A) = \sum_{i>k,j<\ell} A_{i,j}A_{k,\ell}$. We also denote by $N(A)$ the number of -1's in A. Let us also define the column vector $v = (n-1, n-2, .., 1, 0)^t$, and for each ASM A we denote by $m_i(A) := (Av)_i$. Then we have the explicit formula, obtained by taking $\lambda = -q$ for the λ-determinant of the $\alpha\times\alpha$ Vandermonde matrix V_α:

$$\prod_{1\le i<j\le\alpha} (v_i - qv_j) = \sum_{A\in ASM_n} (-q)^{I(A)-N(A)}(1-q)^{N(A)} \prod_{i=1}^{n} v_i^{m_i(A)}$$

Combining this with (4-9), we deduce the following compact expression for the quantum determinant:

Theorem 4.6. *The quantum determinant of the matrix* $\left(M_{a_j+i-j}\right)_{1\le i,j\le\alpha}$ *reads:*

$$M_{a_1,\dots,a_\alpha} = \left|\left(M_{a_j+i-j}\right)_{1\le i,j\le\alpha}\right|_q \tag{4-10}$$

$$= \sum_{A\in ASM_\alpha} (-q)^{I(A)-N(A)}(1-q)^{N(A)} \prod_{i=1}^{\alpha} M_{a_i+\alpha-i-m_i(A)} \tag{4-11}$$

Finally, we use a known isomorphism Schiffmann and Vasserot [2011] between the Spherical DAHA with the Elliptic Hall algebra (EHA) to make the connection between generalized Macdonald operators and a functional representation of the EHA Di Francesco and Kedem [2017a]. From this connection, we obtain new algebraic relations between the operators $\mathfrak{M}_{\alpha,n}$, inherited from the "skinny triangle" relations of Burban and Schiffmann [2012].

To conclude, the results of this section are so far valid for the A type only. It would be interesting to investigate the (t-deformed) algebraic structures hiding behind the B,C,D cases as well.

5 Open problems

In this note, we described various instances of discrete or continuous integrability in combinatorial problems. A recurrent theme is the ability to rephrase said combinatorial problems in terms of paths or trees.

Paths are very important objects. Equipped with non-commutative weights, paths allow to describe non-commutative monomials in finitely generated non-commutative algebras. We have encountered a few instances of this in the present note. A crucial question remains open: how to deal with *families* of non-intersecting non-commutative paths? We have found specific answers in the cases where the non-commutativity is "under control", e.g. in the case of quantum path weights with specific q-commuting relations Di Francesco [2011b]. More general non-commuting weighted paths can be described via the theory of quasideterminants I. Gelfand, S. Gelfand, Retakh, and Wilson [2005], however it remains to find a good theory for non-intersecting non-commutative paths, and perhaps a non-commutative version of the Lindström-Gessel-Viennot (LGV) determinant formula.

Interacting paths are a combinatorial version of interacting fermions. Starting from NILP, we can turn on some interaction, by for instance allowing paths to touch without crossing (osculating paths) and attaching a contact energy to such instances. As shown in the case of the 6V/ASM model, such interactions still allow for solving, together with the machinery of integrable lattice models. As another indication, he so-called tangent method for determining phase transitions in large random tilings (arctic curves) seems to apply to interacting paths as well. The determinant form of the partition function for the 6V model with DWBC would point to the fact that there should exist determinant formulas for interacting paths that generalize LGV. It seems that a number of interacting path problems are still open, and a systematic study is in order.

References

Roger E. Behrend, Philippe Di Francesco, and P. Zinn-Justin (2012). "On the weighted enumeration of alternating sign matrices and descending plane partitions". *J. Combin. Theory Ser. A* 119.2, pp. 331–363. MR: 2860598 (cit. on p. 2607).

Roger E. Behrend, Philippe Di Francesco, and Paul Zinn-Justin (2013). "A doubly-refined enumeration of alternating sign matrices and descending plane partitions". *J. Combin. Theory Ser. A* 120.2, pp. 409–432. arXiv: 1202.1520. MR: 2995049 (cit. on p. 2607).

J. Bouttier, Philippe Di Francesco, and E. Guitter (2003). "Geodesic distance in planar graphs". *Nuclear Phys. B* 663.3, pp. 535–567. MR: 1987861 (cit. on pp. 2604, 2605).

Igor Burban and Olivier Schiffmann (2012). "On the Hall algebra of an elliptic curve, I". *Duke Math. J.* 161.7, pp. 1171–1231. MR: 2922373 (cit. on p. 2613).

Luigi Cantini and Andrea Sportiello (2011). "Proof of the Razumov-Stroganov conjecture". *J. Combin. Theory Ser. A* 118.5, pp. 1549–1574. MR: 2771600 (cit. on p. 2607).

Ivan Cherednik (2005). *Double affine Hecke algebras*. Vol. 319. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, pp. xii+434. MR: 2133033 (cit. on p. 2611).

F. Colomo and A. Sportiello (2016). "Arctic curves of the six-vertex model on generic domains: the tangent method". *J. Stat. Phys.* 164.6, pp. 1488–1523. MR: 3541190 (cit. on p. 2608).

Philippe Di Francesco (2005). "Geodesic distance in planar graphs: an integrable approach". *Ramanujan J.* 10.2, pp. 153–186. MR: 2194521 (cit. on p. 2605).

– (2011a). "Discrete integrable systems, positivity, and continued fraction rearrangements". *Lett. Math. Phys.* 96.1-3, pp. 299–324. MR: 2788915 (cit. on p. 2609).

– (2011b). "Quantum A_r Q-system solutions as q-multinomial series". *Electron. J. Combin.* 18.1, Paper 176, 17. MR: 2836811 (cit. on p. 2613).

Philippe Di Francesco, P. Ginsparg, and J. Zinn-Justin (1995). "2D gravity and random matrices". *Phys. Rep.* 254.1-2, p. 133. MR: 1320471 (cit. on p. 2603).

Philippe Di Francesco, E. Guitter, and C. Kristjansen (2000). "Integrable 2D Lorentzian gravity and random walks". *Nuclear Phys. B* 567.3, pp. 515–553. MR: 1741653 (cit. on pp. 2600, 2602).

Philippe Di Francesco and Rinat Kedem (n.d.). work in progress (cit. on p. 2611).

– (2009). "Q-systems as cluster algebras. II. Cartan matrix of finite type and the polynomial property". *Lett. Math. Phys.* 89.3, pp. 183–216. MR: 2551179 (cit. on p. 2610).

– (2010). "Q-systems, heaps, paths and cluster positivity". *Comm. Math. Phys.* 293.3, pp. 727–802. MR: 2566162 (cit. on p. 2609).

– (2011). "Non-commutative integrability, paths and quasi-determinants". *Adv. Math.* 228.1, pp. 97–152. MR: 2822229 (cit. on p. 2610).

– (2014). "Quantum cluster algebras and fusion products". *Int. Math. Res. Not. IMRN* 10, pp. 2593–2642. MR: 3214279 (cit. on p. 2609).

– (Apr. 2017a). "(t,q) Q-systems, DAHA and quantum toroidal algebras via generalized Macdonald operators". arXiv: 1704.00154 (cit. on p. 2612).

– (2017b). "Quantum Q systems: from cluster algebras to quantum current algebras". *Lett. Math. Phys.* 107.2, pp. 301–341. arXiv: 1606.09052. MR: 3600918 (cit. on pp. 2611, 2612).

– (2018). "Difference equations for graded characters from quantum cluster algebra". *Transform. Groups*, pp. 1–34. arXiv: 1505.01657 (cit. on pp. 2609, 2610, 2612).

Philippe Di Francesco, Rinat Kedem, and Bolor Turmunkh (2017). "A path model for Whittaker vectors". *J. Phys. A* 50.25, pp. 255201, 35. MR: 3659140 (cit. on pp. 2608, 2609).

Philippe Di Francesco and D. Kutasov (1991). "Correlation functions in 2D string theory". *Phys. Lett. B* 261.4, pp. 385–390. MR: 1110973 (cit. on p. 2603).

Philippe Di Francesco and Matthew F. Lapa (2018). "Arctic Curves in path models from the Tangent Method". *J. Phys. A* 51, pp. 155202, 63. arXiv: 1711.03182 (cit. on p. 2608).

Philippe Di Francesco and P. Zinn-Justin (2005). "Around the Razumov-Stroganov conjecture: proof of a multi-parameter sum rule". *Electron. J. Combin.* 12, Research Paper 6, 27. MR: 2134169 (cit. on p. 2607).

– (2006). "Inhomogeneous model of crossing loops and multidegrees of some algebraic varieties". *Comm. Math. Phys.* 262.2, pp. 459–487. arXiv: 0412031. MR: 2200268 (cit. on p. 2607).

Jintai Ding and Kenji Iohara (1997). "Generalization of Drinfeld quantum affine algebras". *Lett. Math. Phys.* 41.2, pp. 181–193. MR: 1463869 (cit. on p. 2611).

B. Feigin and S. Loktev (1999). "On generalized Kostka polynomials and the quantum Verlinde rule". In: *Differential topology, infinite-dimensional Lie algebras, and applications*. Vol. 194. Amer. Math. Soc. Transl. Ser. 2. Amer. Math. Soc., Providence, RI, pp. 61–79. MR: 1729359 (cit. on p. 2609).

Israel Gelfand, Sergei Gelfand, Vladimir Retakh, and Robert Lee Wilson (2005). "Quasideterminants". *Adv. Math.* 193.1, pp. 56–141. MR: 2132761 (cit. on p. 2613).

G. Hatayama, A. Kuniba, M. Okado, T. Takagi, and Y. Yamada (1999). "Remarks on fermionic formula". In: *Recent developments in quantum affine algebras and related topics (Raleigh, NC, 1998)*. Vol. 248. Contemp. Math. Amer. Math. Soc., Providence, RI, pp. 243–291. MR: 1745263 (cit. on p. 2609).

Rinat Kedem (2008). "Q-systems as cluster algebras". *J. Phys. A* 41.19, pp. 194011, 14. MR: 2452184 (cit. on p. 2610).

A. N. Kirillov and N. Yu. Reshetikhin (1987). "Representations of Yangians and multiplicities of the inclusion of the irreducible components of the tensor product of representations of simple Lie algebras". *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 160.Anal. Teor. Chisel i Teor. Funktsiĭ. 8, pp. 211–221, 301. MR: 906858 (cit. on p. 2609).

M. L. Kontsevich (1991). "Intersection theory on the moduli space of curves". *Funktsional. Anal. i Prilozhen.* 25.2, pp. 50–57, 96. MR: 1142208 (cit. on p. 2603).

Bertram Kostant (1978). "On Whittaker vectors and representation theory". *Invent. Math.* 48.2, pp. 101–184. MR: 507800 (cit. on p. 2608).

Greg Kuperberg (1996). "Another proof of the alternating-sign matrix conjecture". *Internat. Math. Res. Notices* 3, pp. 139–150. MR: 1383754 (cit. on p. 2607).

I. G. Macdonald (1995). *Symmetric functions and Hall polynomials*. Second. Oxford Mathematical Monographs. With contributions by A. Zelevinsky, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, pp. x+475. MR: 1354144 (cit. on p. 2611).

Kei Miki (2007). "A (q, γ) analog of the $W_{1+\infty}$ algebra". *J. Math. Phys.* 48.12, pp. 123520, 35. MR: 2377852 (cit. on p. 2611).

W. H. Mills, David P. Robbins, and Howard Rumsey Jr. (1983). "Alternating sign matrices and descending plane partitions". *J. Combin. Theory Ser. A* 34.3, pp. 340–359. MR: 700040 (cit. on p. 2607).

A. V. Razumov and Yu. G. Stroganov (2004). "Combinatorial nature of the ground-state vector of the $\mathrm{O}(1)$ loop model". *Teoret. Mat. Fiz.* 138.3, pp. 395–400. MR: 2077318 (cit. on p. 2607).

David P. Robbins and Howard Rumsey Jr. (1986). "Determinants and alternating sign matrices". *Adv. in Math.* 62.2, pp. 169–184. MR: 865837 (cit. on p. 2612).

Gilles Schaeffer (1997). "Bijective census and random generation of Eulerian planar maps with prescribed vertex degrees". *Electron. J. Combin.* 4.1, Research Paper 20, 14. MR: 1465581 (cit. on p. 2603).

O. Schiffmann and E. Vasserot (2011). "The elliptic Hall algebra, Cherednik Hecke algebras and Macdonald polynomials". *Compos. Math.* 147.1, pp. 188–234. MR: 2771130 (cit. on p. 2612).

Doron Zeilberger (1996). "Proof of the alternating sign matrix conjecture". *Electron. J. Combin.* 3.2. The Foata Festschrift, Research Paper 13, approx. 84. MR: 1392498 (cit. on p. 2607).

P. Zinn-Justin and Philippe Di Francesco (2008). "The quantum Knizhnik–Zamolodchikov equation, completely symmetric self-complementary plane partitions, and alternating-sign matrices". *Teoret. Mat. Fiz.* 154.3, pp. 387–408. arXiv: 0703015. MR: 2431554 (cit. on p. 2607).

Received 2017-11-21.

PHILIPPE DI FRANCESCO

Department of Mathematics
University of Illinois at Urbana-Champaign
1409 West Green Street
Urbana, IL 61801
USA

and

Institut de Physique Théorique du Commissariat à l'Energie Atomique
Unité de Recherche associée du CNRS
Université Paris-Saclay
CEA Saclay/IPhT/Bat 774
F-91191 Gif sur Yvette Cedex
France
philippe@illinois.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2615–2634)

CONFORMAL FIELD THEORY, VERTEX OPERATOR ALGEBRAS AND OPERATOR ALGEBRAS

Yasuyuki Kawahigashi (河東泰之)

Abstract

We present recent progress in theory of local conformal nets which is an operator algebraic approach to study chiral conformal field theory. We emphasize representation theoretic aspects and relations to theory of vertex operator algebras which gives a different and algebraic formulation of chiral conformal field theory.

1 Introduction

Quantum field theory is a vast area in physics and two-dimensional conformal field theory has caught much attention recently. A two-dimensional conformal field theory decomposes into two chiral conformal field theories, and here we present mathematical studies of a chiral conformal field theory based on operator algebras. It is within a scope of what is called algebraic quantum field theory and our mathematical object is called a local conformal net.

The key idea in algebraic quantum field theory is to work on operator algebras generated by observables in a spacetime region rather than quantum fields. In chiral conformal field theory, the spacetime becomes a one-dimensional circle and a spacetime region is an interval in it, which is a nonempty, nondense, open and connected set in the circle, so we deal with a continuous family of operator algebras parameterized by intervals. This is what a local conformal net is.

Each operator algebra of a local conformal net acts on the same Hilbert space from the beginning, but we also consider its representation theory on another Hilbert space. Such a representation corresponds to a notion of a charge, and a unitary equivalence class of a representation is called a superselection sector. In the Doplicher-Haag-Roberts theory Doplicher, Haag, and Roberts [1971], a representation is realized with a DHR endomorphism of one operator algebra, and such an endomorphism produces a subfactor in the sense of the Jones theory Jones [1983], Jones [1985]. Subfactor theory plays an important

---

MSC2010: primary 81T40; secondary 17B69, 18D10, 46L37, 81T05.

role in this approach. It has revolutionized theory of operator algebras and revealed its surprising deep relations to 3-dimensional topology, quantum groups and solvable lattice models. Its connection to quantum field theory was clarified by Longo [1989] and it has been an important tool also in conformal field theory since then.

Representation theory of a local conformal net gives a powerful tool to study chiral conformal field theory. We present α-induction, a certain induction procedure for representation theory of a local conformal net, and its use for classification theory.

A vertex operator algebra gives another axiomatization of a chiral conformal field theory and it has started with the famous Moonshine conjecture Conway and Norton [1979]. The axiomatic framework has been established in Frenkel, Lepowsky, and Meurman [1988] and we have had many research papers on this topic. This is an algebraic axiomatization of Fourier expansions of a family of operator-valued distributions on the one-dimensional circle. Since a local conformal net and a vertex operator algebra give different axiomatizations of the same physical theory, it is natural to expect that they have many common features. There have been many parallel results in the two theories, but a precise relation between the two were not known until recently. We have established that if a vertex operator algebra satisfies unitarity and an extra mild assumption called strong locality, then we can construct the corresponding local conformal net and also recover the original vertex operator algebra from the local conformal net. Strong locality is known to be satisfied for most examples and we do not know any example of a vertex operator algebra which does not have strong locality.

There are many open problems to study in the operator algebraic approach to chiral conformal field theory. We present some of them in this article.

We refer a reader to lecture notes Kawahigashi [2015b] for more details with an extensive bibliography.

This work was supported in part by Research Grants and the Grants-in-Aid for Scientific Research, JSPS.

2 Algebraic quantum field theory and local conformal nets

In a common approach to quantum field theory such as the Wightman axioms, we deal with quantum fields which are a certain kind of operator-valued distributions on the spacetime acting on the same Hilbert space together with a spacetime symmetry group. An operator-valued distribution T applied to a test function f gives $\langle T, f\rangle$ which is an (often unbounded) operator. Handling distributions, rather than functions, and unbounded operators causes technical difficulties, so an idea of algebraic quantum field theory of Haag-Kastler is to study operator algebras generated by observables in a spacetime region. Let T be an operator-valued distribution and f be a test function supported in O which is a spacetime region. Then $\langle T, f\rangle$ gives an observable in O (if it is self-adjoint). Let $A(O)$ be the von Neumann algebra generated by these observables. (A von Neumann algebra is an algebra of bounded linear operators on a Hilbert space containing the identity operator which is closed under the $*$-operation and the strong operator topology. This topology is given by pointwise convergence on the Hilbert space.) We have a family $\{A(O)\}$ of von Neumann algebras. We impose physically natural axioms on such a family and make a mathematical study of these axioms.

We apply the above general idea to 2-dimensional conformal field theory. We first consider the 2-dimensional Minkowski space with space coordinate x and time coordinate t. We have a certain restriction procedure of a conformal field theory on the Minkowski space to the two light rays $\{x = \pm t\}$. In this way, we can regard one light ray as a kind of spacetime though it has only one dimension. Then conformal symmetry can move the point at infinity of this light ray, so our space should be now S^1, the one-point compactification of a light ray. A spacetime region is now a nonempty nondense open connected subset of S^1 and such a set is called an interval. Our mathematical object is a family of von Neumann algebras $\{A(I)\}$ parametrized by an interval $I \subset S^1$. We impose the following axioms on this family.

1. (Isotony) For two intervals $I_1 \subset I_2$, we have $A(I_1) \subset A(I_2)$.

2. (Locality) When two intervals I_1, I_2 are disjoint, we have $[A(I_1), A(I_2)] = 0$.

3. (Möbius covariance) We have a unitary representation U of $PSL(2, \mathbb{R})$ on the Hilbert space such that we have $U(g)A(I)U(g)^* = A(gI)$ for all $g \in PSL(2, \mathbb{R})$, where g acts on S^1 as a fractional linear transformation on $\mathbb{R} \cup \{\infty\}$ and $S^1 \setminus \{-1\}$ is identified with $\mathbb{R}$ through the Cayley transform $C(z) = -i(z-1)/(z+1)$.

4. (Conformal covariance) We have a projective unitary representation, still denoted by U, of Diff(S^1), the group of orientation preserving diffeomorphisms of S^1, extending the unitary representation U of $PSL(2, \mathbb{R})$ such that

$$U(g)\mathfrak{A}(I)U(g)^* = A(gI), \quad g \in \mathrm{Diff}(S^1),$$
$$U(g)xU(g)^* = x, \quad x \in A(I),\ g \in \mathrm{Diff}(I'),$$

   where I' is the interior of the complement of I and Diff(I') is the set of diffeomorphisms of S^1 which are the identity map on I.

5. (Positive energy condition) The generator of the restriction of U to the rotation subgroup of S^1, the conformal Hamiltonian, is positive.

6. (Existence of the vacuum vector) We have a unit vector Ω, called the vacuum vector, such that Ω is fixed by the representation U of $PSL(2, \mathbb{R})$ and $(\bigvee_{I \subset S^1} A(I))\Omega$ is

dense in the Hilbert space, where $\bigvee_{I\subset S^1} A(I)$ is the von Neumann algebra generated by $A(I)$'s.

7. (Irreducibility) The von Neumann algebra $\bigvee_{I\subset S^1} A(I)$ is the algebra of all the bounded linear operators on the Hilbert space.

Isotony is natural because a larger spacetime domain should have more observables. Locality comes from the Einstein causality in the 2-dimensional Minkowski space that observables in spacelike separated regions should commute with each other. Note that we have a simple condition of disjointness instead of spacelike separation. Conformal covariance represents an infinite dimensional symmetry. This gives a reason for the name "conformal" field theory. The positive energy condition expresses positivity of the eigenvalues of the conformal Hamiltonian. The vacuum state is a physically distinguished state of the Hilbert space. Irreducibility means that our Hilbert space is irreducible.

It is non-trivial to construct an example of a local conformal net. Basic sources of constructions are as follows. These are also sources of constructions of vertex operator algebras as we see below.

1. Affine Kac-Moody algebras Gabbiani and Fröhlich [1993], Wassermann [1998], Toledano Laredo [1999]

2. Virasoro algebra Xu [2000a], Kawahigashi and Longo [2004a]

3. Even lattices Kawahigashi and Longo [2006], Dong and Xu [2006]

When we have some examples of local conformal nets, we have the following methods to construct new ones.

1. Tensor product

2. Simple current extension Böckenhauer and Evans [1998]

3. Orbifold construction Xu [2000b]

4. Coset construction Xu [2000a]

5. Extension by a Q-system Longo and Rehren [1995], Kawahigashi and Longo [2004a], Xu [2007]

The first four constructions were first studied for vertex operator algebras. The last one was first studied for a local conformal net and later for vertex operator algebras Huang, Kirillov, and Lepowsky [2015]. The Moonshine net, the operator algebraic counterpart of the famous Moonshine vertex operator algebra, is constructed from the Leech lattice, an even lattice of rank 24, with a combination of the orbifold construction and a simple current extension Kawahigashi and Longo [2006], for example. This is actually given by a 2-step simple current extension as in Kawahigashi and Suthichitranont [2014]. The Q-system in the last construction was first introduced in Longo [1994]. It is also known under the name of a Frobenius algebra in algebraic literature.

Irreducibility implies that each $A(I)$ has a trivial center. Such an algebra is called a factor. It turns out that each algebra $A(I)$ is isomorphic to the Araki-Woods factor of type III_1 because the split property automatically holds by Morinelli, Tanimoto, and Weiner [2018] and it implies hyperfiniteness of $A(I)$. This shows that each single algebra $A(I)$ contains no information about a local conformal net and what is important is relative relation among $A(I)$'s.

3 Representation theory and superselection sectors

We now present representation theory of a local conformal net. Each algebra $A(I)$ of a local conformal net $\{A(I)\}$ acts on the same Hilbert space having the vacuum vector from the beginning, but we also consider a representation of an algebra $A(I)$ on another common Hilbert space (without a vacuum vector).

The Haag duality $A(I') = A(I)'$ automatically holds from the axioms, where the prime on the right hand side denotes the commutant, and this implies that each representation is represented with an endomorphism λ of $A(I)$ for some fixed I. This is a standard Doplicher-Haag-Roberts theory adapted to a local conformal net Fredenhagen, Rehren, and Schroer [1989]. An endomorphism λ produces $\lambda(A(I))$ which is subalgebra of $A(I)$ and a factor, so it is called a subfactor. It is an object in the Jones theory of subfactors Jones [1983]. The relative size of the subfactor $\lambda(A(I))$ with respect to $A(I)$ is called the Jones index $[A(I) : \lambda(A(I))]$. It turns out that the square root $[A(I) : \lambda(A(I))]^{1/2}$ of the Jones index gives a proper notion of the dimension of the corresponding representation of $\{A(I)\}$ Longo [1989]. The dimension $\dim(\lambda)$ takes its value in the interval $[1, \infty]$.

It is important to have a notion of a tensor product of representations of a local conformal net. Note that while it is easy to define a tensor product of representations of a group, we have no notion of a tensor product of representations of an algebra. It turns out that a composition of endomorphisms of $A(I)$ for a fixed I gives a right notion of a tensor product of representations Doplicher, Haag, and Roberts [1971]. In this way, we have a tensor category of finite dimensional representations of $\{A(I)\}$. The original action of $A(I)$ on the Hilbert space is called the vacuum representation and has dimension 1. It plays a role of a trivial representation. In the original setting of the Doplicher-Haag-Roberts theory on the higher dimensional Minkowski space, the tensor product operation is commutative in a natural sense and we have a symmetric tensor category. Now in the setting of chiral conformal field theory, the commutativity is more subtle, and we have a structure of braiding Fredenhagen, Rehren, and Schroer [1989]. We thus have a braided tensor category of finite dimensional representations.

We are often interested in a situation where we have only finitely many irreducible representations and such finiteness is usually called rational. (This rationality is well-studied in a context of representation theory of quantum groups at roots of unity in connection to quantum invariants in 3-dimensional topology.) We have defined complete rationality for a local conformal net, which means we have only finitely many irreducible representations up to unitary equivalence and all of them have finite dimensions, and given its operator algebraic characterization in terms of finiteness of the Jones index of a certain subfactor in Kawahigashi, Longo, and Müger [2001]. (We originally assumed two more properties for complete rationality, but they have been shown to be automatic by Longo and Xu [2004], Morinelli, Tanimoto, and Weiner [2018], respectively.) This characterization is given by only studying the vacuum representation. We have further proved that complete rationality implies that the braiding of the representations is non-degenerate, that is, we have the following theorem in Kawahigashi, Longo, and Müger [2001].

Theorem 3.1. *The tensor category of finite dimensional representations of a completely rational local conformal net is modular.*

It is an important open problem to decide which modular tensor category arises as the representation category of a completely rational local conformal net. The history of classification theory of factors, group actions and subfactors in theory of von Neumann algebras due to Connes, Haagerup, Jones, Ocneanu and Popa culminating in Popa [1994] tells us that as long as we have an analytic condition, generally called amenability, we have no nontrivial obstruction to realization of algebraic invariants. This strongly suggests that any modular tensor category is realized as the representation category of some local conformal net, because we now have amenability automatically. This conjecture has caught much attention these days because of recent work of Jones. We turn to this problem again in the next section.

4 Subfactors and tensor categories

In the Jones theory of subfactors, we study an inclusion $N \subset M$ of factors. In the original setting of Jones [1983], one considers type II_1 factors, but one has to deal with type III factors in conformal field theory. The Jones theory has been extended to type III factors by Pimsner-Popa and Kosaki, and many algebraic arguments are now more or less parallel in the type II_1 and type III cases. For simplicity, we assume factors are of type II_1 in this section. We refer reader to Evans and Kawahigashi [1998] for details of subfactor theory.

We start with a subfactor $N \subset M$. The Jones index $[M : N]$ is a number in the interval $[1, \infty]$. In this section, we assume that the index is finite. On the algebra M, we have the left and right actions of M itself. We restrict the left action to the subalgebra N, and we have a bimodule ${}_N M_M$. We make the completion of M with respect to the inner product arising from the trace functional and obtain the Hilbert space $L^2(M)$. For simplicity, we still write ${}_N M_M$ for this Hilbert space with the left action of N and the right action of M. We make relative tensor powers such as ${}_N M \otimes_M M \otimes_N M \otimes_M \cdots$ and their irreducible decomposition gives four kinds of bimodules, N-N, N-M, M-N and M-M. If we have only finitely many irreducible bimodules in this way, we say that the subfactor $N \subset M$ is of finite depth. In this case, finite direct sums of these irreducible N-N (and M-M) bimodules (up to isomorphism) give a fusion category. Note that the relative tensor product is not commutative in general and we have no braiding structure.

If we have a free action of a finite group G on a factor M, we have a subfactor $N = M^G \subset M$. The index is the order of G and the fusion category of N-N bimodules is the representation category of G. There are other constructions of subfactors from actions of finite groups and their quantum group versions give many interesting examples of subfactors. If the index is less than 4, the set of all the possible values is $\{4\cos^2 \pi/n \mid n = 3, 4, 5, \ldots\}$ Jones [1983]. Classification of subfactors with index less than 4 has been given in Ocneanu [1988] and this is well-understood today in terms of quantum groups or conformal field theory. Such classification of subfactors has been extended to index value 5 Jones, Morrison, and Snyder [2014] recently.

There are some exceptional subfactors which do not seem to arise from such constructions involving (quantum) groups. The most notable examples are the Haagerup subfactor Asaeda and Haagerup [1999], the Asaeda-Haagerup subfactor Asaeda and Haagerup [ibid.] and the extended Haagerup subfactor Bigelow, Peters, Morrison, and Snyder [2012] in the index range $(4, 5)$. (The first two were constructed along an extension of the line of Ocneanu [1988] and the last one is based on the planar algebra of Jones.) Such a subfactor produces an exceptional fusion category and then it produces an exceptional modular tensor category through the Drinfeld center construction. (See Izumi [2000] for an operator algebraic treatment of this.) Such a modular tensor category does not seem to arise from a combination of other known constructions applied to the Wess-Zumino-Witten models. The above three subfactors were found through a combinatorial search for a very narrow range of index values. This strongly suggests that there is a huge variety of exceptional fusion categories and modular tensor categories beyond what is known today. History of classification theory of subfactors even strongly suggests that there is a huge variety of exceptional modular tensor categories even up to Witt equivalence ignoring Drinfeld centers, because it seems impossible to exhaust all examples by prescribing construction methods.

As explained in the previous section, we strongly believe that all such exceptional modular tensor categories do arise from local conformal nets. This would mean that there is a huge variety of chiral conformal field theories beyond what is known today. For the Haagerup subfactor, a partial evidence for this conjecture is given in Evans and Gannon [2011].

5 α-induction, modular invariants and classification theory

We next present an important tool to study representation of a local conformal net. For a subgroup H of another group G and a representation of H, we have a notion of an induced representation of G. We have some similar notion for a representation of a local conformal net. Let $\{A(I) \subset B(I)\}$ be an inclusion of local conformal nets and assume the index $[B(I) : A(I)]$ is finite. For a a representation of $\{A(I)\}$ which is given by an endomorphism λ of a factor $A(I)$ for some fixed interval I, we extend λ to an endomorphism of $B(I)$. This extension depends on a choice of positive and negative crossings in the braiding structure of representations of $\{A(I)\}$ and we denote it with $\alpha^{\pm}_{\lambda}$ where $\pm$ stands for the choice of positive and negative crossings. This gives an "almost" representation of $\{B(I)\}$ and it is called a soliton endomorphism. This induction machinery is called α-induction. It was first introduced in Longo and Rehren [1995] and studied in detail in Xu [1998], Böckenhauer and Evans [1998]. Ocneanu had a graphical calculus in a very different context involving the A-D-E Dynkin diagrams and the two methods were unified in Böckenhauer, Evans, and Kawahigashi [1999], Böckenhauer, Evans, and Kawahigashi [2000]. It turns out that the intersection of irreducible endomorphisms of $B(I)$ arising from α^+-induction and α^--induction exactly gives those corresponding the representations of $\{B(I)\}$ by Kawahigashi, Longo, and Müger [2001], Böckenhauer, Evans, and Kawahigashi [1999], Böckenhauer, Evans, and Kawahigashi [2000].

Let $\{A(I)\}$ be completely rational in the above setting. Then $\{B(I)\}$ is automatically also completely rational. (The converse also holds.) The modular tensor category of $\{A(I)\}$ gives a (finite dimensional) unitary representation of $SL(2,\mathbb{Z})$ from its braiding. (The dimension of the representation is the number of irreducible representations of $\{A(I)\}$ up to unitary equivalence.) Define the matrix $(Z_{\lambda\mu})$ by $Z_{\lambda\mu} = \dim \mathrm{Hom}(\alpha^+_{\lambda}, \alpha^-_{\mu})$ where λ, μ denote endomorphisms of $A(I)$ corresponding to irreducible representations of $\{A(I)\}$. Then we have the following in Böckenhauer, Evans, and Kawahigashi [1999].

Theorem 5.1. *The matrix Z commutes with the above unitary representation of $SL(2,\mathbb{Z})$.*

Such Z also satisfies $Z_{\lambda\mu} \in \{0, 1, 2, \dots\}$ and $Z_{00} = 1$ where 0 denotes the vacuum representation of $\{A(I)\}$. Such a matrix is called a modular invariant of the representation of $SL(2,\mathbb{Z})$. The number of modular invariants for a given local conformal net $\{A(I)\}$ is always finite and often quite limited. This gives the following classification method of all possible irreducible extensions $\{B(I)\}$ for a given local conformal net $\{A(I)\}$. (Any irreducible extension automatically has a finite index by Izumi, Longo, and Popa [1998].)

1. Find all possible modular invariants $(Z_{\lambda\mu})$ for the modular tensor category arising from representations of $\{A(I)\}$.
2. For each $(Z_{\lambda\mu})$, determine all possible Q-systems corresponding to $\bigoplus Z_{0\lambda}\lambda$.
3. Pick up only local Q-systems.

Consider a local conformal net $\{A(I)\}$. The projective unitary representation of $\mathrm{Diff}(S^1)$ gives a representation of the Virasoro algebra and it gives a positive real number c called the central charge. This is a numerical invariant of a local conformal net and the value of c is in the set $\{1-6/n(n+1) \mid n=3,4,5,\ldots\} \cup [1,\infty)$. We now restrict ourselves to the case $c<1$. Let $\mathrm{Vir}_c(I)$ be the von Neumann algebra generated by $U(g)$ where $g \in \mathrm{Diff}(S^1)$ acts trivially on I'. This gives an extension $\{\mathrm{Vir}_c(I) \subset A(I)\}$. It turns out $\{\mathrm{Vir}_c(I)\}$ is completely rational and we can apply the above method to classify all possible $\{A(I)\}$. The modular invariants have been classified in Cappelli, Itzykson, and Zuber [1987], and locality and a certain 2-cohomology argument imply that the extensions exactly correspond to so-called type I modular invariants. We thus have a complete classification of local conformal nets with $c<1$ as follows Kawahigashi and Longo [2004a].

Theorem 5.2. *Any local conformal net with $c<1$ is one of the following.*

1. *The Virasoro nets $\{\mathrm{Vir}_c(I)\}$ with $c<1$.*
2. *Their simple current extensions with index 2.*
3. *Four exceptionals at $c=21/22, 25/26, 144/145, 154/155$.*

The four exceptionals correspond to the Dynkin diagrams E_6 and E_8. Three of them are identified with certain coset constructions, but the remaining one with $c=144/145$ does not seem to be related to any other known constructions so far. All these four are given by an extension by a Q-system. Note that this appearance of modular invariants is different from its original context in 2-dimensional conformal field theory.

6 Vertex operator algebras

A vertex operator algebra gives another mathematical axiomatization of a chiral conformal field theory. It deals with Fourier expansions of operator-valued distributions, vertex operators, on S^1 in an algebraic manner.

Recall that we have a complete list of finite simple groups today as follows Frenkel, Lepowsky, and Meurman [1988].

1. Cyclic groups of prime order.
2. Alternating groups of degree 5 or higher.
3. 16 series of groups of Lie type over finite fields.
4. 26 sporadic finite simple groups.

The largest group among the 26 groups in the fourth in terms of the order is called the Monster group, and its order is approximately 8×10^{53}. This group was first constructed by Griess. It has been known that the smallest dimension of a non-trivial irreducible representation of the Monster group is 196883.

The next topic in this section is the j-function. This is a function of a complex number τ with $\operatorname{Im} \tau > 0$ with the following expansion.

$$j(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + 864299970q^3 + \cdots,$$

where we set $q = \exp(2\pi i \tau)$.

This function has modular invariance property

$$j(\tau) = j\left(\frac{a\tau + b}{c\tau + d}\right),$$

for

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}),$$

and this property and the condition that the top term of the Laurent series of q start with q^{-1} determine the j-function uniquely except for the constant term.

McKay noticed that the first non-trivial coefficient of the Laurent expansion of the j-function except for the constant term is 196884 which is "almost" 196883. Extending this idea, Conway and Norton [1979] formulated the Moonshine conjecture as follows.

Conjecture 6.1. *1. We have some graded infinite dimensional $\mathbb{C}$-vector space $V = \bigoplus_{n=0}^{\infty} V_n$ $(\dim V_n < \infty)$ with some natural algebraic structure and its automorphism group is the Monster group.*

2. Each element g of the Monster group acts on each V_n linearly. The Laurent series

$$\sum_{n=0}^{\infty} (\operatorname{Tr} g|_{V_n}) q^{n-1}$$

arising from the trace value of the g-action on V_n is a classical function called a Hauptmodul corresponding to a genus 0 subgroup of $SL(2, \mathbb{R})$. (The case g is the identity element is the j-function without the constant term.)

"Some natural algebraic structure" in the above conjecture has been formulated as a vertex operator algebra in Frenkel, Lepowsky, and Meurman [1988] and the full Moonshine conjecture has been proved by Borcherds [1992]. The axioms of a vertex operator algebra are given as follows.

Let V be a $\mathbb{C}$-vector space. We say that a formal series $a(z) = \sum_{n\in\mathbb{Z}} a_{(n)} z^{-n-1}$ with coefficients $a_{(n)} \in \mathrm{End}(V)$ is a field on V, if for any $b \in V$, we have $a_{(n)}b = 0$ for all sufficiently large n.

A $\mathbb{C}$-vector space V is called a vertex algebra if we have the following properties.

1. (State-field correspondence) For each $a \in V$, we have a field $Y(a,z) = \sum_{n\in\mathbb{Z}} a_{(n)} z^{-n-1}$ on V.

2. (Translation covariance) We have a linear map $T \in \mathrm{End}(V)$ such that we have $[T, Y(a,z)] = \frac{d}{dz} Y(a,z)$ for all $a \in V$.

3. (Existence of the vacuum vector) We have a vector $\Omega \in V$ with $T\Omega = 0$, $Y(\Omega, z) = \mathrm{id}_V$, $a_{(-1)}\Omega = a$.

4. (Locality) For all $a, b \in V$, we have $(z-w)^N [Y(a,z), Y(b,w)] = 0$ for a sufficiently large integer N.

We then call $Y(a,z)$ a vertex operator. (The locality axiom is one representation of the idea that $Y(a,z)$ and $Y(b,w)$ should commute for $z \neq w$.)

Let V be a $\mathbb{C}$-vector space and $L(z) = \sum_{n\in\mathbb{Z}} L_n z^{-n-2}$ be a field on V. If the endomorphisms L_n satisfy the Virasoro algebra relations

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{(m^3-m)\delta_{m+n,0}}{12} c,$$

with central charge $c \in \mathbb{C}$, then we say $L(z)$ is a Virasoro field. If V is a vertex algebra and $Y(\omega, z) = \sum_{n\in\mathbb{Z}} L_n z^{-n-2}$ is a Virasoro field, then we say $\omega \in V$ is a Virasoro vector. A Virasoro vector ω is called a conformal vector if $L_{-1} = T$ and L_0 is diagonalizable on V, that is, V is an algebraic direct sum of the eigenspaces of L_0. Then the corresponding vertex operator $Y(\omega, z)$ is called the energy-momentum field and L_0 the conformal Hamiltonian. A vertex algebra with a conformal vector is called a conformal vertex algebra. We then say V has central charge $c \in \mathbb{C}$.

A nonzero element a of a conformal vertex algebra in $\mathrm{Ker}(L_0 - \alpha)$ is said to be a homogeneous element of conformal weight $d_a = \alpha$. We then set $a_n = a_{(n+d_a-1)}$ for $n \in \mathbb{Z} - d_a$. For a sum a of homogeneous elements, we extend a_n by linearity.

A homogeneous element a in a conformal vertex algebra V and the corresponding field $Y(a,z)$ are called quasi-primary if $L_1 a = 0$ and primary if $L_n a = 0$ for all $n > 0$.

We say that a conformal vertex algebra V is of CFT type if we have $\mathrm{Ker}(L_0-\alpha)\neq 0$ only for $\alpha\in\{0,1,2,3,\dots\}$ and $V_0=\mathbb{C}\Omega$.

We say that a conformal vertex algebra V is a vertex operator algebra if we have the following.

1. We have $V=\bigoplus_{n\in\mathbb{Z}} V_n$, where $V_n=\mathrm{Ker}(L_0-n)$.
2. We have $V_n=0$ for all sufficiently small n.
3. We have $\dim(V_n)<\infty$ for $n\in\mathbb{Z}$.

Basic sources of constructing vertex operator algebras are affine Kac-Moody and Virasoro algebras due to Frenkel-Zhu and even lattices due to Frenkel-Lepowsky-Meurman. Methods to construct new examples from known examples are a tensor product, a simple current extension due Schellekens-Yankielowicz, orbifold construction due to Dijkgraaf-Vafa-Verlinde-Verlinde, coset construction due to Frenkel-Zhu, and an extension by a Q-system due to Huang-Kirillov-Lepowsky. These are parallel to constructions of local conformal nets, but constructions of vertex operator algebras are earlier except for the extension by a Q-system.

A representation theory of a vertex operator algebra is known as a theory of modules. It has been shown by Huang that we have a modular tensor category for a well-behaved vertex operator algebra. (The well-behavedness condition is basically the so-called C_2-cofiniteness.)

7 From a vertex operator algebra to a local conformal net and back

We now would like to construct a local conformal net from a vertex operator algebra V. First of all, we need a Hilbert space of states, and it should be the completion of V with respect to some natural inner product. A vertex operator algebra with such an inner product is called unitary. Many vertex operator algebras are unitary, but also many others are non-unitary. In order to have the corresponding local conformal net, we definitely have to assume that V is unitary. We now give a precise definition of a unitary vertex operator algebra.

An invariant bilinear form on a vertex operator algebra V is a bilinear form $(\cdot,\cdot)$ on V satisfying

$$(Y(a,z)b,c)=(b,Y(e^{zL_1}(-z^{-2})^{L_0}a,z^{-1})c)$$

for all $a,b,c\in V$.

For a vertex operator algebra V with a conformal vector ω, an automorphism g as a vertex algebra is called a VOA automorphism if we have $g(\omega)=\omega$.

Let V be a vertex operator algebra and suppose we have a positive definite inner product $(\cdot \mid \cdot)$, where we assume this is antilinear in the first variable. We say the inner product is normalized if we have $(\Omega \mid \Omega) = 1$. We say that the inner product is invariant if there exists a VOA antilinear automorphism θ of V such that $(\theta\cdot \mid \cdot)$ is an invariant bilinear form on V. We say that θ is a PCT operator associated with the inner product.

If we have an invariant inner product, we automatically have $(L_n a \mid b) = (a \mid L_{-n} b)$ for $a, b \in V$ and also $V_n = 0$ for $n < 0$. The PCT operator θ is unique and we have $\theta^2 = 1$ and $(\theta a \mid \theta b) = (b \mid a)$ for all $a, b \in V$. (See Carpi, Kawahigashi, Longo, and Weiner [n.d., Section 5.1] for details.)

A unitary vertex operator algebra V is a pair of a vertex operator algebra and a normalized invariant inner product. It is simple if we have $V_0 = \mathbb{C}\Omega$.

Now suppose V is a unitary vertex operator algebra. A vertex operator $Y(a,z)$ should mean a Fourier expansion of an operator-valued distribution on S^1. For a test function f with Fourier coefficients $\hat{f}_n$, the action of the distribution $Y(a,z)$ applied to the test function f on $b \in V$ should be $\sum_{n\in\mathbb{Z}} \hat{f}_n a_n b$. In order to make sense out of this, we need convergence of this infinite sum. To insure such convergence, we introduce the following notion of energy-bounds.

Let $(V, (\cdot \mid \cdot))$ be a unitary vertex operator algebra. We say that $a \in V$ (or $Y(a,z)$) satisfies energy-bounds if we have positive integers s, k and a constant $M > 0$ such that we have

$$\|a_n b\| \le M(|n|+1)^s \|(L_0+1)^k b\|,$$

for all $b \in V$ and $n \in \mathbb{Z}$. If every $a \in V$ satisfies energy-bounds, we say V is energy-bounded.

We have the following Proposition in Carpi, Kawahigashi, Longo, and Weiner [ibid.].

Proposition 7.1. *If V is a simple unitary vertex operator algebra generated by V_1 and $F \subset V_2$ where F is a family of quasi-primary θ-invariant Virasoro vectors, then V is energy-bounded.*

We now assume V is energy-bounded. Let H be the completion of V with respect to the inner product. For any $a \in V$ and $n \in \mathbb{Z}$, we regard $a_{(n)}$ as a densely defined operator on H. This turns out to be closable. Let $f(z)$ be a smooth function on $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ with Fourier coefficients

$$\hat{f}_n = \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta} \frac{d\theta}{2\pi}$$

for $n \in \mathbb{Z}$. For every $a \in V$, we define the operator $Y_0(a, f)$ with domain V by

$$Y_0(a,f)b = \sum_{n\in\mathbb{Z}} \hat{f}_n a_n b$$

for $b \in V$. The convergence follows from the energy-bounds and $Y_0(a, f)$ is a densely defined operator. This is again closable. We denote by $Y(a, f)$ the closure of $Y_0(a, f)$ and call it a smeared vertex operator.

We define $\mathcal{A}_{(V,(\cdot|\cdot))}(I)$ to be the von Neumann algebra generated by the (possibly unbounded) operators $Y(a, f)$ with $a \in V$, $f \in C^\infty(S^1)$ and supp $f \subset I$. The family $\{\mathcal{A}_{(V,(\cdot|\cdot))}(I)\}$ clearly satisfies isotony. We can verify that $(\bigvee_I \mathcal{A}_{(V,(\cdot|\cdot))}(I))\Omega$ is dense in H. A proof of conformal covariance is nontrivial, but can be done as in Toledano Laredo [1999] by studying the representations of the Virasoro algebra and Diff(S^1). We also have the vacuum vector Ω and the positive energy condition. However, locality is not clear at all from our construction, so we make the following definition.

We say that a unitary vertex operator algebra $(V, (\cdot \mid \cdot))$ is strongly local if it is energy-bounded and we have $\mathcal{A}_{(V,(\cdot|\cdot))}(I) \subset \mathcal{A}_{(V,(\cdot|\cdot))}(I')'$ for all intervals $I \subset S^1$.

A strongly local unitary vertex operator algebra produces a local conformal net through the above procedure by definition, but the definition of strong locality looks like we assume what we want to prove, and it would be useless unless we have a good criterion for strong locality. The following theorem gives such a criterion Carpi, Kawahigashi, Longo, and Weiner [n.d.].

Theorem 7.2. *Let V be a simple unitary vertex operator algebra generated by $V_1 \cup F$ where $F \subset V_2$ is a family of quasi-primary θ-invariant Virasoro vectors, then V is strongly local.*

The above criteria applies to vertex operator algebras arising from the affine Kac-Moody and Virasoro algebras. We also have the following theorem which we can apply to many examples Carpi, Kawahigashi, Longo, and Weiner [ibid.].

Theorem 7.3. *(1) Let V_1, V_2 be simple unitary strongly local vertex operator algebras. Then $V_1 \otimes V_2$ is also strongly local.*

(2) Let V be a simple unitary strongly local vertex operator algebra and W its subalgebra. Then W is also strongly local.

The second statement of the above theorem shows that strong locality passes to orbifold and coset constructions, in particular.

For a unitary vertex operator algebra V, we write Aut(V) for the automorphism group of V. For a local conformal net $\{A(I)\}$, we have a notion of the automorphism group and we write Aut(A) for this. We have the following in Carpi, Kawahigashi, Longo, and Weiner [ibid.].

Theorem 7.4. *Let V be a strongly local unitary vertex operator algebra and $\{A_{(V,(\cdot|\cdot))}(I)\}$ the corresponding local conformal net. Suppose* Aut(V) *is finite. Then we have* Aut$(A_{(V,(\cdot|\cdot))}) =$ Aut(V).

The Moonshine vertex operator algebra $V^\natural$ is strongly local and unitary, so we can apply the above result to this to obtain the Moonshine net. It was first constructed in Kawahigashi and Longo [2006] with a more ad-hoc method.

For the converse direction, we have the following Carpi, Kawahigashi, Longo, and Weiner [n.d.].

Theorem 7.5. *Let V be a simple unitary strongly local vertex operator algebra and $\{\mathcal{A}_{(V,(\cdot|\cdot))}(I)\}$ be the corresponding local conformal net. Then one can recover the vertex operator algebra structure on V, which is an algebraic direct sum of the eigenspaces of the conformal Hamiltonian, from the local conformal net $\{\mathcal{A}_{(V,(\cdot|\cdot))}(I)\}$.*

This is proved by using the Tomita-Takesaki theory and extending the methods in Fredenhagen and Jörß [1996]. Establishing correspondence between the representation theories of a vertex operator algebra and a local conformal net is more difficult, though we have some recent progress due to Carpi, Weiner and Xu. The method of Tener [2017] may be more useful for this. We list the following conjecture on this. (For a representation of a local conformal net, we define the character as $\mathrm{Tr}(q^{L_0-c/24})$ when it converges for some small values of q. We have a similar definition for a module of a vertex operator algebra.)

Conjecture 7.6. *We have a bijective correspondence between completely rational local conformal nets and simple unitary C_2-cofinite vertex operator algebras. We also have equivalence of tensor categories for finite dimensional representations of a completely rational local conformal net and modules of the corresponding vertex operator algebra. We further have coincidence of the corresponding characters of the irreducible representations of a completely rational local conformal net and irreducible modules of the corresponding vertex operator algebra.*

Recall that we have a classical correspondence between Lie algebras and Lie groups. The correspondence between affine Kac-Moody algebras and loop groups is similar to this, but "one step higher". Our correspondence between vertex operator algebras and local conformal nets is something even one more step higher.

Finally we discuss the meaning of strong locality. We have no example of a unitary vertex operator algebra which is known to be not strongly local. If there should exist such an example, it would not correspond to a chiral conformal field theory in a physical sense. This means that one of the following holds: any simple unitary vertex operator algebra is strongly local or the axioms of unitary vertex operator algebras are too weak to exclude non-physical examples.

8 Other types of conformal field theories

Here we list operator algebraic treatments of conformal field theories other than chiral ones.

Full conformal field theory is a theory on the 2-dimensional Minkowski space. We axiomatize a net of von Neumann algebras $\{B(I \times J)\}$ parameterized by double cones (rectangles) in the Minkowski space in a similar way to the case of local conformal nets. From this, a restriction procedure produces two local conformal nets $\{A_L(I)\}$ and $\{A_R(I)\}$. We assume both are completely rational. Then we have a subfactor $A_L(I) \otimes A_R(J) \subset B(I \times J)$ which automatically has a finite index, and the study of $\{B(I \times J)\}$ is reduced to studies of $\{A_L(I)\}$, $\{A_R(I)\}$ and this subfactor. A modular invariant again naturally appears here and we have a general classification theory. For the case of central charge less than 1, we obtain a complete and concrete classification result as in Kawahigashi and Longo [2004b].

A boundary conformal field theory is a quantum field theory on the half-Minkowski space $\{(t, x) \in \mathfrak{M} \mid x > 0\}$. The first general theory to deal with this setting was given in Longo and Rehren [2004]. We have more results in Carpi, Kawahigashi, and Longo [2013] and Bischoff, Kawahigashi, and Longo [2015]. In this case, a restriction procedure gives one local conformal net. We assume that this is completely rational. Then we have a non-local, but relatively local extension of this completely rational local conformal net which automatically has a finite index. The study of a boundary conformal field theory is reduced to studies of this local conformal net and a non-local extension. For the case of central charge less than 1, we obtain a complete and concrete classification result as in Kawahigashi, Longo, Pennig, and Rehren [2007] along the line of this general theory.

We also have results on the phase boundaries and topological defects in the operator algebraic framework in Bischoff, Kawahigashi, Longo, and Rehren [2016], Bischoff, Kawahigashi, Longo, and Rehren [2015]. See Fuchs, Runkel, and Schweigert [2004] for earlier works on topological defects.

A superconformal field theory is a version of $\mathbb{Z}_2$-graded conformal field theory having extra supersymmetry. We have operator algebraic versions of $N = 1$ and $N = 2$ superconformal field theories as in Carpi, Kawahigashi, and Longo [2008] and Carpi, Hillier, Kawahigashi, Longo, and Xu [2015] based on $N = 1$ and $N = 2$ super Virasoro algebras, and there we have superconformal nets rather than local conformal nets. We also have relations of this theory to noncommutative geometry in Kawahigashi and Longo [2005], Carpi, Hillier, Kawahigashi, and Longo [2010], Carpi, Hillier, Kawahigashi, Longo, and Xu [2015].

9 Future directions

We list some problems and conjectures for the future studies at the end of this article.

Conjecture 9.1. *For a completely rational local conformal net, we have convergent characters for all irreducible representations and they are closed under modular transformations of $SL(2,\mathbb{Z})$. Furthermore, the S-matrix defined with braiding gives transformation rules of the characters under the transformation $\tau \mapsto -1/\tau$.*

This conjecture was made in Gabbiani and Fröhlich [1993, page 625] and follows from Conjecture 7.6.

We say that a local conformal net is holomorphic if its only irreducible representation is the vacuum representation. The following is Xu [2009, Conjecture 3.4] which is the operator algebraic counterpart of the famous uniqueness conjecture of the Moonshine vertex operator algebra.

Conjecture 9.2. *A holomorphic local conformal net with $c = 24$ and the eigenspace of L_0 with eigenvalue 1 being 0 is unique up to isomorphism.*

A reason to expect such uniqueness from an operator algebraic viewpoint is that a set of simple algebraic invariants should be a complete invariant as long as we have some kind of amenability, which is automatic in the above case.

The following is an operator algebraic counterpart of Höhn [2003, Conjecture 3.5].

Conjecture 9.3. *Fix a modular tensor category $\mathcal{C}$ and a central charge c. Then we have only finitely many local conformal nets with representation category $\mathcal{C}$ and central charge c.*

From an operator algebraic viewpoint, the following problem is also natural.

Problem 9.4. *Suppose a finite group G is given. Construct a local conformal net whose automorphism group is G in some canonical way.*

This "canonical" method should produce the Moonshine net if G is the Monster group. We may have to consider some superconformal nets rather than local conformal nets to get a nice solution.

Conformal field theory on Riemann surfaces has been widely studied and conformal blocks play a important role there. It is not clear at all how to formulate this in our operator algebraic approach to conformal field theory, so we have the following problem.

Problem 9.5. *Formulate a conformal field theory on a Riemann surface in the operator algebraic approach.*

It is expected that the $N = 2$ full superconformal field theory is related to Calabi-Yau manifolds, so we also list the following problem.

Problem 9.6. *Construct an operator algebraic object corresponding to a Calabi-Yau manifold in the setting of $N = 2$ full superconformal field theory and study the mirror symmetry in this context.*

The structure of a modular tensor category naturally appears also in the context of topological phases of matters and anyon condensation as in Kawahigashi [2015a], Kawahigashi [2017], Kong [2014]. (The results in Böckenhauer, Evans, and Kawahigashi [2001] can be also seen in this context.) We list the following problem.

Problem 9.7. *Relate local conformal nets directly with topological phases of matters and anyon condensation.*

References

M. Asaeda and U. Haagerup (1999). "Exotic subfactors of finite depth with Jones indices $(5+\sqrt{13})/2$ and $(5+\sqrt{17})/2$". *Comm. Math. Phys.* 202.1, pp. 1–63. MR: 1686551 (cit. on p. 2621).

Stephen Bigelow, Emily Peters, Scott Morrison, and Noah Snyder (2012). "Constructing the extended Haagerup planar algebra". *Acta Math.* 209.1, pp. 29–82. MR: 2979509 (cit. on p. 2621).

Marcel Bischoff, Yasuyuki Kawahigashi, and Roberto Longo (2015). "Characterization of 2D rational local conformal nets and its boundary conditions: the maximal case". *Doc. Math.* 20, pp. 1137–1184. MR: 3424476 (cit. on p. 2630).

Marcel Bischoff, Yasuyuki Kawahigashi, Roberto Longo, and Karl-Henning Rehren (2015). *Tensor categories and endomorphisms of von Neumann algebras—with applications to quantum field theory*. Vol. 3. SpringerBriefs in Mathematical Physics. Springer, Cham, pp. x+94. MR: 3308880 (cit. on p. 2630).

– (2016). "Phase boundaries in algebraic conformal QFT". *Comm. Math. Phys.* 342.1, pp. 1–45. MR: 3455144 (cit. on p. 2630).

J. Böckenhauer and D. E. Evans (1998). "Modular invariants, graphs and α-induction for nets of subfactors. I". *Comm. Math. Phys.* 197.2, pp. 361–386. MR: 1652746 (cit. on pp. 2618, 2622).

Jens Böckenhauer, David E. Evans, and Yasuyuki Kawahigashi (1999). "On α-induction, chiral generators and modular invariants for subfactors". *Comm. Math. Phys.* 208.2, pp. 429–487. MR: 1729094 (cit. on p. 2622).

– (2000). "Chiral structure of modular invariants for subfactors". *Comm. Math. Phys.* 210.3, pp. 733–784. MR: 1777347 (cit. on p. 2622).

– (2001). "Longo-Rehren subfactors arising from α-induction". *Publ. Res. Inst. Math. Sci.* 37.1, pp. 1–35. MR: 1815993 (cit. on p. 2632).

Richard E. Borcherds (1992). "Monstrous moonshine and monstrous Lie superalgebras". *Invent. Math.* 109.2, pp. 405–444. MR: 1172696 (cit. on p. 2625).

A. Cappelli, C. Itzykson, and J.-B. Zuber (1987). "The A-D-E classification of minimal and $A_1^{(1)}$ conformal invariant theories". *Comm. Math. Phys.* 113.1, pp. 1–26. MR: 918402 (cit. on p. 2623).

S. Carpi, Y. Kawahigashi, R. Longo, and M. Weiner (n.d.). "From vertex operator algebras to conformal nets and back". To appear in *Mem. Amer. Math. Soc.* (cit. on pp. 2627–2629).

Sebastiano Carpi, Robin Hillier, Yasuyuki Kawahigashi, and Roberto Longo (2010). "Spectral triples and the super-Virasoro algebra". *Comm. Math. Phys.* 295.1, pp. 71–97. MR: 2585992 (cit. on p. 2630).

Sebastiano Carpi, Robin Hillier, Yasuyuki Kawahigashi, Roberto Longo, and Feng Xu (2015). "$N = 2$ superconformal nets". *Comm. Math. Phys.* 336.3, pp. 1285–1328. MR: 3324145 (cit. on p. 2630).

Sebastiano Carpi, Yasuyuki Kawahigashi, and Roberto Longo (2008). "Structure and classification of superconformal nets". *Ann. Henri Poincaré* 9.6, pp. 1069–1121. MR: 2453256 (cit. on p. 2630).

– (2013). "How to add a boundary condition". *Comm. Math. Phys.* 322.1, pp. 149–166. MR: 3073161 (cit. on p. 2630).

J. H. Conway and S. P. Norton (1979). "Monstrous moonshine". *Bull. London Math. Soc.* 11.3, pp. 308–339. MR: 554399 (cit. on pp. 2616, 2624).

Chongying Dong and Feng Xu (2006). "Conformal nets associated with lattices and their orbifolds". *Adv. Math.* 206.1, pp. 279–306. MR: 2261756 (cit. on p. 2618).

Sergio Doplicher, Rudolf Haag, and John E. Roberts (1971). "Local observables and particle statistics. I". *Comm. Math. Phys.* 23, pp. 199–230. MR: 0297259 (cit. on pp. 2615, 2619).

David E. Evans and Terry Gannon (2011). "The exoticness and realisability of twisted Haagerup-Izumi modular data". *Comm. Math. Phys.* 307.2, pp. 463–512. MR: 2837122 (cit. on p. 2622).

David E. Evans and Yasuyuki Kawahigashi (1998). *Quantum symmetries on operator algebras*. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, pp. xvi+829. MR: 1642584 (cit. on p. 2620).

K. Fredenhagen, K.-H. Rehren, and B. Schroer (1989). "Superselection sectors with braid group statistics and exchange algebras. I. General theory". *Comm. Math. Phys.* 125.2, pp. 201–226. MR: 1016869 (cit. on pp. 2619, 2620).

Klaus Fredenhagen and Martin Jörß (1996). "Conformal Haag-Kastler nets, pointlike localized fields and the existence of operator product expansions". *Comm. Math. Phys.* 176.3, pp. 541–554. MR: 1376431 (cit. on p. 2629).

Igor Frenkel, James Lepowsky, and Arne Meurman (1988). *Vertex operator algebras and the Monster*. Vol. 134. Pure and Applied Mathematics. Academic Press, Inc., Boston, MA, pp. liv+508. MR: 996026 (cit. on pp. 2616, 2623, 2625).

Jürgen Fuchs, Ingo Runkel, and Christoph Schweigert (2004). "TFT construction of RCFT correlators. III. Simple currents". *Nuclear Phys. B* 694.3, pp. 277–353. MR: 2076134 (cit. on p. 2630).

Fabrizio Gabbiani and Jürg Fröhlich (1993). "Operator algebras and conformal field theory". *Comm. Math. Phys.* 155.3, pp. 569–640. MR: 1231644 (cit. on pp. 2618, 2631).

Gerald Höhn (2003). "Genera of vertex operator algebras and three-dimensional topological quantum field theories". In: *Vertex operator algebras in mathematics and physics (Toronto, ON, 2000)*. Vol. 39. Fields Inst. Commun. Amer. Math. Soc., Providence, RI, pp. 89–107. MR: 2029792 (cit. on p. 2631).

Yi-Zhi Huang, Alexander Kirillov Jr., and James Lepowsky (2015). "Braided tensor categories and extensions of vertex operator algebras". *Comm. Math. Phys.* 337.3, pp. 1143–1159. MR: 3339173 (cit. on p. 2618).

Masaki Izumi (2000). "The structure of sectors associated with Longo-Rehren inclusions. I. General theory". *Comm. Math. Phys.* 213.1, pp. 127–179. MR: 1782145 (cit. on p. 2621).

Masaki Izumi, Roberto Longo, and Sorin Popa (1998). "A Galois correspondence for compact groups of automorphisms of von Neumann algebras with a generalization to Kac algebras". *J. Funct. Anal.* 155.1, pp. 25–63. MR: 1622812 (cit. on p. 2623).

V. F. R. Jones (1983). "Index for subfactors". *Invent. Math.* 72.1, pp. 1–25. MR: 696688 (cit. on pp. 2615, 2619–2621).

Vaughan F. R. Jones (1985). "A polynomial invariant for knots via von Neumann algebras". *Bull. Amer. Math. Soc. (N.S.)* 12.1, pp. 103–111. MR: 766964 (cit. on p. 2615).

Vaughan F. R. Jones, Scott Morrison, and Noah Snyder (2014). "The classification of subfactors of index at most 5". *Bull. Amer. Math. Soc. (N.S.)* 51.2, pp. 277–327. MR: 3166042 (cit. on p. 2621).

Yasuyuki Kawahigashi (2015a). "A remark on gapped domain walls between topological phases". *Lett. Math. Phys.* 105.7, pp. 893–899. MR: 3357983 (cit. on p. 2632).

– (2015b). "Conformal field theory, tensor categories and operator algebras". *J. Phys. A* 48.30, pp. 303001, 57. MR: 3367967 (cit. on p. 2616).

– (2017). "A relative tensor product of subfactors over a modular tensor category". *Lett. Math. Phys.* 107.11, pp. 1963–1970. MR: 3714771 (cit. on p. 2632).

Yasuyuki Kawahigashi and Roberto Longo (2004a). "Classification of local conformal nets. Case $c < 1$". *Ann. of Math. (2)* 160.2, pp. 493–522. MR: 2123931 (cit. on pp. 2618, 2623).

– (2004b). "Classification of two-dimensional local conformal nets with $c < 1$ and 2-cohomology vanishing for tensor categories". *Comm. Math. Phys.* 244.1, pp. 63–97. MR: 2029950 (cit. on p. 2630).

– (2005). "Noncommutative spectral invariants and black hole entropy". *Comm. Math. Phys.* 257.1, pp. 193–225. MR: 2163574 (cit. on p. 2630).

– (2006). "Local conformal nets arising from framed vertex operator algebras". *Adv. Math.* 206.2, pp. 729–751. MR: 2263720 (cit. on pp. 2618, 2619, 2629).

Yasuyuki Kawahigashi, Roberto Longo, and Michael Müger (2001). "Multi-interval subfactors and modularity of representations in conformal field theory". *Comm. Math. Phys.* 219.3, pp. 631–669. MR: 1838752 (cit. on pp. 2620, 2622).

Yasuyuki Kawahigashi, Roberto Longo, Ulrich Pennig, and Karl-Henning Rehren (2007). "The classification of non-local chiral CFT with $c < 1$". *Comm. Math. Phys.* 271.2, pp. 375–385. MR: 2287908 (cit. on p. 2630).

Yasuyuki Kawahigashi and Noppakhun Suthichitranont (2014). "Construction of holomorphic local conformal framed nets". *Int. Math. Res. Not. IMRN* 11, pp. 2924–2943. MR: 3214309 (cit. on p. 2619).

Liang Kong (2014). "Anyon condensation and tensor categories". *Nuclear Phys. B* 886, pp. 436–482. MR: 3246855 (cit. on p. 2632).

R. Longo and K.-H. Rehren (1995). "Nets of subfactors". *Rev. Math. Phys.* 7.4. Workshop on Algebraic Quantum Field Theory and Jones Theory (Berlin, 1994), pp. 567–597. MR: 1332979 (cit. on pp. 2618, 2622).

Roberto Longo (1989). "Index of subfactors and statistics of quantum fields. I". *Comm. Math. Phys.* 126.2, pp. 217–247. MR: 1027496 (cit. on pp. 2616, 2619).

– (1994). "A duality for Hopf algebras and for subfactors. I". *Comm. Math. Phys.* 159.1, pp. 133–150. MR: 1257245 (cit. on p. 2619).

Roberto Longo and Karl-Henning Rehren (2004). "Local fields in boundary conformal QFT". *Rev. Math. Phys.* 16.7, pp. 909–960. MR: 2097363 (cit. on p. 2630).

Roberto Longo and Feng Xu (2004). "Topological sectors and a dichotomy in conformal field theory". *Comm. Math. Phys.* 251.2, pp. 321–364. MR: 2100058 (cit. on p. 2620).

Vincenzo Morinelli, Yoh Tanimoto, and Mihály Weiner (2018). "Conformal Covariance and the Split Property". *Comm. Math. Phys.* 357.1, pp. 379–406. MR: 3764574 (cit. on pp. 2619, 2620).

Adrian Ocneanu (1988). "Quantized groups, string algebras and Galois theory for algebras". In: *Operator algebras and applications, Vol. 2*. Vol. 136. London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, pp. 119–172. MR: 996454 (cit. on p. 2621).

Sorin Popa (1994). "Classification of amenable subfactors of type II". *Acta Math.* 172.2, pp. 163–255. MR: 1278111 (cit. on p. 2620).

James E. Tener (2017). "Construction of the unitary free fermion Segal CFT". *Comm. Math. Phys.* 355.2, pp. 463–518. MR: 3681383 (cit. on p. 2629).

Valerio Toledano Laredo (1999). "Integrating unitary representations of infinite-dimensional Lie groups". *J. Funct. Anal.* 161.2, pp. 478–508. MR: 1674631 (cit. on pp. 2618, 2628).

Antony Wassermann (1998). "Operator algebras and conformal field theory. III. Fusion of positive energy representations of $\mathrm{LSU}(N)$ using bounded operators". *Invent. Math.* 133.3, pp. 467–538. MR: 1645078 (cit. on p. 2618).

Feng Xu (1998). "New braided endomorphisms from conformal inclusions". *Comm. Math. Phys.* 192.2, pp. 349–403. MR: 1617550 (cit. on p. 2622).

– (2000a). "Algebraic coset conformal field theories". *Comm. Math. Phys.* 211.1, pp. 1–43. MR: 1757004 (cit. on p. 2618).

– (2000b). "Algebraic orbifold conformal field theories". *Proc. Natl. Acad. Sci. USA* 97.26, pp. 14069–14073. MR: 1806798 (cit. on p. 2618).

– (2007). "Mirror extensions of local nets". *Comm. Math. Phys.* 270.3, pp. 835–847. MR: 2276468 (cit. on p. 2618).

– (2009). "An application of mirror extensions". *Comm. Math. Phys.* 290.1, pp. 83–103. MR: 2520508 (cit. on p. 2631).

Received 2017-11-30.

Yasuyuki Kawahigashi (河東泰之)
Graduate School of Mathematical Sciences
The University of Tokyo
Komaba, Tokyo, 153-8914
Japan

and

Kavli IPMU (WPI)
University of Tokyo
5-1-5 Kashiwanoha
Kashiwa, 277-8583
Japan
yasuyuki@ms.u-tokyo.ac.jp

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2635–2660)

VARIATIONAL FORMULAE FOR THE CAPACITY INDUCED BY SECOND-ORDER ELLIPTIC DIFFERENTIAL OPERATORS

Claudio Landim

Abstract

We review recent progress in potential theory of second-order elliptic operators and on the metastable behavior of Markov processes.

There has been many recent progress in the potential theory of non-reversible Markov processes. We review in this article some of these advances. In Section 1, we present a brief historical overview of potential theory and we introduce the main notions which will appear throughout the article. In Section 2, we present two variational formulae for the capacity between two sets induced by second-order elliptic operators non necessarily self-adjoint. In the following three sections we present applications of these results. In Section 3, we discuss recurrence of Markov processes; in Section 4, we present a sharp estimate for the transition time between two wells in a dynamical system randomly perturbed; and in Section 5, we prove the metastable behavior of this process.

1 Potential theory

We present in this section a brief historical introduction to the Dirichlet principle. The interested reader will find in Kellogg [1967] a full account and references.

From Newton's law of universal gravitation to Laplace's equation. In 1687, Newton enunciated the Law of universal gravitation which states that "every particle of matter in the universe attracts every other particle with a force whose direction is that of the line joining the two, and whose magnitude is directly as the product of their masses, and inversely as the square of their distance from each other". The magnitude F of the force between two particules, one of mass m_1 situated at $x \in \mathbb{R}^3$ and one of mass m_2 situated at $y \in \mathbb{R}^3$ is thus given by

$$F \;=\; \kappa \, \frac{m_1 \, m_2}{\|x-y\|^2} \;, \qquad (1.1)$$

MSC2010: primary 60J45; secondary 60J60, 65Q30.

where $\|(z_1, z_2, z_3)\| = \sqrt{z_1^2 + z_2^2 + z_3^2}$ stands for the Euclidean distance, and κ for a constant which depends only on the units used. In order to avoid the constant κ we choose henceforth the unit of force so that $\kappa = 1$.

Once Newton's gravitation law has been formulated, it is natural to calculate the force exerted on a particle of unit mass by different types of bodies. Consider a body $\mathfrak{B}$ occupying a portion Ω of the space $\mathbb{R}^3$. Assume that its density ρ at each point $z \in \Omega$ is well defined and that it is continuous and bounded as a function of z. By density at z we mean the limit of the ratio between the mass of a portion containing z and the volume of the portion, as the volume of the portion vanishes. By Equation (1.1), the force at a point $x \in \mathbb{R}^3$ is given by

$$F(x) \;=\; \int_\Omega \frac{z - x}{\|z - x\|^3}\, \rho(z)\, dz\;. \tag{1.2}$$

Note that the force is well defined in Ω because $x \mapsto \|x\|^{-2}$ is integrable in a neighborhood of the origin and we assumed the density ρ to be bounded. Equation (1.2) defines, therefore, a vector field $F = (F_1, F_2, F_3) : \mathbb{R}^3 \to \mathbb{R}^3$.

The force field F introduced in Equation (1.2) turns out to be *divergence free* in Ω^c:

$$(\nabla \cdot F)(x) \;:=\; \sum_{j=1}^{3} (\partial_{x_j} F_j)(x) \;=\; 0\,, \quad x \in \Omega^c\,, \tag{1.3}$$

where ∂_{x_j} represents the partial derivative with respect to x_j. It is also *conservative*: Fix a point $x \in \mathbb{R}$, and let $\gamma : [0,1] \to \mathbb{R}^3$ be a smooth, closed path such that $\gamma(0) = \gamma(1) = x$. The integral of the force field along the cycle γ is given by

$$\int F \cdot d\gamma \;:=\; \sum_{j=1}^{3} \int F_j(\gamma)\, d\gamma_j \;=\; \sum_{j=1}^{3} \int_0^1 F_j(\gamma(t))\, \gamma_j'(t)\, dt \;=\; 0\,.$$

As the force field is conservative and the space is simply connected [any two paths with the same endpoints can be continuously deformed one into the other], we may associate a potential $V : \mathbb{R}^3 \to \mathbb{R}$ to the vector field F. Fix a point $x_0 \in \mathbb{R}^3$ and a constant C_0, and let

$$V(x) \;=\; C_0 \;+\; \int F \cdot d\gamma\,, \tag{1.4}$$

where γ is a continuous path from x_0 to x. The potential V is well defined because the force field is conservative, and it is unique up to an additive constant. By requiring it to vanish at infinity, it becomes

$$V(x) \;=\; -\int_{\mathbb{R}^3} \frac{1}{\|z - x\|}\, \rho(z)\, dz\,, \tag{1.5}$$

and it is called the Newton potential of the measure $\rho(z)\,dz$. Moreover, if we represent by ∇V the gradient of V, $\nabla V = (\partial_{x_1} V\,,\, \partial_{x_2} V\,,\, \partial_{x_3} V)$,

$$\nabla V \;=\; F\;. \tag{1.6}$$

Hence, since the force field is divergence-free [Equation (1.3)],

$$\Delta V \;:=\; \nabla\cdot\nabla V \;=\; 0 \quad \text{on } \Omega^c\;, \tag{1.7}$$

which is known as Laplace's differential equation.

This last identity provides an alternative way to compute the force field induced by a body whose density is unknown. Let $\mathfrak{B}$ be a body occupying a portion Ω of the space $\mathbb{R}^3$. Assume that Ω^c is a domain [open and connected] which has a smooth, simply connected boundary, denoted by $\partial\Omega$. Assume, further, that the force field F exerted by the body $\mathfrak{B}$ can be measured at the boundary of Ω. Fix a point $x_0 \in \partial\Omega$, set $V(x_0) = 0$, and extend the definition of V to $\partial\,\Omega$ through Equation (1.4). By Equation (1.7), the potential V solves the equation

$$\begin{cases}(\Delta W)(x) \;=\; 0\,, & x\in\Omega^c\,,\\ W(x) \;=\; V(x)\,, & x\in\partial\Omega\,.\end{cases} \tag{1.8}$$

To derive F, it remains to solve the linear equation (1.8) and to retrieve F from V by Equation (1.6).

The problem of proving the existence of a function satisfying Equation (1.8) or of finding it when it exists is known as the Dirichlet problem, or the first boundary problem of potential theory.

Dirichlet's principle. In 1850, Dirichlet proposed the following argument to prove the existence of a solution to Equation (1.8). It is simpler to present it in the context of masses distributed along surfaces. If mass points disturb, on may think in terms of charges since, according to Coulomb's law, two point charges exert forces on each other which are given by Newton's law with the word mass replaced by charge, except that charges may attract or repel each other.

Consider a bounded domain Ω whose boundary, represented by $\partial\Omega$, is smooth. Let ζ be a surface density on $\partial\Omega$. By Equation (1.5), the potential associated to this mass distribution is given by

$$V(x) \;=\; -\int_{\partial\Omega} \frac{1}{\|z-x\|}\,\zeta(z)\,\sigma(dz)\;,$$

where $\sigma(dz)$ stands for the surface measure. The surface density can be recovered from the potential. By Theorem VI of Chapter VI in Kellogg [1967],

$$\frac{\partial V}{\partial n_+}(x) \;-\; \frac{\partial V}{\partial n_-}(x) \;=\; -\,4\pi\,\zeta(x)\;, \tag{1.9}$$

where n_+, resp. n_-, represents the outward, resp. inward, pointing unit normal vector to $\partial\Omega$.

Denote by $E(\zeta)$ the potential energy of the mass distribution ζ. It corresponds to the total work needed to assemble the distribution from a state of infinite dispersion, and it is given by

$$E(\zeta) \;=\; \frac{1}{2}\int_{\partial\Omega} V(x)\,\zeta(x)\,\sigma(dx)\;.$$

Since, by Equation (1.9), the surface density can be expressed in terms of the potential, we may consider the energy as a function of the potential. After this replacement, as the potential satisfies Laplace's equation (1.7) on $(\partial\Omega)^c$, applying the divergence theorem, we obtain that

$$E(V) \;=\; \frac{1}{8\pi}\int_{\mathbb{R}^3} \|\nabla V(x)\|^2\,dx\;.$$

It is a principle of physics that equilibrium is characterized by the least potential energy consistent with the constraints of the problem. Thus, to prove the existence of a solution of the differential equation (1.8), Dirichlet proposed to consider the variational problem

$$\inf_f \int_{\mathbb{R}^3} \|(\nabla f)(x)\|^2\,dx\;,$$

where the infimum is carried over all smooth functions $f:\mathbb{R}^3\to\mathbb{R}$ such that $f=V$ on $\partial\Omega$.

Mathematicians objected to the argument at an early date, pointing that the infimum might not be attained at an element of the class of functions considered. Weierstrass gave an example showing that the principle was false, and in 1899, Hilbert provided conditions on the surface, the boundary values and the class of functions f admitted, for which the Dirichlet principel could be proved.

Condenser capacity. In electrostatics, the capacity of an isolated conductor is the the total charge the conductor can hold at a given potential.

Let $\Omega_1\subset\Omega_2$ be bounded domains with smooth boundaries represented by Σ_1, Σ_2, respectively. Assume that the closure of Ω_1 is contained in Ω_2. Consider the potential which is equal to 1 at Ω_1, 0 at Ω_2^c, and which satisfies Laplace's equation on $\mathcal{R}=\Omega_2\setminus\overline{\Omega}_1$. Since V satisfies Laplace's equation on $(\Sigma_1\cup\Sigma_2)^c$, this potential can be obtained from a surface distribution concentrated on $\Sigma_1\cup\Sigma_2$. The total mass [charge] on Σ_1 is given by

$$\int_{\Sigma_1}\zeta(x)\,\sigma(dx) \;=\; -\,\frac{1}{4\pi}\int_{\Sigma_1}\frac{\partial V}{\partial n_+}(x)\,\sigma(dx)\;,$$

where the last identity follows from Equation (1.9) and from the fact that the inward derivative vanishes because V is constant in Ω_1. The condenser capacity of Ω_1 relative to Ω_2 is given by

$$\operatorname{cap}(\Omega_1, \Omega_2^c) \;=\; -\,\frac{1}{4\pi}\int_{\Sigma_1} \frac{\partial V}{\partial n_+}(x)\,\sigma(dx)\;, \tag{1.10}$$

The measure $\nu = -\,(1/4\pi)\,(\partial V/\partial n_+)(x)\,\sigma(dx)$ on Σ_1 is called the *harmonic measure*. The capacity of Ω_1 is obtained by letting Ω_2 increase to $\mathbb{R}^d$.

As the potential V is equal to 1 on Ω_1 and 0 at Σ_2, we may insert $V(x)$ in the previous integral, add the integral of the same expression over Σ_2, and then use the divergence theorem and the fact that V is harmonic on the annulus $\mathcal{R} = \Omega_2 \setminus \Omega_1$ to conclude that the previous expression is equal to

$$\frac{1}{4\pi}\int_{\mathcal{R}} \|(\nabla V)(x)\|^2\,dx \;=\; \frac{1}{4\pi}\inf_f \int_{\mathbb{R}^3} \|(\nabla f)(x)\|^2\,dx\;,$$

where the infimum is carried over all smooth functions f such that $f = 1$ on Ω_1 and $f = 0$ on Ω_2^c. This latter formula provides a variational formula for the capacity defined by Equation (1.10), called the Dirichlet principle.

In the next section, we present two variational formulae for the capacity induced by a second-order elliptic operator which is not self-adjoint with respect to the stationary state [as it is the case of the Laplace operator with respect to the Lebesgue measure]. We then present some applications of the formulae.

2 Dirichlet and Thomson principles

In this section, we extend the notion of capacity to the context of general second order differential operators not necessarily self-adjoint. We then provide two variational formulae for the capacity, the so-called Dirichlet and Thomson principles. We will not be precise on the smoothness conditions of the functions and of the boundary of the sets. The interested reader will find in the references rigorous statements.

To avoid integrability conditions at infinity, we state the Dirichlet and the Thomson principles on a finite cube with periodic boundary conditions. Fix $d \geq 1$, and denote by $\mathbb{T}^d = [0,1)^d$ the d-dimensional torus of length 1. Denote by $a(x)$ a uniformly positive-definite matrix whose entries $a_{i,j}$ are smooth functions: There exist $c_0 > 0$ such that for all $x \in \mathbb{T}^d$, $\xi \in \mathbb{R}^d$,

$$\xi\cdot a(x)\,\xi \;\geq\; c_0\,\|\xi\|^2\;, \tag{2.1}$$

where $\eta\cdot\xi$ represents the scalar product in $\mathbb{R}^d$.

Generator. Denote by $\mathfrak{L}$ the generator given by

$$\mathfrak{L} f \;=\; \nabla\cdot(a\,\nabla f) \;+\; b\cdot\nabla f\;, \tag{2.2}$$

where $b : \mathbb{T}^d \to \mathbb{R}^d$ is a smooth vector field. By modifying the drift b we could assume the matrix a to be symmetric. We will not assume this condition for reasons which will become clear in the next sections. There exists a unique Borel probability measure such that $\mu\mathfrak{L} = 0$. This measure is absolutely continuous, $\mu(dx) = m(x)dx$, where m is the unique solution to

$$\nabla\cdot(a^\dagger\,\nabla m) \;-\; \nabla\cdot(b\,m) \;=\; 0\,, \tag{2.3}$$

where $a^\dagger$ stands for the transpose of a. For existence, uniqueness and regularity conditions of solutions of elliptic equations, we refer to Gilbarg and Trudinger [1977]. Let $V(x) = -\log m(x)$, so that $m(x) = e^{-V(x)}$.

Throughout this section A, B represent two closed, disjoint subsets of $\mathbb{T}^d$

(2.4) which are the closure of open sets with smooth boundaries.

For such a set A, denote by $\mu_A(dx)$ the measure $m(x)\sigma(dx)$ on ∂A, where $\sigma(dx)$ represents the surface measure. Hence, for every smooth vector field φ,

$$\oint_{\partial A}\varphi(x)\cdot n_A(x)\,\mu_A(dx) \;=\; \oint_{\partial A}\varphi(x)\cdot n_A(x)\,e^{-V(x)}\,\sigma(dx)\;,$$

where n_A represents the *inward* normal vector to ∂A.

We may rewrite the generator $\mathfrak{L}$ introduced in Equation (2.2) as

$$\mathfrak{L}f \;=\; e^V\nabla\cdot\big(e^{-V}a\nabla f\big) \;+\; c\cdot\nabla f\;,$$

where $c = b + a^\dagger\nabla V$. It follows from (2.3) that

$$\nabla\cdot(e^{-V}c) \;=\; 0\,. \tag{2.5}$$

This implies that the operator $c\cdot\nabla$ is skew-adjoint in $L_2(\mu)$: for any smooth functions f, $g : \mathbb{T}^d \to \mathbb{R}$,

$$\int f\,c\cdot\nabla g\,d\mu \;=\; -\int g\,c\cdot\nabla f\,d\mu\;, \tag{2.6}$$

and that for any open set D with a smooth boundary,

$$\oint_{\partial D} c\cdot n_D\,d\mu_D \;=\; \int_{\mathbb{T}^d\setminus D} e^V\,\nabla\cdot(e^{-V}c)\,d\mu \;=\; 0\,. \tag{2.7}$$

In view of Equation (2.6), the adjoint of $\mathfrak{L}$ in $L_2(\mu)$, represented by $\mathfrak{L}^*$, is given by

$$\mathfrak{L}^*f \;=\; e^V\,\nabla\cdot\big(e^{-V}a^\dagger\nabla f\big) \;-\; c\cdot\nabla f\;,$$

while the symmetric part, denoted by $\mathfrak{L}^s$, $\mathfrak{L}^s = (1/2)(\mathfrak{L} + \mathfrak{L}^*)$, takes the form

$$\mathfrak{L}^s f \;=\; e^V \nabla \cdot \big(e^{-V} a_s \nabla f\big) \;. \tag{2.8}$$

where a_s stands for the symmetrization of the matrix a: $a_s = (1/2)[a + a^\dagger]$.

Capacity. Recall that A, B are closed sets satisfying Equation (2.4). Let

$$\Omega \;=\; \mathbb{T}^d \setminus (A \cup B) \;.$$

Denote by $h = h_{A,B}$, resp. $h^* = h^*_{A,B}$, the unique solutions to the linear elliptic equations

$$\begin{cases} \mathfrak{L}h = 0 & \text{on } \Omega, \\ h = \chi_A & \text{on } A \cup B, \end{cases} \qquad \begin{cases} \mathfrak{L}^*h = 0 & \text{on } \Omega, \\ h^* = \chi_A & \text{on } A \cup B, \end{cases} \tag{2.9}$$

where χ_C, $C \subset \mathbb{T}^d$, represents the indicator function of the set C. The functions h, h^* are called the *equilibrium potentials* between A and B. A function f such that $(\mathfrak{L} f)(x) = 0$ is said to harmonic at x. If it is harmonic at all points in some domain Ω, it is said to be *harmonic* in Ω.

By analogy to the electrostatic definition (1.10) of the capacity of a set, define the *capacity* between the sets A, B of $\mathbb{T}^d$ as

$$\text{cap}(A, B) \;:=\; \oint_{\partial A} a\, \nabla h \cdot n_A \, d\mu_A \;, \quad \text{cap}^*(A, B) \;:=\; \oint_{\partial A} a^\dagger\, \nabla h^* \cdot n_A \, d\mu_A \;. \tag{2.10}$$

Since $h = 1$ at ∂A and $h = 0$ at ∂B, we may insert h in the integral and add the surface integral of the same expression over ∂B. Applying then the divergence theorem, we obtain that

$$\text{cap}(A, B) \;=\; \int_\Omega \nabla h \cdot a\, \nabla h \, d\mu \;+\; \int_\Omega h\, e^V\, \nabla \cdot \big(e^{-V}\, a\, \nabla h\big)\, d\mu \;.$$

As ∇h vanishes on $A \cup B$, we may extend the integrals to $\mathbb{T}^d$. The integrand in the second term can be written as $h\,[\mathfrak{L}h - c \cdot \nabla h]$. Since $c \cdot \nabla$ is skew-adjoint and h is harmonic on $(\partial A \cup \partial B)^c$, we conclude that

$$\text{cap}(A, B) \;=\; \int \nabla h \cdot a\, \nabla h \, d\mu \;, \quad \text{cap}^*(A, B) \;=\; \int \nabla h^*_{A,B} \cdot a^\dagger\, \nabla h^*_{A,B} \, d\mu \;. \tag{2.11}$$

In the previous formulae, we may replace a, $a^\dagger$ by their symmetric part a_s, and we may restrict the integrals to Ω.

Lemma 2.1. *Let A, B be two closed subsets satisfying the conditions* (2.4). *Then,*

$$\text{cap}(A, B) \;=\; \text{cap}(B, A) \;=\; \text{cap}^*(A, B) \;.$$

Moreover,

$$\mathrm{cap}(A,B) \;=\; \oint_{\partial A} (a\,\nabla h + h\,c)\cdot n_A\, d\mu_A\;. \tag{2.12}$$

Proof. It is clear that $\mathrm{cap}(B,A) = \mathrm{cap}(A,B)$ since $h_{B,A} = 1 - h_{A,B}$ as $A\cap B = \varnothing$. The proof of Equation (2.12) is similar to the one which led from the definition of the capacity to Equation (2.11). One has just to recall from Equation (2.5) that $\nabla\cdot(e^{-V}c) = 0$.

We turn to the proof that $\mathrm{cap}(A,B) = \mathrm{cap}^*(A,B)$ It relies on the claim that

$$\mathrm{cap}(A,B) \;=\; \int_{\mathbb{T}^d} \{\nabla h^*\cdot a\,\nabla h - h^*\, c\cdot\nabla h\}\, d\mu \;=\; \int_{\mathbb{T}^d} \{\nabla h\cdot a^\dagger\,\nabla h^* + h\,c\cdot\nabla h^*\}\, d\mu\;.$$

To prove this claim, repeat the calculations carried out to derive Equation (2.12) to conclude that

$$\int_{\mathbb{T}^d} \nabla h^*\cdot a\,\nabla h\, d\mu \;=\; \int_{\Omega} \nabla h^*\cdot a\,\nabla h\, d\mu \;=\; \oint_{\partial A} a\,\nabla h\cdot n_A\, d\mu_A \;+\; \int_{\Omega} h^*\, c\cdot\nabla h\, d\mu\;.$$

Since ∇h vanishes on $A\cup B$, we may carry the second integral over $\mathbb{T}^d$. This proves the first identity of the claim because the first term on the right hand side is equal to the capacity between A and B.

The same computation inverting the roles of h and h^* gives that

$$\int_{\mathbb{T}^d} \nabla h\cdot a^\dagger\,\nabla h^*\, d\mu \;=\; \mathrm{cap}^*(A,B) \;-\; \int_{\mathbb{T}^d} h\,c\cdot\nabla h^*\, d\mu\;.$$

Compare this identity with the previous one. The left-hand sides coincide. As $c\cdot\nabla$ is skew-adjoint, the second terms on the right-hand sides are also equal. Hence, $\mathrm{cap}(A,B) = \mathrm{cap}^*(A,B)$ because the first term on the right-hand side of the penultimate equation is $\mathrm{cap}(A,B)$. The previous identity together with the fact that $\mathrm{cap}^*(A,B) = \mathrm{cap}(A,B)$ yields the second identity of the claim. □

Considering $\mathfrak{L}^*$ in place of $\mathfrak{L}$ we obtain from the previous lemma that

$$\mathrm{cap}^*(A,B) \;=\; \oint_{\partial A} \Big(a^\dagger\,\nabla h^* - h^*\, c\Big)\cdot n_A\, d\mu_A \tag{2.13}$$

Variational formulae for the capacity. Let $\mathfrak{F}$ be the Hilbert space of vector fields $\varphi : \Omega\to\mathbb{R}^d$ endowed with the scalar product $\langle\cdot,\cdot\rangle$ given by:

$$\langle\varphi,\psi\rangle \;:=\; \int_{\Omega} \varphi\cdot a_s^{-1}\,\psi\, d\mu\;.$$

Let $\mathfrak{F}_\gamma$, $\gamma \in \mathbb{R}$, be the closure in $\mathfrak{F}$ of the space of smooth vector fields $\varphi \in \mathfrak{F}$ such that

$$\nabla \cdot (e^{-V} \varphi) = 0 \, , \qquad \oint_{\partial A} \varphi \cdot n_A \, d\mu_A \;=\; -\gamma \, . \tag{2.14}$$

Let $\mathfrak{C}_{\alpha,\beta}$, α, $\beta \in \mathbb{R}$, be the space of smooth functions $f : \Omega \to \mathbb{R}$ such that $f \equiv \alpha$ on A and $f \equiv \beta$ on B. For $f \in \mathfrak{C}_{\alpha,\beta}$ define

$$\Psi_f \;:=\; a_s \, \nabla f \;, \quad \Phi_f \;:=\; a^\dagger \, \nabla f \;-\; f \, c \; .$$

Note that

$$\langle \Psi_h, \Psi_h \rangle \;=\; \int_\Omega \nabla h \cdot a_s \, \nabla h \, d\mu \;=\; \mathrm{cap}(A, B) \; . \tag{2.15}$$

Lemma 2.2. *For every $\varphi \in \mathfrak{F}_\gamma$ and $f \in \mathfrak{C}_{\alpha,0}$ we have that*

$$\langle \Phi_f - \varphi \, , \, \Psi_h \rangle \;=\; \gamma \;+\; \alpha \, \mathrm{cap}(A, B) \; .$$

Proof. Fix $\varphi \in \mathfrak{F}_\gamma$ and $f \in \mathfrak{C}_{\alpha,0}$. By definition of Φ_f,

$$\langle \Phi_f - \varphi, \Psi_h \rangle \;=\; \int_\Omega \Big(a^\dagger \, \nabla f \;-\; f \, c \;-\; \varphi \Big) \cdot \nabla h \, d\mu \; .$$

Writing $a^\dagger \, \nabla f \cdot \nabla h$ as $\nabla f \cdot a \nabla h$, and integrating by parts, since $f = \alpha$ on ∂A and $f = 0$ on ∂B, the previous term becomes

$$-\int_\Omega \Big(f \, e^V \nabla \cdot \big(e^{-V} a \, \nabla h \big) \;+\; f \, c + \varphi \Big) \cdot \nabla h \, d\mu \;+\; \alpha \oint_{\partial A} a \, \nabla h \cdot n_A \, d\mu_A \; .$$

By definition, the last integral is the capacity between A and B, while the expression involving f is equal to $-f \, \mathfrak{L} h$. This expression vanishes because h is $\mathfrak{L}$-harmonic in Ω. Hence, since $h = \chi_A$ on $\partial A \cup \partial B$, by an integration by part, the previous expression is equal

$$\int_\Omega h \, e^V \nabla \cdot (e^{-V} \varphi) \, d\mu \;-\; \oint_{\partial A} \varphi \cdot n_A \, d\mu_A \;+\; \alpha \, \mathrm{cap}(A, B) \; .$$

By Equation (2.14), this expression is equal to $\gamma + \alpha \, \mathrm{cap}(A, B)$, as claimed. □

Theorem 2.3 (Dirichlet principle)**.** *Fix two disjoint subsets A, B of $\mathbb{T}^d$ satisfying Equation (2.4). Then,*

$$\mathrm{cap}(A, B) \;=\; \inf_{f \in \mathfrak{C}_{1,0}} \inf_{\varphi \in \mathfrak{F}_0} \langle \Phi_f - \varphi, \Phi_f - \varphi \rangle \; ,$$

and the infimum is attained for $h_\star = (1/2)(h + h^*)$ and $\varphi_\star = \Phi_{h_\star} - \Psi_h$.

Proof. Fix $f \in \mathfrak{C}_{1,0}$ and $\varphi \in \mathfrak{F}_0$. By Lemma 2.2, applied with $\gamma = 0$ and $\alpha = 1$, and by Schwarz inequality,

$$\mathrm{cap}(A,B)^2 \;=\; \langle \Phi_f - \varphi\,,\, \Psi_h\rangle^2 \;\le\; \langle \Phi_f - \varphi\,,\, \Phi_f - \varphi\rangle\, \langle \Psi_h, \Psi_h\rangle\;.$$

By Equation (2.15), the last term is equal to $\mathrm{cap}(A,B)$, so that $\mathrm{cap}(A,B) \le \langle \Phi_f - \varphi, \Phi_f - \varphi\rangle$ for every f in $\mathfrak{C}_{1,0}$ and φ in $\mathfrak{F}_0$.

Recall from the statement of the theorem the definition of $h_\star$ and $\varphi_\star$. Since $\Phi_{h_\star} - \varphi_\star = \Psi_h$, by Equation (2.15), $\mathrm{cap}(A,B) = \langle \Phi_{h_\star} - \varphi_\star\,,\, \Phi_{h_\star} - \varphi_\star\rangle$. Therefore, to complete the proof of the theorem, it remains to check that $h_\star$ belongs to $\mathfrak{C}_{1,0}$, and $\varphi_\star$ to $\mathfrak{F}_0$. It is easy to check the first condition. For the second one, observe that

$$\nabla\cdot(e^{-V}\varphi_\star) \;=\; \frac{1}{2}\, e^{-V}\, (\mathfrak{L}^* h^* - \mathfrak{L} h) \;-\; \frac{1}{2}\,(h + h^*)\,\nabla\cdot(e^{-V} c)\;.$$

This expression vanishes on Ω by the harmonicity of h, h^* and in view of Equation (2.5). On the other hand,

$$\oint_{\partial A} \varphi_\star\cdot n_A\, d\mu_A \;=\; \frac{1}{2}\oint_{\partial A} (a^\dagger\,\nabla h^* - h^* c)\cdot n_A\, d\mu_A \;-\; \frac{1}{2}\oint_{\partial A} (a\,\nabla h + h\, c)\cdot n_A\, d\mu_A\;.$$

By Lemma 2.1 and identity Equation (2.13), the previous expression is equal to $(1/2)\{\mathrm{cap}^*(A,B) - \mathrm{cap}(A,B)\} = 0$, which completes the proof of the theorem. $\square$

Theorem 2.4 (Thomson principle). *Fix two disjoint subsets A, B of $\mathbb{T}^d$ satisfying Equation (2.4). Then,*

$$\frac{1}{\mathrm{cap}(A,B)} \;=\; \inf_{\varphi\in\mathfrak{F}_1}\, \inf_{f\in\mathfrak{C}_{0,0}} \langle \Phi_f - \varphi\,,\, \Phi_f - \varphi\rangle\;.$$

Moreover, the infimum is attained at $h_\star = (h - h^*)/2\,\mathrm{cap}(A,B)$ and $\varphi_\star = \Phi_{h_\star} - \Psi_{h/\mathrm{cap}(A,B)}$.

Proof. Fix φ in $\mathfrak{F}_1$ and f in $\mathfrak{C}_{0,0}$. By Lemma 2.2 (applied with $\alpha = 0$ and $\gamma = 1$), by Schwarz inequality, and by Equation (2.15),

$$1 \;=\; \langle \Phi_f - \varphi\,,\, \Psi_h\rangle^2 \;\le\; \langle \Phi_f - \varphi\,,\, \Phi_f - \varphi\rangle\,\langle \Psi_h\,,\, \Psi_h\rangle \;=\; \langle \Phi_f - \varphi\,,\, \Phi_f - \varphi\rangle\, \mathrm{cap}(A,B)\,.$$

By definition of $h_\star$, $\varphi_\star$, $\Phi_{h_\star} - \varphi_\star = \Psi_{h/\mathrm{cap}(A,B)}$, so that by Equation (2.15),

$$\langle \Phi_{h_\star} - \varphi_\star, \Phi_{h_\star} - \varphi_\star\rangle \;=\; \langle \Psi_{h/\mathrm{cap}(A,B)}\,,\, \Psi_{h/\mathrm{cap}(A,B)}\rangle \;=\; 1/\mathrm{cap}(A,B)\;.$$

It remains to check that $h_\star \in \mathfrak{C}_{0,0}$, and $\varphi_\star \in \mathfrak{F}_1$. It is easy to verify the first condition. For the second one, observe that

$$\varphi_\star \;=\; \frac{-1}{2\,\mathrm{cap}(A,B)}\Big\{\big[\,a\,\nabla h \,+\, h\,c\,\big] + \big[\,a^\dagger\,\nabla h^* \,-\, h^*\,c\,\big]\Big\}\;.$$

Therefore, $2\,\mathrm{cap}(A,B)\,\nabla\cdot(e^{-V}\varphi_\star) = -\,e^{-V}[\mathfrak{L}h + \mathfrak{L}^* h^*] = 0$ on Ω. Moreover,

$$
\begin{aligned}
&-\,2\,\mathrm{cap}(A,B)\oint_{\partial A}\varphi_\star\cdot n_A\,d\mu_A \\
&\quad = \oint_{\partial A}(\,a\,\nabla h + hc\,)\cdot n_A\,d\mu_A \;+\; \oint_{\partial A}(\,a^\dagger\,\nabla h^* - h^* c\,)\cdot n_A\,d\mu_A\;.
\end{aligned}
$$

By Lemma 2.1 and Equation (2.13), the right-hand side is equal to $\mathrm{cap}(A,B) + \mathrm{cap}^*(A,B) = 2\,\mathrm{cap}(A,B)$. This proves that $\varphi_\star$ belongs to $\mathfrak{F}_1$, and completes the proof of the theorem. □

Reversible case. In the reversible case, $c = 0$, a symmetric, the optimal flow φ in the Dirichlet principle is the null one, so that

$$
\mathrm{cap}(A,B) \;=\; \inf_{f\in\mathfrak{C}_{1,0}}\langle\Phi_f,\Phi_f\rangle \;=\; \inf_{f\in\mathfrak{C}_{1,0}}\int_\Omega \nabla f\cdot a\,\nabla f\,d\mu\;. \tag{2.16}
$$

In the last identity we used the fact that $\Phi_f = \Psi_f = a\nabla f$. We thus recover the Dirichlet principle in the reversible context.

Similarly, in the reversible case, the optimal function f in the Thomson principle is the null one, so that

$$
\frac{1}{\mathrm{cap}(A,B)} \;=\; \inf_{\varphi\in\mathfrak{F}_1}\langle\varphi\,,\,\varphi\rangle\;,
$$

which is the Thomson's principle in the reversible case.

We conclude this subsection comparing the capacity induced by the generator $\mathfrak{L}$ with the one induced by the symmetric part of the generator, $\mathfrak{L}^s$ given by Equation (2.8).

Fix two disjoint subsets A, B satisfying Equation (2.4). Denote by $\mathrm{cap}_s(A,B)$, the capacity between A and B induced $\mathfrak{L}^s$. Since h belongs to $\mathfrak{C}_{1,0}$, by Equation (2.11) and Equation (2.16),

$$
\mathrm{cap}_s(A,B) \;\le\; \mathrm{cap}(A,B)\;.
$$

In the case of Markov chains taking value in a countable state-space, it is proved in Lemma 2.6 of Gaudillière and Landim [2014] that if the generator satisfies a sector condition with constant C_0,

$$
\Big(\int(\mathfrak{L}f)\,g\,d\mu\Big)^2 \;\le\; C_0\int(-\,\mathfrak{L}f)\,f\,d\mu\int(-\,\mathfrak{L}g)\,g\,d\mu\;,
$$

for all smooth functions f, g, then $\mathrm{cap}(A,B)\le C_0\,\mathrm{cap}_s(A,B)$.

Stochastic representation. The operators $\mathfrak{L}$ and $\mathfrak{L}^*$ are generators of Markov processes on $\mathbb{T}^d$ with invariant measure μ. More precisely, $\mathfrak{L}$ is the generator of the solution of the stochastic differential equation

$$
dX_t \;=\; \{-\,a^\dagger\,\nabla V \;+\; \nabla\cdot a \;+\; c\}(X_t)\,dt \;+\; \sqrt{2\,a_s}\,dW_t\;, \tag{2.17}
$$

where W_t is a standard d-dimensional Brownian motion, $\sqrt{2a_s}$ represents the symmetric, positive-definite square root of $2a_s$, and $\nabla\cdot a$ is the vector field whose j-th coordinate is $(\nabla\cdot a)_j = \sum_{1\le i\le d}\partial_{x_i} a_{i,j}$. For $\mathfrak{L}^*$, one has to replace the drift in Equation (2.17) by $-a\,\nabla V + \nabla\cdot a^\dagger - c$.

Denote by $C([0,+\infty);\mathbb{T}^d)$ the space of continuous functions $X : [0,+\infty)\to\mathbb{T}^d$ endowed with the topology of locally uniform convergence. Let $\{\mathbb{P}_x : x\in\mathbb{T}^d\}$, resp. $\{\mathbb{P}^*_x : x\in\mathbb{T}^d\}$, be the probability measures on $C([0,+\infty);\mathbb{T}^d)$ induced by the Markov process associated to the generator $\mathfrak{L}$, resp. $\mathfrak{L}^*$, starting from x.

Denote by H_C, C a closed subset of $\mathbb{T}^d$, the hitting time of C:

$$H_C \;=\; \inf\{t\ge 0 : X_t\in C\}\,.$$

Lemma 2.5. *Let C be the closure of an open set with smooth boundary. Consider two continuous functions b, $f:\mathbb{T}^d\to\mathbb{R}$. Let $u:\mathbb{T}^d\to\mathbb{R}$ be given by*

$$u(x) \;:=\; \mathbb{E}_x\Big[\, b(X(H_C)) \;+\; \int_0^{H_C} f(X(t))\,dt\,\Big]\,.$$

Then, u is the unique solution to

$$\begin{cases} \mathfrak{L}u = -f & \text{on } \mathbb{T}^d\setminus C,\\ u = b & \text{on } \partial C.\end{cases} \tag{2.18}$$

This result provides a stochastic representation for the harmonic functions $h = h_{A,B}$, $h^* = h^*_{A,B}$ introduced previously:

$$h(x) \;=\; \mathbb{P}_x[H_A < H_B]\,,\quad h^*(x) \;=\; \mathbb{P}^*_x[H_A<H_B]\,.$$

Harmonic measure. In view of the definition (2.10) of the capacity, define the probability measure $\nu\equiv\nu_{A,B}$ as the harmonic measure on $\partial A\cup\partial B$ conditioned to ∂A as

$$d\nu :=\; \frac{1}{\mathrm{cap}(A,B)}\, a^\dagger\,\nabla h^*\cdot n_A\,d\mu_A\,.$$

Proposition 2.6. *For each continuous function $f:\mathbb{T}^d\to\mathbb{R}$,*

$$\mathbb{E}_\nu\Big[\int_0^{H_B} f(X_t)\,dt\Big] \;=\; \frac{1}{\mathrm{cap}(A,B)}\int h^*\,f\,d\mu\,. \tag{2.19}$$

Proof. Fix a continuous function f, and let $\Omega_B = \mathbb{T}^d\setminus B$. Denote by u the unique solution of the elliptic equation (2.18) with $C=B$, $b\equiv 0$. In view of Lemma 2.5 and by definition of the harmonic measure ν, the left hand side of Equation (2.19) is equal to

$$\frac{1}{\mathrm{cap}(A,B)}\oint_{\partial A} u\,[a^\dagger\,\nabla h^*]\cdot n_A\,d\mu_A\,.$$

The integral of the same expression at ∂B vanishes because u vanishes on ∂B. Hence, by the divergence theorem, this expression is equal to

$$\frac{1}{\mathrm{cap}(A,B)} \int_\Omega e^V \, \nabla \cdot \{ e^{-V} \, [a^\dagger \, \nabla h^*] \, u \} \, d\mu \, .$$

Since the equilibrium potential h^* is harmonic on Ω, the previous equation is equal to

$$\frac{1}{\mathrm{cap}(A,B)} \int_\Omega \nabla h^* \, a \, \nabla u \, d\mu \; + \; \frac{1}{\mathrm{cap}(A,B)} \int_\Omega u \, c \cdot \nabla h^* \, d\mu \, .$$

Consider the first term. Since ∇h^* vanishes on A, we may extend the integral to $\Omega_B = \mathbb{T}^d \setminus B$. By the divergence theorem and since the equilibrium potential h^* vanishes on ∂B, this expression is equal to

$$- \frac{1}{\mathrm{cap}(A,B)} \int_{\Omega_B} h^* \, e^V \, \nabla \cdot \{ e^{-V} \, a \, \nabla u \} \, d\mu \, .$$

As $\mathfrak{L} u = -f$ on Ω_B, this expression is equal to

$$\frac{1}{\mathrm{cap}(A,B)} \int_{\Omega_B} h^* \, c \cdot \nabla u \, d\mu \; + \; \frac{1}{\mathrm{cap}(A,B)} \int_{\Omega_B} h^* \, f \, d\mu \, .$$

Since the equilibrium potential h^* vanishes on B, we may replace Ω_B by $\mathbb{T}^d$ in the last integral.

Up to this point we proved that the left-hand side of Equation (2.19) is equal to

$$\frac{1}{\mathrm{cap}(A,B)} \Big\{ \int_{\mathbb{T}^d} h^* \, f \, d\mu \; + \; \int_{\Omega_B} h^* \, c \cdot \nabla u \, d\mu \; + \; \int_\Omega u \, c \cdot \nabla h^* \, d\mu \Big\} \, .$$

Since ∇h^* vanishes on $A \cup B$ and h^* on B, we may extend the last two integrals to $\mathbb{T}^d$. By Equation (2.6), the sum of the last two terms vanishes, which completes the proof of the proposition. □

Proposition 2.6 is due to Bovier, Eckhoff, Gayrard, and Klein [2001] for reversible Markov chains. A generalization to non-reversible chains can be found in Beltrán and Landim [2012b]. A Dirichlet principle, as a double variational formula of type $\inf_f \sup_g$ involving functions, was proved by Pinsky [1988a,b, 1995] in the context of diffusions. It has been derived by Doyle [1994] and by Gaudillière and Landim [2014] for Markov chains. The Dirichlet principle, stated in Theorem 2.3, appeared in Gaudillière and Landim [ibid.] for Markov chains and is due to Landim, Mariani, and Seo [2017] in the context of diffusion processes. The Thomson principle, stated in Theorem 2.4, is due to Slowik [2012] in the context of Markov chains and appeared in Landim, Mariani, and Seo [2017] for diffusions.

3 Recurrence of Markov chains

The capacity is an effective tool to prove the recurrence or transience of Markov processes whose stationary state are explicitly known.

Consider the following open problem. Let $\boldsymbol{X} = \{(X_k, Y_k) : k \in \mathbb{Z}\}$, be a sequence of independent, identically distributed random variables such that $P[(X_0, Y_0) = (\pm 1, \pm 1)] = 1/4$ for all 4 combinations of signs. Given a random environment $\boldsymbol{X}$ consider the discrete-time random walk on $\mathbb{Z}^2$ whose jump probabilities are given by

$$p\big((j,k)\,,\,(j+Y_k,k)\big) \;=\; p\big((j,k)\,,\,(j,k+X_j)\big) \;=\; 1/2 \;\text{ for all } (j,k)\in\mathbb{Z}^2\,. \tag{3.1}$$

Denote by $Z_t = (Z_t^1, Z_t^2)$ the position at time $t \in \mathbb{Z}$ of the random walk. Equation (3.1) states that in the horizontal line $\{(p,q) : q = k\}$ Z only jumps from (j,k) to $(j+Y_k, k)$ for all j. In other words, on each horizontal line the random walk is totally asymmetric, but the direction of the jumps may be differ from line to line. Similarly, on the vertical lines $\{(p,q) : p = j\}$ the random walk is totally asymmetric and only jumps from (j,k) to $(j, k+X_j)$. It is not known if this random walk is recurrent or not [almost surely with respect to the random environment].

Fix an environment $\boldsymbol{X}$, and let $\boldsymbol{P}^{\boldsymbol{X}}_{(j,k)}$ be the distribution of the random walk Z which starts at time $t = 0$ from (j,k). Denote by H_0^+ the return time to 0: $H_0^+ = \inf\{t \geq 1 : Z_t = 0\}$. The random walk is recurrent if and only if $\boldsymbol{P}^{\boldsymbol{X}}_0[H_0^+ = \infty] = 0$. Let $\{B_N : N \geq 1\}$ be an increasing sequence of finite sets such that $\cup_N B_N = \mathbb{Z}^2$, and note that

$$\boldsymbol{P}^{\boldsymbol{X}}_0[H_0^+ = \infty] \;=\; \lim_{N\to\infty} \boldsymbol{P}^{\boldsymbol{X}}_0[H_{B_N^c} < H_0^+]\,, \tag{3.2}$$

where $H_{B_N^c}$ stands for the hitting time of B_N^c.

In the context of discrete-time Markov chains evolving on a countable state-space the capacity between two disjoint sets A, B is given by

$$\operatorname{cap}(A,B) \;=\; \sum_{x\in A} M(x)\, P_x[H_B < H_A^+]\,,$$

where M represents the stationary state of the Markov chain and H_B, resp. H_A^+, the hitting time of the set B, resp. the return time to the set A.

By the previous identity, the right-hand side of Equation (3.2) can be rewritten as

$$\frac{1}{M_{\boldsymbol{X}}(0)} \lim_{N\to\infty} \operatorname{cap}_{\boldsymbol{X}}\big(\{0\}\,,\; B_N^c\big)\,,$$

where $M_{\boldsymbol{X}}$ represents the stationary state of the random walk. It is easy to show that $M_{\boldsymbol{X}}$ does not depend on the environment and is constant, $M_{\boldsymbol{X}}(z) = 1$ for all $z \in \mathbb{Z}^2$.

In view of the Dirichlet principle, to prove that the random walk is recurrent, one needs to find a sequence of functions f_N in $\mathcal{C}_{1,0}$ and of vector fields φ_N in $\mathfrak{F}_0$ [with $A = \{0\}$ $B = B_N^c$ and depending on the environment X] for which $\langle \Phi_{f_N} - \varphi_N , \Phi_{f_N} - \varphi_N \rangle$ vanishes asymptotically.

This has not been achieved yet. However, this is the simplest way to prove that the symmetric, nearest-neighbor random walk on $\mathbb{Z}^2$ is recurrent $[p((j,k),(j,k\pm 1)) = p((j,k),(j\pm 1,k)) = 1/4]$. In this case also $M(z) = 1$ for all $z \in \mathbb{Z}^2$. Consider $B_N = \{-(N-1), \dots, N-1\}^2$, and set $\varphi_N = 0$, $f_N(x) = 1 - \log|x|_m / \log N$, $x \in B_N$, where $|0|_m = 1$, $|(j,k)| = \max\{|j|, |k|\}$. For these sequences,

$$\langle \Phi_{f_N} - \varphi_N , \Phi_{f_N} - \varphi_N \rangle \;=\; \frac{1}{4} \sum_{j=1}^{2} \sum_{x\in\mathbb{Z}^2} [f_N(x+e_j) - f_N(x)]^2 \;\le\; \frac{C_0}{\log N} \;,$$

where $\{e_1, e_2\}$ stands for the canonical basis of $\mathbb{R}^2$ and C_0 for a finite constant independent of N. This proves that the 2-dimensional, nearest-neighbor, symmetric random walk is recurrent.

4 Eyring-Kramers formula for the transition time

We examine in this section the stochastic differential equation (2.17) as a small perturbation of a dynamical system $\dot x(t) = F(x(t))$, by introducing a small parameter $\epsilon > 0$ in the equation.

To reduce the noise in Equation (2.17), we substitute $\sqrt{2\, a_s}$ in the second term of the right-hand side by $\sqrt{2\,\epsilon\, a_s}$. At this point, to keep the structure of the equation, we have to replace in the first term a by $\epsilon\, a$. To avoid the term $-\,a^\dagger\nabla V$ to become small, we change V to V/ϵ. After these modifications the Equation (2.17) becomes

$$dX^\epsilon_t \;=\; \big\{ -\, a^\dagger\,\nabla V \,+\, \epsilon\,\nabla\cdot a \,+\, c\,\big\}(X^\epsilon_t)\, dt \;+\; \sqrt{2\,\epsilon\, a_s}\, dW_t \;. \tag{4.1}$$

The diffusion X^ϵ_t is a small perturbation of the dynamical system $\dot x(t) = -\,[a^\dagger\,\nabla V](x(t)) + c(x(t))$. For the equilibrium points of this ODE to be the critical points of V, we require V to be a Lyapounov functional. This is the case if $c\cdot\nabla V = 0$ on $\mathbb{T}^d$.

The generator of the diffusion X^ϵ_t, denoted by $\mathfrak{L}_\epsilon$, is given by

$$\mathfrak{L}_\epsilon f \;=\; \epsilon\, e^{V/\epsilon}\,\nabla\cdot\big\{ e^{-V/\epsilon}\, a\,\nabla f \big\} \;+\; c\cdot\nabla f \;.$$

Let μ_ϵ be the probability measure given by

$$\mu_\epsilon(dx) \;=\; \frac{1}{Z_\epsilon}\, \exp\{-V(x)/\epsilon\}\, dx \;, \tag{4.2}$$

where Z_ϵ is the normalizing constant, $Z_\epsilon := \int_{\mathbb{T}^d} \exp\{-V(x)/\epsilon\}\, dx$. We have seen in the previous section that μ_ϵ is the stationary state of the process X_t^ϵ provided $\nabla \cdot (e^{-V/\epsilon} c) = 0$. Since $c \cdot \nabla V$ vanishes, this equation becomes $\nabla \cdot c = 0$. We assume therefore that

$$c \cdot \nabla V \;=\; 0 \;\text{ and }\; \nabla \cdot c \;=\; 0 \;\text{ on }\; \mathbb{T}^d\,. \tag{4.3}$$

We examine the transition time in the case where V is a double well potential. Assume that there exists an open set $\mathcal{G} \subset \mathbb{T}^d$ such that

(H1) The potential V has a finite number of critical points in $\mathcal{G}$. Exactly two of them, denoted by $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$, are local minima. The Hessian of V at each of these minima has d strictly positive eigenvalues.

(H2) There is one and only one saddle point between $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ in $\mathcal{G}$, denoted by $\boldsymbol{\sigma}$. The Hessian of V at $\boldsymbol{\sigma}$ has exactly one strictly negative eigenvalue and $(d-1)$ strictly positive eigenvalues.

(H3) We have that $V(\boldsymbol{\sigma}) < \inf_{x\in\partial\mathcal{G}} V(x)$.

Assume without loss of generality that $V(\boldsymbol{m}_2) \le V(\boldsymbol{m}_1)$, so that $\boldsymbol{m}_2$ is the global minimum of the potential V in $\mathcal{G}$. Denote by Ω the level set of the potential defined by saddle point, $\Omega = \{x \in \mathcal{G} : V(x) < V(\boldsymbol{\sigma})\}$. Let $\mathcal{V}_1$, $\mathcal{V}_2$ be two domains with smooth boundary containing $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$, respectively, and contained in Ω:

$$\boldsymbol{m}_i \;\in\; \mathcal{V}_i \;\subset\; \{x \in \mathcal{G} : V(x) < V(\boldsymbol{\sigma}) - \kappa\} \tag{4.4}$$

for some $\kappa > 0$.

Denote by $\nabla^2 V(x)$ the Hessian of V at x. By Lemma 10.1 of Landim and Seo [2016b], both $[\nabla^2 V\, a](\boldsymbol{\sigma})$ and $[\nabla^2 V\, a^\dagger](\boldsymbol{\sigma})$ have a unique (and the same) negative eigenvalue. Denote by $-\mu$ this common negative eigenvalue.

Let $\mathbb{P}_x^\epsilon$, $x \in \mathbb{T}^d$, be the probability measure on $C(\mathbb{R}_+, \mathbb{T}^d)$ induced by the Markov process X_t^ϵ starting from x. Expectation with respect to $\mathbb{P}_x^\epsilon$ is represented by $\mathbb{E}_x^\epsilon$.

Theorem 4.1 (Eyring-Kramers formula)**.** *We have that*

$$\mathbb{E}^\epsilon_{\boldsymbol{m}_1}\,[H_{\mathcal{V}_2}] \;=\; [1 + o_\epsilon(1)]\; \mathfrak{p}\, e^{\Lambda/\epsilon}\;, \quad \text{where } \mathfrak{p} \;=\; \frac{2\pi}{\mu}\, \frac{\sqrt{-\det[\nabla^2 V(\boldsymbol{\sigma})]}}{\sqrt{\det[(\nabla^2 V)(\boldsymbol{m}_1)]}} \tag{4.5}$$

and $\Lambda = V(\boldsymbol{\sigma}) - V(\boldsymbol{m}_1)$.

The term $\mathfrak{p}$ is called the prefactor. It can be understood as the first-order term in the expansion in ϵ of the exponential barrier. Let $\mathbb{E}^\epsilon_{\boldsymbol{m}_1}[H_{\mathcal{V}_2}] = \exp\{\Lambda(\epsilon)/\epsilon\}$. Theorem 4.1 states that $\Lambda(\epsilon) = \Lambda + \epsilon \log \mathfrak{p} + o(\epsilon)$.

The proof of this theorem in the case $c = 0$ and a independent of x, $a(x) = a$, can be found in Landim, Mariani, and Seo [2017]. Uniqueness of local minima and of saddle points connecting the wells is not required there. The same argument should apply to the general case under the hypotheses (2.1), (4.3), but the proof has not been written.

The 0-th order term in the expansion, Λ, can be obtained from Freidlin and Wentzell large deviations theory of random perturbations of dynamical systems Freidlin and Wentzell [1998]. The pre-factor $\mathfrak{p}$ has been calculated rigorously for reversible diffusions by Sugiura [1995, 2001] [based on asymptotics of the principal eigenvalue and eigenfunction for a Dirichlet boundary value problem in a bounded domain], and independently, Bovier, Eckhoff, Gayrard, and Klein [2004] [based on potential theory]. We refer to Berglund [2013] for a recent review.

In the context of chemical reactions, the transition time $\mathbb{E}^{\epsilon}_{\boldsymbol{m}_1}[H_{\mathcal{V}_2}]$ corresponds to the inverse of the rate of a reaction. The so-called "Arrenhuis law" relates the rate of a reaction to the absolute temperature. It seems to have been first discovered empirically by Hood [1878]. van't Hoff [1896] proposed a thermodynamical derivation of the law, and Arrhenius [1889] physical arguments based on molecular dynamics. In the self-adjoint case, the pre-factor $\mathfrak{p}$ first appeared in Eyring [1935] and in more explicit form in Kramers [1940]. Bouchet and Reygner [2016] derived the formula in the non-reversible situation.

5 Metastability

We developed in these last years a robust method to prove the metastable behavior of Markov processes based on potential theory. We report in this section recent developments which rely on asymptotic properties of elliptic operators.

We first define metastability. Let $Z_\epsilon(t)$ be a sequence of Markov processes taking values in some space E_ϵ. Let $\{\mathcal{E}^1_\epsilon, \dots, \mathcal{E}^n_\epsilon, \Delta_\epsilon\}$ be a partition of the set E_ϵ, and set $\mathcal{E}_\epsilon = \mathcal{E}^1_\epsilon \cup \cdots \cup \mathcal{E}^n_\epsilon$.

Fix a sequence of positive numbers θ_ϵ, and denote by $\widehat{Z}_\epsilon(t)$ the process $Z_\epsilon(t)$ speeded-up by θ_ϵ: $\widehat{Z}_\epsilon(t) = Z_\epsilon(t\,\theta_\epsilon)$. Denote by $\widehat{\boldsymbol{P}}_{\epsilon,x}$, $x \in E_\epsilon$, the distribution of the process $\widehat{Z}_\epsilon(t)$ starting from x. Let $S = \{1, \dots, n\}$, $S_0 = \{0\} \cup S$, and let $\Upsilon_\epsilon : E_\epsilon \to S_0$ be the projection given by

$$\Upsilon_\epsilon(x) = \sum_{j=1}^{n} j\, \chi_{\mathcal{E}^j_\epsilon}(x)\,. \tag{5.1}$$

Note that points in Δ_ϵ are mapped to 0. Denote by $z_\epsilon(t)$ the S_0-valued process defined by

$$z_\epsilon(t) = \Upsilon_\epsilon(\widehat{Z}_\epsilon(t)) = \Upsilon_\epsilon(Z_\epsilon(t\theta_\epsilon))\,.$$

The process $z_\epsilon(t)$ is usually not Markovian.

Definition 5.1. [Metastability]. *We say that the process $Z_\epsilon(t)$ is metastable in the time scale θ_ϵ, with metastable sets $\mathcal{E}_\epsilon^1, \dots, \mathcal{E}_\epsilon^n$ if there exists a S-valued, continuous-time Markov chain $z(t)$ such that for all $x \in \mathcal{E}_\epsilon$ the $\widehat{\boldsymbol{P}}_{\epsilon,x}$-finite-dimensional distributions of $z_\epsilon(t)$ converge to the finite-dimensional distributions of $z(t)$.*

The Markov chain $z(t)$ is called the *reduced chain*. Mind that the reduced chain does not take the value 0. The sojourns of $\widehat{Z}_\epsilon(t)$ at Δ_ϵ are washed-out in the limit. Of course, the same process $Z_\epsilon(t)$ may exhibit different metastable behaviors in different time-scales or even different metastable behaviors in the same time-scale but in different regions of the space, inaccessible one to the other in that time-scale.

In some examples Jara, Landim, and Teixeira [2011], Jara, Landim, and Teixeira [2014], and Beltrán, Chavez, and Landim [2017] the set S may be countably infinite. In these cases Υ_ϵ is a projection from E_ϵ to a finite set $S_\epsilon \cup \{0\}$, where S_ϵ increases to a countable set S, and we require $\#E_\epsilon / \#S_\epsilon \to 0$.

In the remaining part of this section we prove that under certain hypotheses the diffusion X_t^ϵ is metastable. Some of these conditions are not needed, but they simplify the presentation. The reader will find in the references finer results.

We assume from now on that the potential V fulfills the following set of assumptions. There exists an open set $\mathcal{G}$ of $\mathbb{T}^d$ such that

(H1') The function V has a finite number of critical points in $\mathcal{G}$. The global minima of V are represented by $m_1, \dots, m_n$. They all belong to $\mathcal{G}$ and they are all at the same height: $V(m_i) = V(m_j)$ for all i, j. The Hessian of V at each of these minima has d strictly positive eigenvalues.

(H2') Denote by $\{\sigma_1, \dots, \sigma_\ell\}$ the set of saddle points between the global minima. Assume that all saddle points are at the same height and that the Hessian of V at these points has exactly one strictly negative eigenvalue and $(d-1)$ strictly positive eigenvalues.

(H3') We have that $V(\sigma_1) < \inf_{x \in \partial \mathcal{G}} V(x)$.

Denote by Ω the level set of the potential defined by the height of the saddle points: $\Omega = \{x \in \mathcal{G} : V(x) < V(\sigma_1)\}$. Let $\mathcal{W}_1, \dots, \mathcal{W}_p$ be the connected components of Ω. Assume that each of these sets contains one and only one global minima, so that $p = n$. Denote by $\mathcal{V}_1, \dots, \mathcal{V}_n$ domains with smooth boundaries satisfying Equation (4.4) for $1 \le i \le n$, and let

$$\mathcal{V} = \bigcup_{j=1}^{n} \mathcal{V}_j \,, \quad \Delta = \mathbb{T}^d \setminus \mathcal{V} \,, \quad \breve{\mathcal{V}}_j = \bigcup_{k : k \neq j} \mathcal{V}_k \,. \tag{5.2}$$

Recall from Equation (4.5) the definition of Λ. Denote by $\widehat{X}_t^\epsilon$ the process X_t^ϵ speeded-up by $\theta_\epsilon = e^{\Lambda/\epsilon}$. This is the diffusion on $\mathbb{T}^d$ whose generator, denoted by $\widehat{\mathfrak{L}}_\epsilon$, is given by

$\widehat{\mathfrak{L}}_\epsilon = \theta_\epsilon \, \mathfrak{L}_\epsilon$. Denote by $\mathbb{P}^\epsilon_x$, resp. $\widehat{\mathbb{P}}^\epsilon_x$, $x \in \mathbb{T}^d$, the probability measure on $C(\mathbb{R}_+, \mathbb{T}^d)$ induced by the diffusion X^ϵ_t, resp. $\widehat{X}^\epsilon_t$, starting from x. Expectation with respect to $\mathbb{P}^\epsilon_x$, is represented by $\mathbb{E}^\epsilon_x$.

Let $S = \{1, \dots, n\}$, $S_0 = \{0\} \cup S$. Denote by $\Upsilon : \mathbb{T}^d \to S_0$ the projection given by Equation (5.1) with $\mathcal{E}^j_\epsilon$ replaced by $\mathcal{V}_j$, and let $x_\epsilon(t)$ be the S_0-valued process defined by

$$x_\epsilon(t) \;=\; \Upsilon(\widehat{X}^\epsilon_t) \;=\; \Upsilon(X^\epsilon(t\theta_\epsilon)) \,.$$

Note that $x_\epsilon(t)$ is not Markovian.

The proof of the metastable behavior of the diffusion X^ϵ_t is divided in four steps. We first show that in the time scale θ_ϵ the process X^ϵ_t spends a negligible amount of time in the set Δ. Then, we derive a candidate for the S-valued Markov chain which is supposed to describe the asymptotic behavior of the process among the wells. In the third step, we prove that the projection of the trace of $\widehat{X}^\epsilon_t$ on $\mathcal{V}$ converges to the S-valued Markov chain introduced in the second step. Finally, we show that the previous results together with an extra condition yield the convergence of the finite-dimensional distributions of $x_\epsilon(t)$.

Step 1: The set Δ is negligible. We first examine in the next lemma the time spent on the set Δ.

Lemma 5.2. *For all $t > 0$,*

$$\lim_{\epsilon\to 0} \sup_{x\in\mathcal{V}} \mathbb{E}^\epsilon_x\Big[\int_0^t \chi_\Delta(X(s\theta_\epsilon))\, ds \Big] \;=\; 0 \,. \tag{5.3}$$

Proof. Here is a sketch of the proof of this result which highlights the relevance of the variational formulae for the capacity. Denote by $\mathrm{cap}_\epsilon(\mathcal{A}, \mathcal{B})$ the capacity between two disjoint subsets $\mathcal{A}$, $\mathcal{B}$ with respect to the diffusion X^ϵ_t.

Fix $1 \le j \le n$ and assume that x belongs to $\mathcal{V}_j$. The time scale θ_ϵ is of the order of the transition time $H_{\breve{\mathcal{V}}_j}$, where the $\breve{\mathcal{V}}_j$ has been introduced in Equation (5.2). The expectation appearing in the statement of the lemma is therefore of the same order of

$$\frac{1}{\theta_\epsilon}\, \mathbb{E}^\epsilon_x\Big[\int_0^{H_{\breve{\mathcal{V}}_j}} \chi_\Delta(X(s))\, ds \Big] \;\sim\; \frac{1}{\theta_\epsilon \mathrm{cap}_\epsilon(\mathcal{V}_j, \breve{\mathcal{V}}_j)} \int_{\mathbb{T}^d} \chi_\Delta \, h_{\mathcal{V}_j, \breve{\mathcal{V}}_j}\, d\mu_\epsilon \;,$$

where last step follows from Proposition 2.6. It would be an identity if we had the harmonic measure in place of the Dirac measure concentrated on x, but these expectations should not be very different because x belongs to the basin of attraction of m_j. Since $\mu_\epsilon(\Delta) \to 0$, the proof is completed if we can show, using the variational principles, that $\theta_\epsilon \mathrm{cap}_\epsilon(\mathcal{V}_j, \breve{\mathcal{V}}_j)$ converges to a positive value. □

Step 2: The reduced chain. The time-scale θ_ϵ at which the process X_t^ϵ evolves among the wells should be of the order of the transition time $\mathbb{E}_{m_j}^\epsilon[\,H(\breve{\mathcal{V}}_j)\,]$. Hence, by Proposition 2.6,

$$\theta_\epsilon \;\sim\; \mathbb{E}_{m_j}^\epsilon\big[\,H(\breve{\mathcal{V}}_j)\,\big] \;\sim\; \frac{1}{\mathrm{cap}_\epsilon(\mathcal{V}_j\,,\,\breve{\mathcal{V}}_j)}\int h_{\mathcal{V}_j,\breve{\mathcal{V}}_j^*}\,d\mu\;.$$

Since m_j is a global minimum of V, the last integral is of order 1 because the harmonic function $h_{\mathcal{V}_j,\breve{\mathcal{V}}_j^*}$ is equal to 1 at $\mathcal{V}_j$. We conclude that the time-scale θ_ϵ should be of the order $\mathrm{cap}_\epsilon(\mathcal{V}_j\,,\,\breve{\mathcal{V}}_j)^{-1}$.

It is proved in Beltrán and Landim [2010, 2012b], in the context of Markov chains taking values in a countable state space, that under certain assumptions

$$\lambda_j \;:=\; \lim_{\epsilon\to 0}\theta_\epsilon\,\frac{1}{\mu_\epsilon(\mathcal{V}_j)}\,\mathrm{cap}_\epsilon(\mathcal{V}_j\,,\,\breve{\mathcal{V}}_j)$$

represents the holding time at j of the reduced chain. Moreover, in the reversible case, the jump rates $r(j,k)$ of the reduced chain are given by

$$r(j,k)=\lim_{\epsilon\to 0}\frac{1}{2\,\mu_\epsilon(\mathcal{V}_j)}\Big\{\mathrm{cap}_\epsilon(\mathcal{V}_j\,,\,\breve{\mathcal{V}}_j)+\mathrm{cap}_\epsilon(\mathcal{V}_k\,,\,\breve{\mathcal{V}}_k)-\mathrm{cap}_\epsilon(\mathcal{V}_j\cup\mathcal{V}_k\,,\,\mathcal{V}\setminus[\mathcal{V}_j\cup\mathcal{V}_k])\Big\}\,.$$

In the non-reversible case, the jump rates are more difficult to derive. By Beltrán and Landim [2012b, Proposition 4.2], still in the context of Markov chains taking values in a countable state space,

$$r(j,k) \;=\; \lambda_j\,\lim_{\epsilon\to 0}\overline{\mathbb{P}}_{m_j}^\epsilon\big[\,H(\mathcal{V}_k) < H\big(\mathcal{V}\setminus[\mathcal{V}_j\cup\mathcal{V}_k]\big)\,\big]\,,$$

where $\overline{\mathbb{P}}_{m_j}^\epsilon$ represents the distribution of the process in which the well $\mathcal{V}_j$ has been collapsed to the point m_j. Estimates on the harmonic function appearing on the right-hand of this equation are obtained by showing that this function solves a variational problem, similar to the one for the capacity, and then that to be optimal, a function has to take a precise value at the set $\mathcal{V}_j$. We refer to Landim [2014] and Landim and Seo [2016b] for details, where this program has been successfully undertaken for two different models.

Assume that one can compute the asymptotic jump rates through the previous formulae or that one can guess by other means the jump rates of the reduced chain. Denote by $\boldsymbol{L}$ the generator of the S-valued continuous-time Markov chain induced by these jump rates. Let $D(\mathbb{R}_+,E)$, E a metric space, be the space of E-valued, right-continuous functions with left-limits endowed with the Skorohod topology, and let $\boldsymbol{Q}_j$, $j\in S$, the measure on $D(\mathbb{R}_+,S)$ induced by the Markov chain with generator $\boldsymbol{L}$ starting from j.

Step 3: Convergence of the trace. We turn to the convergence of the trace process. Recall that $\widehat{X}_t^\epsilon$ represents the process X_t^ϵ speeded-up by θ_ϵ. Denote by $T_{\mathcal{V}}(t)$, $t\geq 0$, the total time spent by the diffusion $\widehat{X}^\epsilon$ on the set $\mathcal{V}$ in the time interval $[0,t]$:

$$T_{\mathcal{V}}(t) \;:=\; \int_0^t \chi_{\mathcal{V}}(\widehat{X}^\epsilon_s)\, ds \;,$$

Denote by $\{S_{\mathcal{V}}(t) : t \geq 0\}$ the generalized inverse of $T_{\mathcal{V}}(t)$:

$$S_{\mathcal{V}}(t) \;:=\; \sup\{s \geq 0 : T_{\mathcal{V}}(s) \leq t\}\,.$$

Clearly, for all $r \geq 0$, $t \geq 0$,

$$\{S_{\mathcal{V}}(r) \geq t\} \;=\; \{T_{\mathcal{V}}(t) \leq r\}\;. \tag{5.4}$$

It is also clear that for any starting point $x \in \mathbb{T}^d$, $\lim_{t\to\infty} T_{\mathcal{V}}(t) = \infty$ almost surely. Therefore, the random path $\{Y_\epsilon(t) : t \geq 0\}$, given by $Y_\epsilon(t) := \widehat{X}^\epsilon(S_{\mathcal{V}}(t))$, is well defined for all $t \geq 0$ and takes value in the set $\mathcal{V}$. We call the process $Y_\epsilon(t)$ the *trace* of $\widehat{X}^\epsilon_t$ on the set $\mathcal{V}$.

The process $Y_\epsilon(t)$ is Markovian provided the initial filtration is large enough. Indeed, denote by $\{\mathcal{F}_t^0 : t \geq 0\}$ the natural filtration of $C(\mathbb{R}_+, \mathbb{T}^d)$: $\mathcal{F}_t^0 = \sigma(X_s : 0 \leq s \leq t)$. Fix $x_0 \in \mathcal{V}$ and denote by $\{\mathcal{F}_t : t \geq 0\}$ the usual augmentation of $\{\mathcal{F}_t^0 : t \geq 0\}$ with respect to $\mathbb{P}^\epsilon_{x_0}$. We refer to Section III.9 of Rogers and Williams [2000] for a precise definition, and to Landim and Seo [2016a] for a proof of the next result which relies on the identity (5.4).

Lemma 5.3. *For each $t \geq 0$, $S_{\mathcal{V}}(t)$ is a stopping time with respect to the filtration $\{\mathcal{F}_t\}$.*

As $S_{\mathcal{V}}(t)$ is a stopping time with respect to the filtration $\{\mathcal{F}_t\}$, $Y_\epsilon(t)$ is a $\mathcal{V}$-valued, Markov process with respect to the filtration $\mathcal{G}_t := \mathcal{F}_{S(t)}$. Let $\Psi : \mathcal{V} \to S$ be the projection given by

$$\Psi(x) \;=\; \sum_{j=1}^n j\, \chi_{\mathcal{V}_j}(x)\;,$$

and denote by $y_\epsilon(t)$ the S-valued process obtained by projecting $Y_\epsilon(t)$ with Ψ:

$$y_\epsilon(t) \;=\; \Psi(Y_\epsilon(t))\;.$$

Note that the process $y_\epsilon(t)$ is not Markovian.

Denote by $\mathbb{Q}^\epsilon_x$, resp. $\mathbf{Q}^\epsilon_x$, $x \in \mathcal{V}$, the probability measure on $D(\mathbb{R}_+, \mathcal{V})$, resp. $D(\mathbb{R}_+, S)$, induced by the process $Y_\epsilon(t)$, resp. $y_\epsilon(t)$, given that $Y_\epsilon(0) = x$. Fix $j \in S$, $x \in \mathcal{V}_j$. As usual, the proof that $\mathbf{Q}^\epsilon_x$ converges to $\mathbf{Q}_j$ is divided in two steps. We first show that the sequence $\mathbf{Q}^\epsilon_x$ is tight and then we prove the uniqueness of limit points.

Lemma 5.4. *Assume that conditions (5.3) is in force. Suppose, furthermore, that*

$$\lim_{r\to 0}\limsup_{\epsilon\to 0}\max_{1\le j\le n}\sup_{x\in\mathcal{V}_j}\mathbb{P}^\epsilon_x\big[\,H(\breve{\mathcal{V}}_j)\,\le\, r\,\theta_\epsilon\,\big]\;=\;0\,. \tag{5.5}$$

Then, for every $1\le j\le n$, $x_0\in\mathcal{V}_j$, the sequence of measures $\mathbf{Q}^\epsilon_{x_0}$ is tight. Moreover, every limit point $\mathbf{Q}^*$ of the sequence $\mathbf{Q}^\epsilon_{x_0}$ is such that

$$\mathbf{Q}^*\{x : x(0)=j\}\;=\;1\quad\textit{and}\quad\mathbf{Q}^*\{x : x(t)\neq x(t-)\}\;=\;0 \tag{5.6}$$

for every $t>0$.

A proof of this result for one-dimensional diffusions is presented in Landim and Seo [2016a, Lemma 7.5]. Condition (5.5) asserts that in the time-scale θ_ϵ, the process X^ϵ_t may not jump instantaneously from one well to the other. We show in Section 8 of this article that the probability $\mathbb{P}^\epsilon_x[\,H(\breve{\mathcal{V}}_j)\,\le\,r\,\theta_\epsilon\,]$ is bounded by the capacity between two sets for an enlarged process. The proof of this lemma is thus reduced to an estimate of capacities.

The proof of uniqueness relies on the characterization of continuous-time Markov chains as solutions of martingale problems. One needs to show that

$$\boldsymbol{F}(y(t))\;-\;\int_0^t(\boldsymbol{L}\boldsymbol{F})(y(s))\,ds \tag{5.7}$$

is a martingale under $\mathbf{Q}$ for all functions $\boldsymbol{F}:S\to\mathbb{R}$ and all limit point $\mathbf{Q}$ of the sequence $\mathbf{Q}^\epsilon_{x_0}$.

We proved in Beltrán and Landim [2010, 2012b] that this property is in force in the context of countable state spaces provided the mean jump rates converge and if each well $\mathcal{V}_j$ has an element z_j such that

$$\lim_{\epsilon\to 0}\sup_{y\in\mathcal{V}_j,y\neq z_j}\frac{\mathrm{cap}_\epsilon(\mathcal{V}_j,\breve{\mathcal{V}}_j)}{\mathrm{cap}_\epsilon(\{y\},\{z_j\})}\;=\;0\,. \tag{5.8}$$

The point z_j is not special. Typically, if Equation (5.8) holds for a point z_j in the well, it holds for all the other ones. We refer to Beltrán and Landim [2010, 2012b] for details.

Condition (5.8) has been derived for Markov processes which "visit points", that is, for Markov processes which visit all points of a well before reaching another well. This is the case of condensing zero-range processes Beltrán and Landim [2012a] and Landim [2014], random walks in potential fields Landim, Misturini, and Tsunoda [2015] and Landim and Seo [2016b], one-dimensional diffusions Landim and Seo [2016a], and for all processes whose wells are reduced to singletons, as the inclusion processes Bianchi, Dommers, and Giardinà [2017].

We present here an alternative method to deduce Equation (5.7) which relies on certain asymptotic properties of the elliptic operator $\mathscr{L}_\epsilon$. Fix a function $\boldsymbol{F} : S \to \mathbb{R}$, let $\boldsymbol{G} = \boldsymbol{L}\boldsymbol{F}$, and let $g : \mathbb{T}^d \to \mathbb{R}$ be given by

$$g \;=\; \sum_{i=1}^{n} \boldsymbol{G}(i)\, \chi_{\mathcal{V}_i} \;.$$

Assume that there exists a sequence of function $g_\epsilon : \mathbb{T}^d \to \mathbb{R}$ such that

(P1) g_ϵ vanishes on $\mathcal{V}^c$ and converges to g uniformly on $\mathcal{V}$;

(P2) The Poisson equation $\widehat{\mathfrak{L}}_\epsilon f = g_\epsilon$ in $\mathbb{T}^d$ has a solution denoted by f_ϵ. Moreover, there exists a finite constant C_0 such that

$$\sup_{0<\epsilon<1} \sup_{x\in\mathbb{T}^d} |f_\epsilon(x)| \;\le\; C_0 \;, \quad \text{and} \quad \lim_{\epsilon\to 0} \sup_{x\in\mathcal{V}} \big| f_\epsilon(x) - f(x) \big| \;=\; 0 \;,$$

where $f : \mathbb{T}^d \to \mathbb{R}$ is given by $f = \sum_{1\le i\le n} \boldsymbol{F}(i)\, \chi_{\mathcal{V}_i}$.

The natural candidate for g_ϵ in conditions (P1) and (P2) is the function g itself. However, as the process is ergodic, the Poisson equation $\widehat{\mathfrak{L}}_\epsilon f = b$ has a solution only if b has mean zero with respect to μ_ϵ. We need therefore to modify g to obtain a mean-zero function. Denote by π the stationary state of the Markov chain whose generator is $\boldsymbol{L}$. We expect $\mu_\epsilon(\mathcal{V}_i)$ to converge to π_i. Hence,

$$\lim_{\epsilon\to 0} E_{\mu_\epsilon}[g] \;=\; \lim_{\epsilon\to 0} \sum_{i=1}^{n} \boldsymbol{G}(i)\, \mu_\epsilon(\mathcal{V}_i) \;=\; \sum_{i=1}^{n} \boldsymbol{L}\boldsymbol{F}(i)\, \pi_i \;=\; 0 \,.$$

A reasonable candidate for g_ϵ is thus $g - r(\epsilon)\, \chi_{\mathcal{V}_1}$, where $r(\epsilon) = E_{\mu_\epsilon}[g]/\mu_\epsilon(\mathcal{V}_1)$.

Properties (P1), (P2) have been proved in Evans and Tabrizian [2016] and Seo and Tabrizian [2017] for elliptic operators on $\mathbb{R}^d$ of the form $\mathfrak{L}_\epsilon f = e^{V/\epsilon}\nabla\cdot(e^{-V/\epsilon} a \nabla f)$ and in Landim and Seo [2016a] for one-dimensional diffusions with periodic boundary conditions. It is an open problem to prove these conditions in the context of interacting particle systems.

Lemma 5.5. *Fix $1 \le j \le n$ and $x_0 \in \mathcal{V}_j$. Assume that conditions (P1) and (P2) are in force. Let $\mathbf{Q}^*$ be a limit point of the sequence $\mathbf{Q}^\epsilon_{x_0}$ satisfying Equation (5.6). Then, for every $\boldsymbol{F} : S \to \mathbb{R}$, Equation (5.7) is a martingale under the measure $\mathbf{Q}^*$.*

Proof. Fix $1 \le j \le n$, $x_0 \in \mathcal{V}_j$ and a function $\boldsymbol{F} : S \to \mathbb{R}$. Let $f_\epsilon : \mathbb{T}^d \to \mathbb{R}$ be the function given by assumption (P2). Then,

$$M_\epsilon(t) \;=\; f_\epsilon(\widehat{X}^\epsilon_t) \;-\; \int_0^t (\widehat{\mathfrak{L}}_\epsilon f_\epsilon)(\widehat{X}^\epsilon_s)\, ds \;=\; f_\epsilon(\widehat{X}^\epsilon_t) \;-\; \int_0^t g_\epsilon(\widehat{X}^\epsilon_s)\, ds$$

is a martingale with respect to the filtration $\mathscr{F}_t$ and the measure $\widehat{\mathbb{P}}^{\epsilon}_{x_0}$. Since $\{S_{\mathcal{V}}(t) : t \ge 0\}$ are stopping times with respect to $\mathscr{F}_t$,

$$\widehat{M}_\epsilon(t) \;=\; M_\epsilon(S_{\mathcal{V}}(t)) \;=\; f_\epsilon(Y_\epsilon(t)) \;-\; \int_0^{S_{\mathcal{V}}(t)} g_\epsilon(\widehat{X}^\epsilon_s)\, ds$$

is a martingale with respect to the filtration $\mathscr{G}_t$. Since g_ϵ vanishes on $\mathcal{V}^c$, by a change of variables,

$$\int_0^{S_{\mathcal{V}}(t)} g_\epsilon(\widehat{X}^\epsilon_s)\, ds \;=\; \int_0^{S_{\mathcal{V}}(t)} g_\epsilon(\widehat{X}^\epsilon_s)\, \chi_{\mathcal{V}}(\widehat{X}^\epsilon_s)\, ds \;=\; \int_0^t g_\epsilon(\widehat{X}^\epsilon(S_{\mathcal{V}}(s)))\, ds\;.$$

Hence,

$$\widehat{M}_\epsilon(t) \;=\; f_\epsilon(Y_\epsilon(t)) \;-\; \int_0^t g_\epsilon(Y_\epsilon(s))\, ds$$

is a $\{\mathscr{G}_t\}$-martingale under the measure $\mathbb{Q}^\epsilon_{x_0}$.

By (P1) and (P2), g_ϵ, resp. f_ϵ, converge to g, resp. f, uniformly in $\mathcal{V}$ as $\epsilon \to 0$. Hence, since $Y_\epsilon(s) \in \mathcal{V}$ for all $s \ge 0$, we may replace in the previous equation g_ϵ, f_ϵ by g, f, respectively, at a cost which vanishes as $\epsilon \to 0$. Therefore,

$$\widehat{M}_\epsilon(t) \;=\; f(Y_\epsilon(t)) \;-\; \int_0^t g(Y_\epsilon(s))\, ds \;+\; o(1)$$

is a $\{\mathscr{G}_t\}$-martingale under the measure $\mathbb{Q}^\epsilon_{x_0}$.

Since f and g are constant on each set $\mathcal{V}_i$, $f(Y_\epsilon(t)) = \boldsymbol{F}(y_\epsilon(t))$, $g(Y_\epsilon(t)) = \boldsymbol{G}(y_\epsilon(t))$. By the second condition in Equation (5.6), $\mathbf{Q}^*$ is concentrated on trajectories which are continuous at any fixed time with probability 1. We may, therefore, pass to the limit and conclude that $\boldsymbol{F}(y(t)) - \int_0^t (\boldsymbol{L}\boldsymbol{F})(y(s))\, ds$ is a martingale under $\mathbf{Q}^*$. □

Theorem 5.6. *Assume that conditions (P1), (P2), Equation (5.3), Equation (5.5) are in force. Fix $j \in S$ and $x_0 \in \mathcal{V}_j$. The sequence of measures $\mathbf{Q}^\epsilon_{x_0}$ converges, as $\epsilon \to 0$, to the probability measure $\mathbf{Q}_j$.*

Proof. The assertion is a consequence of Lemma 5.4, Lemma 5.5 and the fact that there is only one measure $\mathbf{Q}$ on $D(\mathbb{R}_+, S)$ such that $\mathbf{Q}[x(0) = j] = 1$ and such that Equation (5.7) is a martingale for all $\boldsymbol{F} : S \to \mathbb{R}$. □

Step 4: The finite-dimensional distributions. By Landim, Loulakis, and Mourragui [2017, Proposition 1.1], the finite-dimensional distributions of $x_\epsilon(t)$ converge to the finite-dimensional distributions of $y(t)$ if the process $y_\epsilon(t)$ converges in the Skorohod topology to $y(t)$ [Theorem 5.6], if in the time-scale θ_ϵ the total time spent in Δ is negligible [Lemma 5.2] and if

$$\lim_{\delta\to 0} \limsup_{\epsilon \to 0} \sup_{x\in\mathcal{V}} \sup_{\delta\le s\le 2\delta} \mathbb{P}^\epsilon_x[X^\epsilon(s\theta_\epsilon) \in \Delta] \;=\; 0\;.$$

This completes the argument. The convergence of the finite-dimensional distributions of X_t^ϵ and sharp asymptotics for the transition time in the context of diffusions were first obtained by Sugiura [1995, 2001]. The approach presented in this section to prove the metastable behavior of a Markov process has been proposed by Beltrán and Landim [2010, 2012b]. It has been successfully applied to many models quoted in this section. For further reading on metastability, we refer to the books by Olivieri and Vares [2005] and by Bovier and den Hollander [2015].

Acknowledgments. The results presented in this review are the outcome of long standing collaborations. The author wishes to thank his colleagues and friends, J. Beltrán, E. Chavez, A. Gaudillière, M. Jara, M. Loulakis, M. Mariani, R. Misturini, M. Mourragui, I. Seo, A. Teixeira, K. Tsunoda.

References

Svante Arrhenius (1889). "On the reaction velocity of the inversion of cane sugar by acids". *Zeitschrift fur physikalische Chemie* 4. Partially translated to english in "Selected readings in chemical kinetics" edited by M. Back and K. Laidler, Pergamon press, Oxford 1967., 226ff (cit. on p. 2651).

J. Beltrán, E. Chavez, and C. Landim (2017). *On the reaction velocity of the inversion of cane sugar by acids*. Preprint (cit. on p. 2652).

J. Beltrán and C. Landim (2010). "Tunneling and metastability of continuous time Markov chains". *J. Stat. Phys.* 140.6, pp. 1065–1114. MR: 2684500 (cit. on pp. 2654, 2656, 2659).

– (2012a). "Metastability of reversible condensed zero range processes on a finite set". *Probab. Theory Related Fields* 152.3-4, pp. 781–807. MR: 2892962 (cit. on p. 2656).

– (2012b). "Tunneling and metastability of continuous time Markov chains II, the nonreversible case". *J. Stat. Phys.* 149.4, pp. 598–618. MR: 2998592 (cit. on pp. 2647, 2654, 2656, 2659).

N. Berglund (2013). "Kramers' law: validity, derivations and generalisations". *Markov Process. Related Fields* 19.3, pp. 459–490. MR: 3156961 (cit. on p. 2651).

Alessandra Bianchi, Sander Dommers, and Cristian Giardinà (2017). "Metastability in the reversible inclusion process". *Electron. J. Probab.* 22, Paper No. 70, 34. MR: 3698739 (cit. on p. 2656).

Freddy Bouchet and Julien Reygner (2016). "Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes". *Ann. Henri Poincaré* 17.12, pp. 3499–3532. arXiv: 1507.02104. MR: 3568024 (cit. on p. 2651).

Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein (2001). "Metastability in stochastic dynamics of disordered mean-field models". *Probab. Theory Related Fields* 119.1, pp. 99–161. MR: 1813041 (cit. on p. 2647).

– (2004). "Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times". *J. Eur. Math. Soc. (JEMS)* 6.4, pp. 399–424. MR: 2094397 (cit. on p. 2651).

Anton Bovier and Frank den Hollander (2015). *Metastability*. Vol. 351. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. A potential-theoretic approach. Springer, Cham, pp. xxi+581. MR: 3445787 (cit. on p. 2659).

P Doyle (1994). "Energy for Markov chains". Unpublished manuscript (cit. on p. 2647).

Lawrence C. Evans and Peyam R. Tabrizian (2016). "Asymptotics for scaled Kramers-Smoluchowski equations". *SIAM J. Math. Anal.* 48.4, pp. 2944–2961. MR: 3542005 (cit. on p. 2657).

Henry Eyring (1935). "The activated complex in chemical reactions". *The Journal of Chemical Physics* 3.2, pp. 107–115 (cit. on p. 2651).

M. I. Freidlin and A. D. Wentzell (1998). *Random perturbations of dynamical systems*. Second. Vol. 260. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the 1979 Russian original by Joseph Szücs. Springer-Verlag, New York, pp. xii+430. MR: 1652127 (cit. on p. 2651).

A. Gaudillière and C. Landim (2014). "A Dirichlet principle for non reversible Markov chains and some recurrence theorems". *Probab. Theory Related Fields* 158.1-2, pp. 55–89. MR: 3152780 (cit. on pp. 2645, 2647).

David Gilbarg and Neil S. Trudinger (1977). *Elliptic partial differential equations of second order*. Grundlehren der Mathematischen Wissenschaften, Vol. 224. Springer-Verlag, Berlin-New York, pp. x+401. MR: 0473443 (cit. on p. 2640).

Jacobus Henricus van't Hoff (1896). *Studies in chemical dynamics*. Chemical publishing Company (cit. on p. 2651).

John J Hood (1878). "On the laws of chemical change – Part I". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 6.38, pp. 371–383 (cit. on p. 2651).

M. Jara, C. Landim, and Augusto Teixeira (2014). "Universality of trap models in the ergodic time scale". *Ann. Probab.* 42.6, pp. 2497–2557. MR: 3265173 (cit. on p. 2652).

Milton Jara, Claudio Landim, and Augusto Teixeira (2011). "Quenched scaling limits of trap models". *Ann. Probab.* 39.1, pp. 176–223. MR: 2778800 (cit. on p. 2652).

Oliver Dimon Kellogg (1967). *Foundations of potential theory*. Reprint from the first edition of 1929. Die Grundlehren der Mathematischen Wissenschaften, Band 31. Springer-Verlag, Berlin-New York, pp. ix+384. MR: 0222317 (cit. on pp. 2635, 2637).

H. A. Kramers (1940). “Brownian motion in a field of force and the diffusion model of chemical reactions”. *Physica* 7, pp. 284–304. MR: 0002962 (cit. on p. 2651).

C. Landim (2014). “Metastability for a non-reversible dynamics: the evolution of the condensate in totally asymmetric zero range processes”. *Comm. Math. Phys.* 330.1, pp. 1–32. MR: 3215575 (cit. on pp. 2654, 2656).

C. Landim, M. Loulakis, and M. Mourragui (2017). “Metastable Markov chains”. arXiv: 1703.09481 (cit. on p. 2658).

C. Landim, M. Mariani, and I. Seo (2017). “A Dirichlet and a Thomson principle for non-selfadjoint elliptic operators, Metastability in non-reversible diffusion processes”. arXiv: 1701.00985 (cit. on pp. 2647, 2651).

C. Landim, R. Misturini, and K. Tsunoda (2015). “Metastability of reversible random walks in potential fields”. *J. Stat. Phys.* 160.6, pp. 1449–1482. MR: 3382755 (cit. on p. 2656).

C. Landim and I. Seo (2016a). “Metastability of non-reversible, mean-field Potts model with three spins”. *J. Stat. Phys.* 165.4, pp. 693–726. arXiv: 1710.06672. MR: 3568163 (cit. on pp. 2655–2657).

Claudio Landim and Insuk Seo (2016b). “Metastability of non-reversible random walks in a potential field, the Eyring-Kramers transition rate formula”. To appear in Commun. Pure Appl. Math. arXiv: 1605.01009 (cit. on pp. 2650, 2654, 2656).

Enzo Olivieri and Maria Eulália Vares (2005). *Large deviations and metastability*. Vol. 100. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, pp. xvi+512. MR: 2123364 (cit. on p. 2659).

Ross G. Pinsky (1988a). “A generalized Dirichlet principle for second order nonselfadjoint elliptic operators”. *SIAM J. Math. Anal.* 19.1, pp. 204–213. MR: 924555 (cit. on p. 2647).

– (1988b). “A minimax variational formula giving necessary and sufficient conditions for recurrence or transience of multidimensional diffusion processes”. *Ann. Probab.* 16.2, pp. 662–671. MR: 929069 (cit. on p. 2647).

– (1995). *Positive harmonic functions and diffusion*. Vol. 45. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, pp. xvi+474. MR: 1326606 (cit. on p. 2647).

L. C. G. Rogers and David Williams (2000). *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Itô calculus, Reprint of the second (1994) edition. Cambridge University Press, Cambridge, pp. xiv+480. MR: 1780932 (cit. on p. 2655).

I Seo and Peyam R. Tabrizian (2017). “Asymptotics for scaled Kramers-Smoluchowski equations in several dimensions with general potentials”. Preprint (cit. on p. 2657).

M. Slowik (2012). “A note on variational representations of capacities for reversible and nonreversible Markov chains”. Unpublished, Technische Universität Berlin (cit. on p. 2647).

Makoto Sugiura (1995). “Metastable behaviors of diffusion processes with small parameter”. *J. Math. Soc. Japan* 47.4, pp. 755–788. MR: 1348758 (cit. on pp. 2651, 2659).

– (2001). “Asymptotic behaviors on the small parameter exit problems and the singularly perturbation problems”. *Ryukyu Math. J.* 14, pp. 79–118. MR: 1897280 (cit. on pp. 2651, 2659).

Received 2017-12-06.

Claudio Landim
IMPA
Estrada Dona Castorina 110
22460 Rio de Janeiro
Brasil

and

CNRS UMR 6085
Université de Rouen
Avenue de l’Université, BP.12
Technopôle du Madrillet
F76801 Saint-Étienne-du-Rouvray
France
landim@impa.br

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2661–2686)

TRANSPORT IN PARTIALLY HYPERBOLIC FAST-SLOW SYSTEMS

Carlangelo Liverani

Abstract

I will discuss, from a dynamical systems point of view, some recent attempts to rigorously derive the macroscopic laws of transport (e.g. the heat equation) from deterministic microscopic dynamics.

1 The problem

In physics the world is described at different scales by seemingly very different laws. Once the laws are specified, the problem of explaining their compatibility in spite of their apparent differences becomes a mathematical one. Of course, it is possible to give heuristic explanations, and plenty of them are available. Nevertheless, it turns out that the issue is always very subtle, so that non rigorous explanations are often faulty and our naive intuition is at loss.

In addition, in many instances one is interested in the behaviour of the world in a middle ground, that is at intermediate scales, and to make accurate predictions in such a realm a well grounded theory of how one scale merges in the next is necessary. An example of this kind is given by the current development of nanotechnology in which the systems of interest are mesoscopic: too large to apply easily the microscopic laws and too small for the macroscopic laws to be valid without qualification.

Here we will consider the transition between the macroscopic scale (the one we are used to) of the order of a meter and the microscopic (atomic) scale of the order of at most 10^{-9} meters.

Much of the work here described has been supported by the ERC grant Macroscopic Laws and Dynamical Systems (MALADY) (ERC AdG 246953). The author acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006. The author would finally like to thank Jean-Pierre Eckmann and Stefano Olla for pointing out several errors in a first version of this manuscript.

MSC2010: primary 82C05; secondary 37A25, 37C30, 37D30, 37A50, 82C70.

Keywords: hydrodynamic limit, Central Limit theorem, fas-slow systems, partially hyperbolic systems.

While, ultimately, the cross over from the microscopic to the macroscopic must entail an understanding of the measurement process and of the semiclassical limit of Quantum Mechanics, many issues can be treated also remaining in the purely classical realm. One outstanding conceptual issue, going back to the ancient dispute between Zeno and Democritus 2400 yeas ago, stems from the fact that the world around us looks like a continuum (we describe it using partial differential equations); yet we are aware that it consists of atoms, hence it is discrete in nature (as first conclusively proven by Einstein [1905] who showed how Brownian motion, a mesoscopic phenomena, emerges from the microscopic dynamics).

It should therefore be possible to start from an atomic description and derive, in some appropriate limit that accounts for the difference in scales, a continuous description. In order to carry out such a program the first task is to identify the microscopic quantities that can be recognised and described macroscopically. It turns out that the microscopic dynamics has some quantities that are locally conserved (e.g. mass, energy, momentum, ...) and these tend to evolve much slower than the other degrees of freedom, hence allowing an evolution visible in the macroscopic time scale. In this article I will mainly discuss the situation in which the microscopic dynamics is Hamiltonian and classical and the conserved quantity is the energy. Hence the goal is to describe the energy evolution (energy transport) on the macroscopic scale.

Starting with the work of Boltzmann we understand that the microscopic energy manifests itself as thermal heat. The typical macroscopic law for heat transport is the Fourier law (although important violations of such a law, connected to specific microscopic properties, have been discovered, e.g. Carbonium nanotubes).

For simplicity, we consider heat transport in homogenous non-conducting solids. This implies that we have a constant density and there is no mass flow: the only quantity that evolves is heat. The macroscopic definition of heat is the amount of energy that is needed to change the temperature of a body and it is proportional to the temperature via the *specific heat per unit volume* $c_v(T)$. The Fourier law states that the heat flux J satisfies

$$J = -\hat{\kappa}\nabla T$$

where $T(x,t)$ is the temperature at point x and time t, and $\hat{\kappa}$ is the *heat conductivity* of the material. The assumption that heat is a locally conserved quantity is tantamount to saying that it satisfies a continuity equation:

$$c_v(T)\partial_t T = -\mathrm{div} J = \mathrm{div}(\hat{\kappa}\nabla T).$$

In other words, setting $\kappa(T) = \hat{\kappa}/c_v(T)$, often called *diffusivity*, we have

$$\partial_t T = \kappa(T)\Delta T \tag{1-1}$$

which is nothing else that the heat equation.

The mathematical problem mentioned above reads, in the present contest,

- *Derive rigorously the heat equation from a Hamiltonian dynamics.*

Such a problem is extremely hard, as is explained at length in the review Bonetto, J. L. Lebowitz, and Rey-Bellet [2000]. Such a review is 17 years old, yet it is still actual since, in spite of considerable efforts, little progress has been achieved in the intervening years. Yet, little progress is not zero progress. In the following I will describe some of the mathematical work made in the last years. I start by making precise the kind of systems I want to consider.

2 The models: general considerations

We describe here a very idealised model of a homogeneous, non-conducting solid. In other words, the solid will look more or less the same at all places and the particles that constitute it are not free to move very far from their rest position. This is the simplest possible situation one can think of since the only quantity that can move around is the energy.

The main feature of the problem is that there is a microscopic versus a macroscopic world (and description). As we will restrict ourselves to the classical world (that is, we are ignoring quantum effects) both world can be described by differential equations (either ordinary, stochastic or partial) on $\mathbb{R}^n$, for some $n \in \mathbb{N}$. In fact, the macroscopic world and the microscopic one differ just in the scale. The distinction can be encapsulated in the fact that a scaling parameter L becomes extremely large.

More precisely, since we are going to consider only models of solids, we can restrict our discussion to microscopic models defined in a region $\Lambda_L = \{x \in \mathbb{Z}^d \mid L^{-1}x \in \Lambda\}$ for some nice, fixed, region $\Lambda \subset \mathbb{R}^d$.[1] Accordingly, the region Λ stands for the macroscopic solid we want to consider, while Λ_L is the corresponding region in the microscopic description. Note that Λ_L is discrete in nature as it is a subset of a lattice and it has size of order L when measured in microscopic units.[2] Note that the microscopic unitis are such that the discrete nature of the system is evident at distances of order one, rather than at distance L^{-1}, as it happens in macroscopic units.

At each point $x \in \Lambda_L$ we assume that there is a group of particles (atoms, molecules, defects, ...) that are described by coordinates $q_x, p_x \in M_*$,[3] where M_* is the cotangent bundle of an n_* dimensional compact Riemannian manifold M. One can think of the q_x

[1] Of course we are mainly interested in the cases $d = 1, 2, 3$ (wires, membranes and solids).
[2] The choice of a square lattice is immaterial, any other will do.
[3] The fact that M_* does not depend on x is part of our homogeneity hypothesis.

as the displacement of particles from their equilibrium positions while the p_x are their momenta (or velocities). The fact that q_x belongs to a compact manifold is our non conduction hypothesis: particles are not free to move around the solid. That is, there is no *convection*.[4] Accordingly, the microscopic system is described in a phase space $\mathfrak{M}_L = M_*^{\Lambda_L}$.

On the contrary the macroscopic description consists simply of a temperature field $T(x,t)$, $x \in \Lambda$, $t \in \mathbb{R}$. Since the temperature is a function of the internal energy density u of the body, the heat equation (1-1) can be written as

$$\partial_t u = \operatorname{div}(\kappa(u)\nabla u). \tag{2-1}$$

Equation (2-1) describes the macroscopic dynamics, it remains to describe the microscopic dynamics. Ideally we want a Hamiltonian dynamics, yet it is instructive to allow also stochastic dynamics, since their study is easier and it might provide important ideas to understand the deterministic case. Nevertheless, we will assume that the dynamics is Markov. In other words, given any initial probability distribution $\mathbb{P}_0$, the distribution $\mathbb{P}_t$ at time t is given by, for all $f \in \mathcal{C}^0(\mathfrak{M}, \mathbb{R})$,

$$\mathbb{E}_t(f) = \mathbb{E}_0(\mathcal{L}_t f) \tag{2-2}$$

where $\mathbb{E}_t$, $t \geq 0$, is the expectation with respect to the probability measure $\mathbb{P}_t$ and $\mathcal{L}_t : \mathcal{C}^0 \to \mathcal{C}^0$ is a strongly continuous one parameter semigroup. Of course, in the case of a Hamiltonian dynamics, calling ϕ_t the Hamiltonian flow, we will have $\mathcal{L}_t f = f \circ \phi_t$.

To specify the dynamics we need to discuss also the initial conditions. While the initial conditions of equation (2-1) are simply an initial energy profile $u_0(x)$, the initial conditions for the microscopic model are a much more subtle issue, especially in the case of a Hamiltonian dynamics. Indeed, Hamiltonian dynamics are reversible: for each trajectory there exists an initial condition for which the trajectory is run backward (just take the final configuration of the trajectory and reverse all the velocities). On the contrary the heat equation (2-1) has no such property. So initial conditions must play an important role. To address this problem we must be more specific about the type of dynamics we consider, so we postpone the discussion momentarily (see (4-1) for details). For the time being we require the bare minimum. First of all we consider random initial conditions, that is the initial conditions are described by a *non atomic* measure $\mathbb{P}_0$. This corresponds to the natural fact that the exact positions and velocities of the microscopic particles (that in real life applications may be of order 10^{23}) cannot be known precisely and only a statistical knowledge is possible.

Also, to connect to the macroscopic setting, we must say what we mean for the internal energy $\boldsymbol{u}_x$ of the body at site x. This can be done in several ways that are all essentially

[4] Of course, one can achieve the same when $M_* = \mathbb{R}^{2n_*}$ by introducing some strongly confining potential, see Section 4.1, but for now let us keep things as simple as possible.

equivalent, however they depend on the form of the dynamics which we have not yet described precisely; so we postpone the definition to equation (3-3). However, whatever the exact definition, we are interested in the measures $\mu_{\boldsymbol{u},L}$ defined as, for all $\varphi \in \mathcal{C}^\infty$, [5]

$$\mu_{\boldsymbol{u},L}(\varphi) := \frac{1}{L^d} \sum_{x \in \Lambda_L} \varphi(L^{-1}x)\boldsymbol{u}_x(0) = \frac{1}{L^d} \sum_{x \in \Lambda_L} \boldsymbol{u}_x(0)\delta_{L^{-1}x}(\varphi).$$

Note that $\mu_{\boldsymbol{u},L}$ depends from the microscopic configurations and hence is a random variable under $\mathbb{P}_0$. We then ask that there exists a smooth function u such that

$$\lim_{L\to\infty} \frac{1}{L^d} \sum_{x \in \Lambda_L} \boldsymbol{u}_x(0)\delta_{L^{-1}x}(\varphi) = \int_{\mathbb{R}^d} u(x)\varphi(x)dx$$

where the limits is meant almost surely with respect to $\mathbb{P}_0$. Thus, at time zero the microscopic energy gives rise to a nice energy profile on the macroscopic scale.

We are finally able to specify in which sense the macroscopic dynamics should arise from the microscopic one: for each $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$, consider the measures

$$\mu_{\boldsymbol{u},L,t}(\varphi) := \frac{1}{L^d} \sum_{x \in \Lambda_L} \varphi(L^{-1}x)\boldsymbol{u}_x(L^2t) = \frac{1}{L^d} \sum_{x \in \Lambda_L} \boldsymbol{u}_x(L^2t)\delta_{L^{-1}x}(\varphi).$$

The measure $\mu_{\boldsymbol{u},L,t}$ describes the energy density in the microscopic system at the microscopic time L^2t. The choice for this time scaling (called *parabolic* or *diffusive* scaling) comes from the fact that equation (2-1) is invariant under such a scaling, so it presents itself as the natural one. If the macroscopic dynamics must arise from the microscopic dynamics, the we expect that $\mathbb{P}_0$-a.s.

$$\lim_{L\to\infty} \frac{1}{L^d} \sum_{x \in \Lambda_L} \boldsymbol{u}_x(L^2t)\delta_{L^{-1}x}(\varphi) = \int_{\mathbb{R}^d} u(x,t)\varphi(x)dx,$$

where u satisfies (2-1).

It turns out that the above limit is hard to justify even at the heuristic level, so, as a preliminary step, one would be rather happy even proving its averaged version:[6]

$$\lim_{L\to\infty} \frac{1}{L^d} \sum_{x \in \Lambda_L} \mathbb{E}_0(\boldsymbol{u}_x(L^2t)\delta_{L^{-1}x}(\varphi)) = \int_{\mathbb{R}^d} u(x,t)\varphi(x)dx. \tag{2-3}$$

[5] To simplify matter we assume that the volume of Λ is one.

[6] Note that here we abuse notation and use $\mathbb{P}_0, \mathbb{E}_0$ to designate, respectively, the measure and expectation in path space determined by the initial measure (that we also called $\mathbb{P}_0$, hence the abuse). Of course, in the deterministic case all is determined by the initial condition, but in the random case the measure in path space describes also the randomness of the dynamics.

These type of results are called *hydrodynamic limits* and have been first obtained in some generality in the context of stochastic microscopic dynamics, Kipnis, Marchioro, and Presutti [1982], De Masi, Ianiro, Pellegrinotti, and Presutti [1984], Guo, Papanicolaou, and Varadhan [1988], Varadhan [1993], and Olla, Varadhan, and Yau [1993]. For an overview of the hydrodynamic limit see Spohn [1991].

I have thus specified what can be considered a satisfactory explanation of the emergence of the macroscopic dynamics (2-1) from a microscopic model. Of course, this, rather than being the end of the story, is just a starting point. In fact, what are really relevant for physics and applications are the *finite size effects*. That is, the corrections to the macroscopic law coming form the fact that the scale difference (L) is finite and not infinite. This are the type of results that could prove relevant when working at the mesoscopic scale (e.g, nanotechnology).

3 The models: microscopic dynamics

To make precise the model we have to specify the dynamics. Let us start from a Hamiltonian dynamics: this is the one physicists would ultimately like to study.

By our simplifying homegeneity hypothesis the local systems have all the same local Hamiltonians

$$h(q,p) = \frac{1}{2}\langle p, \boldsymbol{m}^{-1} p\rangle + U(q) \tag{3-1}$$

for some strictly positive matrix $\boldsymbol{m}$ and smooth potential U. To simplify notation we assume $\boldsymbol{m} = \mathbb{1}$.[7] The global Hamiltonian is the sum of the local Hamiltonians and of the interaction between near by systems. For simplicity again we assume that the interaction takes place only among nearest neighbors. Also, since we are considering the system as a bunch of interacting systems, we expect the typical internal energy of a local system (binding energy) to be much larger than the interaction energy. We are thus led to a global Hamiltonian of the form

$$H_{\varepsilon,L}(q,p) = \sum_{x\in\Lambda_L} h(q_x,p_x) + \frac{\varepsilon}{2}\sum_{x\in\Lambda_L}\sum_{\|y-x\|=1} V(q_x,q_y) \tag{3-2}$$

where $V(q,q') = V(q',q)$ is a symmetric smooth potential.

We can now specify what we mean by the internal energy at site x:

$$\boldsymbol{u}_x = h(q_x,p_x) + \frac{\varepsilon}{2}\sum_{\|y-x\|=1} V(q_x,q_y). \tag{3-3}$$

[7] Note that we can always reduce to this situation by changing the definition of the scalar product.

We added to the local Hamiltonian the interaction energy so that $\sum_x \boldsymbol{u}_x = H_{\varepsilon,L}$, thus the $\boldsymbol{u}_x$ account for all the energy in the system.

Remark 3.1. *Note that we are considering the case of a body in isolation. In reality the bodies are in contact with the exterior that can be thought as a thermal reservoir at some given temperature. This is, of course, an extremely important problem but it has several extra difficulties (for example, one has to decide a model for the thermal reservoir, and this is the subject of many debates; the invariant measure of the dynamics is not known explicitly, and even establishing its existence is a challenge, see Section 4.1 for more details). Accordingly, to keep the exposition simple, we will not discuss boundary conditions and we will only consider isolated bodies.*

Let ϕ_t be the flow generated by (3-2) via the usual Hamilton equations

$$\begin{aligned} \dot{q} &= \partial_p H_{\varepsilon,L} \\ \dot{p} &= -\partial_q H_{\varepsilon,L}. \end{aligned} \tag{3-4}$$

We have already mentioned that the semigroup (2-2) is defined as $\mathfrak{L}_t f = f \circ \phi_t$. A simple computation shows that the generator of $\mathfrak{L}_t$ is given by

$$\begin{aligned} \frac{d}{dt}\mathfrak{L}_t f|_{t=0} =: \mathbb{X} f &= \sum_{x\in\Lambda_L} \langle p_x, \partial_{q_x} f\rangle - \langle \nabla U(q_x), \partial_{p_x} f\rangle \\ &\quad - \varepsilon \sum_{x\in\Lambda_L} \sum_{\|y-x\|=1} \langle \partial_{q_x} V(q_x, q_y), \partial_{p_x} f\rangle, \end{aligned}$$

where we have used the symmetry of V.

The first problem in tackling the above dynamics is that we are interested in the properties of the system for a very long time (of order L^2). At the moment the only dynamics that are well understood for arbitrary long times are: a) completely integrable systems; b) strongly chaotic systems. The first possibility is of course much simpler, unfortunately it is very non generic. The interaction between different local systems will typically break the complete integrability leading to a system that we have no tools to analyse.

Of course, one could consider very special global systems that are completely integrable, for example a system in which all the potentials are quadratic (harmonic systems), leading to linear Hamilton equations (3-4). Indeed the exploration of such systems started a long time ago Rieder, J. Lebowitz, and E. Lieb [1967] and Lanford, J. L. Lebowitz, and E. H. Lieb [1977] but it yields an anomalous diffusion due to the existence of many conserved quantities beside the energy Zotos [2002]. This goes against the general consensus that the only locally conserved quantity should be the energy. To ensure such a fact one can introduce some stochasticity in the system (either in the interactions or in the local dynamics) and indeed several very interesting results have been obtained concerning harmonics crystals with some randomness, see Olla and Sasada [2013], Basile and Olla [2014], Jara, Komorowski, and Olla [2015], and Komorowski and Olla [2016, 2017] or the review Basile, Bernardin, Jara, Komorowski, and Olla [2016] and references therein.

In general, the introduction of noise makes the problem much more tractable. If the noise is sufficiently strong, then it is possible to establish the full hydrodynamic limit Olla, Varadhan, and Yau [1993], Liverani and Olla [1996], and Fritz, Liverani, and Olla [1997], but also for a very degenerate noise relevant partial results can be obtained Liverani and Olla [2012].

The alternative is to consider strongly chaotic local dynamics. This point of view has been first advocated in a precise manner by Gaspard and Gilbert [2008], in which they propose to study a billiard type model inspired by Bunimovich, Liverani, Pellegrinotti, and Suhov [1992].[8] This is the point of view I wish to pursue: assume that the Hamiltonian flow associate to the Hamiltonian h is strongly chaotic (an *Anosov* flow with *exponential decay of correlations*).

Nevertheless, as we mentioned already, the introduction of a stochastic part in the dynamics is very instructive. To illustrate this we will consider the case in which the interaction between nearby systems has a very strong random component. We will see that this can be partially justified as a mesoscopic regime, see Theorem 5.1, but for the time being it is just an heuristic tool. We assume that near by systems exchange their velocities and that on each kinetic energy surface takes place a diffusion. To make the statement precise, consider the vector fields

$$\begin{aligned} Y_{i,x} f(q,p) &= \langle y_i(p_x), \partial_{p_x} f(q,p)\rangle \\ X_{x,y} f &= p_x \partial_{p_y} f - p_y \partial_{p_x} f \end{aligned}$$

where the vectors $\{y_i(p)\}$ spans the tangent space of the kinetic energy surface $\{\bar p \in \mathbb{R}^d : \bar p^2 = p^2\}$ at the point p. We then consider the semigroup generated by

$$\begin{aligned} &\mathbb{X}_\nu = (1-\nu)\mathbb{X} + \nu S \\ &S = \sum_{x\in\Lambda_L}\sum_i Y_{i,x}^2 + \sum_{x\in\Lambda_L}\sum_{\|x-y\|=1} X_{x,y}^2. \end{aligned} \tag{3-5}$$

Let $\mathfrak{L}_{\nu,t}$ be the semigroup generated by $\mathbb{X}_\nu$. Note that $\mathfrak{L}_{0,t} = \mathfrak{L}_t$ is the deterministic dynamics, while $\mathfrak{L}_{1,t}$ is a purely stochastic (diffusive) dynamics that does not move the q and preserves the kinetic energy $\sum_x p_x^2$. Thus, it reduces to a purely momenta dynamics in which the positions do not play any role.

[8] Billiards do not fall in the class of models described by (3-2) because the potential is not smooth as they have hard-core interactions. However, they are Hamiltonian and morally similar. We will comment further on hard core models in Section 5.2.

4 Invariant measures, reversibility and currents

It is well known that a Hamiltonian flow leaves invariant the Liouville measure $m_{L,E}$, that is the uniform measure on the energy surface $\mathfrak{M}_{L,E}$. This implies that the total energy is an invariant quantity for the generators $\mathfrak{L}_{\nu,t}$ and the Liouville measures $m_{L,E}$ are invariant. Moreover, for $\varepsilon = 0$, all the Liouville measures $m_{L,\overline{E}}$, $\overline{E} = (E_x)$, supported on the energy surfaces $\mathfrak{M}_{L,\overline{E}} = \{(q,p) \in \mathfrak{M}_{L,E} \;:\; \boldsymbol{u}_x = E_x, \sum_x E_x = E\}$, are invariant.

A satisfactory class of initial measures $\mathbb{P}_0$ we may wish to consider is given by

$$\mathbb{E}_0(f) = \rho(p,q) m_{L,\overline{E}^*} \tag{4-1}$$

for some smooth integrable function ρ and energies $\sum_x E^*_x = E$. That is, we are allowed to fix the energies (slow variables) but not the fast variables.

Note that

$$m_{L,E}(fSf) = -\sum_{x\in\Lambda_L}\sum_i m_{L,E}((Y_{i,x}f)^2) - \sum_{x\in\Lambda_L}\sum_{\|x-y\|=1}((X_{x,y}f)^2),$$

which implies

$$m_{L,E}(f\cdot Sf) = m_{L,E}(Sf\cdot f). \tag{4-2}$$

On the contrary, since $m_{L,E}(f\cdot f\circ\phi_t) = m_{L,E}(f\circ\phi_{-t} f)$, we have

$$m_{L,E}(f\mathbb{X}f) = -m_{L,E}(f\mathbb{X}f). \tag{4-3}$$

A semigroup, with the property (4-2) (that is, its generator is self-adjoint), is called *reversible*.

Reversibility is an important property for Markov systems, which has many relevant consequences Kipnis and Varadhan [1986]. Unfortunately, the generator of a deterministic system is anti-selfadjoint (see (4-3)), which is as far as possible from reversible.

However, there exists a seemingly rather different definition of reversibility in the context of flows. A definition which also has momentous consequences Gallavotti and Cohen [1995]. We call a flow ϕ_t *reversible* if there exists an involution $\boldsymbol{i} : \mathfrak{M}_L \to \mathfrak{M}_L$, that is $\boldsymbol{i}\circ\boldsymbol{i} = \mathrm{id}$, such that

$$\phi_t\circ\boldsymbol{i} = \boldsymbol{i}\circ\phi_{-t}.$$

In the case of a Hamiltonian system it is trivial to check that the flow is reversible with the involution $\boldsymbol{i}(q,p) = (q,-p)$.

Note that the macroscopic evolution associated to the equation (2-1) can also be seen as a semigroup where the generator is $Af = \mathrm{div}(\kappa(f)\nabla f)$.[9] However such equation

[9] Note that the (2-1) is really the equation for the density of a measure, so technically the semigroup here is the adjoint of the one we were discussing above.

is *irreversible*, where the word refers to the fact that backward dynamics is not well defined.[10] Thus the two different definitions of reversibility are yet another manifestation of a long standing conundrum: the relation between microscopic reversibility and macroscopic irreversibility (or: how to explain the time arrow).

We are now ready to discuss another important issue: the *current*. The current simply describes the change in energy of a local system. A direct computation yields, for all $s < t$,

$$\frac{d}{dt}\mathbb{E}_0(\boldsymbol{u}_x(t) \mid \mathfrak{F}_s) = \mathbb{E}_0(\mathbb{X}_\nu \boldsymbol{u}_x(t) \mid \mathfrak{F}_s) = \sum_{\|y-x\|=1} \mathbb{E}_0(j_{x,y}(t) \mid \mathfrak{F}_s)$$

$$j_{x,y} = (1-\nu)\frac{\varepsilon}{2}\left[p_y \partial_{q_y} V(q_x, q_y) - p_x \partial_{q_x} V(q_x, p_y)\right] + \nu(p_y^2 - p_x^2),$$

where $\mathfrak{F}_s$ is the σ-algebra determined by the variables $\{q(\tau), p(\tau)\}_{\tau \le s}$. Note that $j_{x,y} = -j_{y,x}$, so the total energy is conserved.

In the case $\nu = 1$, since the q do not evolve, the energies $u_x(t)$ differ from $p_x^2(t)$ only by a constant, so we get a closed equation for the kinetic energy

$$\frac{d}{dt}\mathbb{E}_0(p_x^2(t)) = \sum_{\|y-x\|=1} \mathbb{E}_0(p_y^2(t)) - \mathbb{E}_0(p_x^2(t)).$$

in this case the current is an exact discrete gradient. Also it is not hard to prove that the measures $\mu_{p^2,L,t}$ are tight, so for any convergent subsequence we have, for each $\varphi \in \mathcal{C}^\infty$,

$$\mu_{p^2,L_j,t}(\varphi) - \mu_{p^2,L,0}(\varphi) = \int_0^t \mu_{p^2,L_j,t}(\Delta\varphi) + \mathcal{O}(L_j^{-1}).$$

The above, using the notation of (2-3), yields

$$\partial_t u(x,t) = \Delta u(x,t)$$

in the weak sense. This is an extreme manifestation of the fact that gradient currents are easier to treat since a Laplacian is already implicit in the current. See Guo, Papanicolaou, and Varadhan [1988] to see how to treat the general gradient case. When the current is not of a gradient type (as in the case $\nu \neq 1$), then much more work is needed, see Varadhan [1993].

The standard tools to deal with the non-gradient case (say $\nu = 0$) seem require two facts:

[10] In a sense the equation is irreversible if considered a deterministic equation (which is its physical meaning). It is instead reversible, at least in the case $\kappa(f) = \kappa$, if the associated semigroup is interpreted as the semigroup describing a random process (Brownian motion). Sorry for the ambiguity.

a) the well posedness of the Green-Kubo formula

$$\kappa_\varepsilon = \beta^2\varepsilon^2 \int_0^\infty \sum_{x\in\mathbb{Z}^d} \mathbb{E}_\beta\left(j_{x,x+1}(t)j_{0,1}(0)\right)dt, \tag{4-4}$$

for the infinite system at equilibrium with inverse temperature β.

b) A spectral gap of order L^{-2} for the dynamics $\mathcal{L}_t$ in a region of size L.

Of course, the latter refers to stochastic systems where the gap is meant in L^2 or in some simple Sobolev space. In the deterministic case on such spaces there is no gap at all. Yet, there can be exponential decay of correlations for smooth observables, Accordingly, it is likely that any possible proof will require a decay of correlations of type $e^{-CL^{-2}t}$ for reasonable observables and the dynamics in a region of size L.

Accordingly, all the known approaches require a sharper, quantitative, information on the rate of convergence in the formula (4-4). Note that, even assuming that the flow on each energy surface of the local dynamics is Anosov, already the study of two interacting systems is currently out of reach. Indeed, when two systems interact only the total energy (and not the individual ones) is conserved. Hence, two interacting systems can be viewed as a *partially hyperbolic flow* with a three dimensional central direction. No result whatsoever is currently available on the rate of mixing for such systems, let alone a larger collection of interacting systems. See Bonatti, Díaz, and Viana [2005] for an overview on partially hyperbolic systems.

4.1 Other models.

Let us briefly discuss other possibile microscopic models. One possibility we already mentioned is that $M_* = \mathbb{R}^{2n_*}$ but the confinement is provided by a potential, these are essentially anharmonic chains (we have already mentioned the harmonic case). The first example of such a model goes back to Fermi, Pasta, and Ulam [1965]. The FPU models have proven extremely difficult to investigate, even numerically Benettin, Livi, and Ponno [2009], Benettin, Christodoulidi, and Ponno [2013], and Dauxois, Peyrard, and Ruffo [2005]. However, the study of FPU models has shown that the route from microscopic to macroscopic is much subtler than one can naively imagine and metastable states can play an important role. From the rigorous point of view almost nothing is known, apart from some zero energy density results that are not so relevant in the present context. On the other hand, if one considers the case in which the system is not isolated but it is in contact with external heat baths, then, in some cases, the existence of a stationary measure is known Eckmann, Pillet, and Rey-Bellet [1999] and Eckmann and Hairer [2001] although its properties are still not well understood.

Another possibility is to consider hard core potential, e.g. billiards. This is also a promising line of thought, very close in spirit to the one presented here. See Section 5.2 for details.

In the last years there have also been attempts to investigate models with mass transport, but with independent particles that can exchange energy only interacting with some array of localised systems, typically discs. Again such systems are in contact with reservoirs that can emit and absorb particles. This are intriguing and illuminating models for which is, at times, possible to establish the existence of a stationary measure and some of its properties Eckmann and Young [2004, 2006], Collet and Eckmann [2009], and Yarmola [2014].

In fact, there are many other relevant papers strictly connected to the matter at hand. It is impossible to quote them all, here is a very partial selection Dolgopyat and Liverani [2008], Dolgopyat, Keller, and Liverani [2008], Bricmont and Kupiainen [2007], Lefevere and Zambotti [2010], and Ruelle [2012].

5 A two steps strategy

By the above discussion, the purely deterministic case seems completely out of reach of current techniques. It is thus necessary to try to devise a line of attack that deals with the problems one at a time. The first, natural, idea is to leverage on our understanding of the dynamics for $\varepsilon = 0$. Of course, when $\varepsilon = 0$ there is no exchange of energy, so we must start to look at the case when ε is "infinitesimally" small. One way to formalise precisely such a situation is to investigate if some universal behaviour takes place for small ε.

5.1 Soft core potentials.

This is the model we have discussed so far in the case $\nu = 0$. Only, now we define the random variables $\mathcal{E}_{L,\varepsilon,x}(t) = u_x(\varepsilon^{-2}t)$ and consider the limit $\varepsilon \to 0$, keeping fixed the size of the system. In other words we look at the local energy when the interaction between near by systems is very small, but rescale time in order to be able to see some evolution.

The choice of the scaling ε^{-2} is due to the fact that, in equilibrium, the currents have zero average, hence we expect the exchange of energy between near by systems to be due to fluctuations. This means that, very naively, the variation of energy at site x and time t can be thought as the sum of t zero average independent random variables of size ε. By the central limit theorem one then expects that a change of energy of order one takes place only at time ε^{-2}.

The above super naive picture can indeed be made rigorous in the special case of *contact* Anosov flows. Indeed, it is known that contact Anosov flows exhibit exponential decay of correlations Liverani [2004]. In the Hamiltonian (3-2), this corresponds to the requirement that the local Hamiltonian (3-1) be of the form $h(q,p)=\frac{1}{2}\langle p,p\rangle$ and that M is a compact manifold of strictly negative curvature. Note that the requirement that the local dynamics be a geodesic flow in negative curvature is not so artificial as it might appear at first sight. Indeed there exists mechanical models for which this is exactly the case Hunt and MacKay [2003]. We have the following result.

Theorem 5.1 (Dolgopyat and Liverani [2011]). *For each $L\in\mathbb{N}$ and $n_*\geq 3$,*[11] *the process $\{\mathcal{E}_{L,\varepsilon,x}(t)\}_{x\in\Lambda_L}$, with initial conditions satisfying $\{\mathcal{E}_{L,\varepsilon,x}(0)=E_x>0\}_{x\in\Lambda_L}$, converges in law to a limit $\{\mathcal{E}_{L,x}\}_{x\in\Lambda_L}$ satisfying the mesoscopic SDE*

$$
\begin{aligned}
d\mathcal{E}_{L,x} &= \sum_{|x-y|=1}\boldsymbol{b}(\mathcal{E}_{L,x},\mathcal{E}_{L,y})dt+\sum_{|x-y|=1}\boldsymbol{a}(\mathcal{E}_{L,x},\mathcal{E}_{L,y})dB_{x,y}\\
\mathcal{E}_{L,x}(0) &= \bar{\boldsymbol{u}}_x>0
\end{aligned}
\tag{5-1}
$$

where $\boldsymbol{b}(\mathcal{E}_{L,x},\mathcal{E}_{L,y})=-\boldsymbol{b}(\mathcal{E}_{L,y},\mathcal{E}_{L,x})$, $\boldsymbol{a}(\mathcal{E}_{L,x},\mathcal{E}_{L,y})=\boldsymbol{a}(\mathcal{E}_{L,y},\mathcal{E}_{L,x})$ and $B_{x,y}=-B_{y,x}$ are independent standard Brownian motions.

The result includes the fact that the SDE is well posed, in the sense of the uniqueness of the martingale problem, Stroock and Varadhan [2006]. To prove the latter it is necessary to show that zero is unreachable. Indeed, if zero were reacheable, then the equation (5-1) wold have to be supplemented by boundary conditions, since by definition energies are positive. In turn, to prove unreachability of zero it is necessary to acquire precise informations on the form of the diffusion coefficient and drift. In Dolgopyat and Liverani [2011] it is shown that $\boldsymbol{b},\boldsymbol{a}\in\mathcal{C}^\infty((0,\infty)^2)$ and, for $\mathcal{E}_{L,x}\leq\mathcal{E}_{L,y}$,

$$
\begin{aligned}
\boldsymbol{a}(\mathcal{E}_{L,x},\mathcal{E}_{L,y})^2 &= \frac{A\mathcal{E}_{L,x}}{\sqrt{2\mathcal{E}_{L,y}}}+\mathcal{O}\left(\mathcal{E}_{L,x}^{\frac{3}{2}}\mathcal{E}_{L,y}^{-1}\right)\\
\boldsymbol{b}(\mathcal{E}_{L,x},\mathcal{E}_{L,y}) &= \frac{An_*}{2\sqrt{2\mathcal{E}_{L,y}}}+\mathcal{O}\left(\mathcal{E}_{L,x}^{\frac{1}{2}}\mathcal{E}_{L,y}^{-1}\right),
\end{aligned}
$$

Note that the only invariant measures for (5-1) are measures absolutely continuous w.r.t. Lebesgue with density $h_\beta=\prod_{x\in\Lambda_L}\mathcal{E}_{L,x}^{\frac{n_*}{2}-1}e^{-\beta\mathcal{E}_{L,x}}$.

The SDE corresponds to a parabolic PDE with generator

$$
\mathbb{X}_L=\frac{1}{2h_0}\sum_{|x-y|=1}(\partial_{\mathcal{E}_{L,x}}-\partial_{\mathcal{E}_{L,y}})h_0\boldsymbol{b}^2(\partial_{\mathcal{E}_{L,x}}-\partial_{\mathcal{E}_{L,y}}).
\tag{5-2}
$$

We have thus a *mesoscopic* equations in which the evolution of all the degree of freedom, apart from the energies, can be ignored. Even more remarkably, the generator $\mathbb{X}_L$, with respect to the invariant measure, is *reversible*.

[11] The Theorem should also be true for $n_*=2$, but it is harder to prove. Instead it does not make sense for $n_*=1$, since in such a case the local Hamiltonian is completely integrable and the flow cannot be Anosov.

This is a consequence of the microscopic reversibility of the flow. Indeed, the involution $(q, p) \to (q, -p)$ that exchanges the direction of time, reduces to the identity in the energy variables. Thus, we see not only how irreversibility arises (equations (5-1) are irreversible in the sense that they display a time arrow: a distribution tends to equilibrium going forward in time, but not going backward), but we see also a non trivial relation between (deterministic) reversibility for the microscopic dynamics and (stochastic) reversibility for the macroscopic one (which shows up already at the mesoscopic level).

The generator (5-2) is a dynamics only on the energies, so it has the same flavour as S in (3-5). However, the associated current is not of gradient type. Yet, it is conceivable that the study of the dynamics (5-1) is much easier than the study of the original deterministic dynamics. So it natural to try to perform the hydrodynamic limit on the mesoscopic dynamics.

To this end, as we have already mentioned, it seems necessary to have a spectral gap of size L^{-2} for the operator $\mathbb{X}_L$. At the moment it is unclear if such a fact holds true or not, the problem stemming from the fact that at high energy the diffusion coefficient vanishes. This is a consequence of the fact that at high energy near by systems (in the original deterministic system) interact very little since the size of the potential is very small compared with the available energy.

The situation improves if one starts from a system with some stochasticity as in Liverani and Olla [2012]. Indeed, in Liverani and Olla [ibid.] is considered an anharmonic chain with and energy preserving noise and we establish the same type of result as in (5-1) but now

$$
\begin{aligned}
\boldsymbol{a}(\mathcal{E}_{L,x}, \mathcal{E}_{L,y})^2 &\sim A\mathcal{E}_x\mathcal{E}_{L,y} \\
\boldsymbol{b}(\mathcal{E}_{L,x}, \mathcal{E}_{L,y}) &\sim \mathcal{E}_{L,x} - \mathcal{E}_{L,y}.
\end{aligned}
$$

For such a $\boldsymbol{b}$ the needed spectral gap has been established by Olla and Sasada [2013] comparing $\mathbb{X}_L$ with the generator of the Kac model. This suffices to prove that the fluctuations in equilibrium satisfy the heat equation Liverani, Olla, and Sasada [n.d.].

Hence there is a concrete hope to obtain the heat equation starting from a deterministic dynamics via a two step procedure: first take a weak coupling limit to obtain a mesoscopic equation involving only the energies, then perform the hydrodynamic limit on the latter dynamics.

This is encouraging, yet a natural question arises: is there any relation between the behaviour of the original model, possible for ε very small, and the result of this two step procedure? To answer precisely to such a question would be equivalent to solve our original problem, however even an heuristic answer is not obvious.

A (non trivial) formal computation, see Bernardin, Huveneers, J. L. Lebowitz, Liverani, and Olla [2015], shows that if κ_ε is the diffusivity, as defined in (4-4), for the original model (that we assume finite) and κ_M the diffusivity of the mesoscopic dynamics (that can be proven finite), then

$$\begin{aligned}\kappa_\varepsilon &= \varepsilon^2\kappa_M + \mathcal{O}(\varepsilon^3)\\ \kappa_M &= \mathbb{E}_0(\boldsymbol{a}(\mathcal{E}_0,\mathcal{E}_1)^2) + \sum_x \int_0^\infty \mathbb{E}_0\left(\boldsymbol{b}(\mathcal{E}_0,\mathcal{E}_1)(0)\boldsymbol{b}(\mathcal{E}_x,\mathcal{E}_{x+1})(s)\right) ds.\end{aligned}$$

This suggests that the mesoscopic dynamics captures the main effect of the energy transport and that it is an effective approximation of the behaviour of the microscopic deterministic dynamics.

If the above were true, then it should be possible to use the stochastic dynamics as a first approximation of the long term statistical properties of the original dynamics well beyond the time scale ε^{-2}, which is the time scale established by Theorem 5.1. This is an intriguing possibility that leads to a rather vast research program.

5.2 Hard core potential. Before continuing the discussion on the possibility to extend Theorem 5.1, it is worth to discuss a different possibility: hard core interactions. Indeed, it is quite possible that the presence of hard core at the microscopic level does manifest itself also at macroscopic level. For example it is likely that hard core interactions do not manifest the property of a decreasing diffusion coefficient at high energies since when there is a collision the velocities change dramatically also at high energies, contrary to the case of soft interactions.

Unfortunately, while hard core interactions may cure a problem they come at a high cost, since in such a case the discontinuity of the dynamics creates formidable technical problems. Yet, it is certainly very important and instructive to investigate this alternative.

In this case, the unperturbed systems (corresponding to the Hamiltonian $H_{0,L}$) would consist of a region of size L filled by disjoint billiards domains each containing a ball that can move freely apart for the elastic collision with the walls. In such a case the kinetic energy of each ball is conserved and there is no transport of mass or of energy, see Figure 1. To perturb the system, instead of introducing a potential, we shrink a bit the obstacles, so that channels appear between the different tables. If the channels are large enough to allow near by particles to collide, but not so large as to allow the particles to escape the region in which they are confined, then we obtain a systems in which mass transport is still impossible, but energy transport is allowed, see Figure 2.

Some systems of these type are known to be ergodic Bunimovich, Liverani, Pellegrinotti, and Suhov [1992] but, unfortunately, nothing is known about their mixing rate. In particular it is unknown if the Green-Kubo formula is well defined. Yet, one can imitate what has been done in the previous section: consider the limiting case in which the interaction are extremely rare and rescale the time so that, in average, a particle has one collision with another particle in a (macroscopic) unit time. This leads, again, to a two step

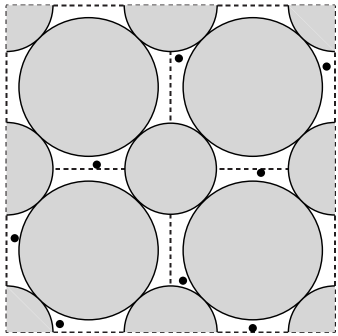


Figure 1: Obstacles gray, particles black. Non interacting particles

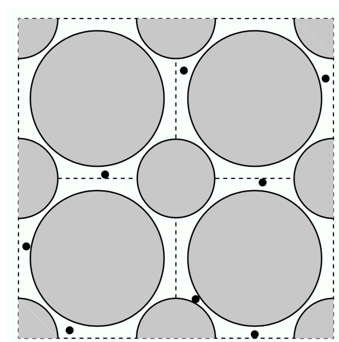


Figure 2: Obstacles gray, particles black. Interacting particles

route to the heat equation. This research program has been put forward in Gaspard and Gilbert [2008] where the authors heuristically derive a mesoscopic equation describing the evolution of the energy with generator

$$\mathbb{X}_L f(\boldsymbol{u}) = \frac{1}{2} \sum_{x \in \Lambda_L} \sum_{\substack{y \in \Lambda_L \\ \|x-y\|=1}} \int_{-\pi}^{\pi} [f(R_\theta^{xy} \boldsymbol{u}) - f(\boldsymbol{u})] \rho(\theta) d\theta \tag{5-3}$$

where R_θ^{xy}, $x \neq y$ is a clockwise rotation of angle θ in the plane $(\boldsymbol{u}_x, \boldsymbol{u}_y)$. The generator (5-3) is the analogous of (5-2) and describe a jump process. This generator is know to have a spectral gap of order L^{-2}, Grigo, Khanin, and Szász [2012] and Sasada [2015]. The situation seems then very promising, unfortunately all attempts to derive rigorously (5-3) have so far failed. Nevertheless, lately there has been some notable technical progresses Baladi, Demers, and Liverani [2018], Dolgopyat and Nándori [n.d.], and Bálint, Nándori, Szász, and Tóth [n.d.] and some relevant results on related models Bálint, Gilbert, Nándori, Szász, and Tóth [2017] and Dolgopyat and Nándori [2016].

6 Partially hyperbolic Fast-slow systems and limit theorems

At the end of Section 5.1 we came to the conclusion that (5-1) might hold for much longer times than ε^{-2} and that this, if true, could help in establishing the Green-Kubo formula and, ultimately, the heat equation. However, to investigate such a possibility is a non trivial task. A task that is best accomplished proceeding by intermediate steps. This leads us to the general problem of studying the long time validity of limit theorems in partially hyperbolic fast-slow systems.

Fast-Slow systems are system in which there are two group of variables that evolve accordingly to very different time scales. For example, the Hamiltonian (3-2), with M a compact Riemannian manifold in negative curvature and the local Hamiltonian (3-1) of the form $h(q,p) = \frac{1}{2}\langle p, p\rangle$, yields a dynamics in which the variables q_x, $v_x = (2u_x)^{-1/2} p_x$ vary on a microscopic time scale of order one, while the variable $u_x = \frac{1}{2} p_x^2$ varies on a timescale of order ε^{-2}. Partial hyperbolicity stems form the fact that for $\varepsilon = 0$ the system foliates in uniformly hyperbolic systems, so the dynamics in the central direction is simply identity, and the central direction persists under perturbations Hirsch, Pugh, and Shub [1977].

If we want to understand the behaviour of such systems for arbitrarily long times it is best to start form the simplest possible example.

6.1 The not so simple simplest example.

We start by defining the one parameter family of maps $F_\varepsilon \in \mathcal{C}^4(\mathbb{T}^2, \mathbb{T}^2)$

$$F_\varepsilon(x,z) = (f(x,z), z + \varepsilon\omega(x,z)), \tag{6-1}$$

where $\partial_x f(x,z) \geq \lambda > 1$ is an expanding map for all z. We then consider the dynamics $(x_n, z_n) = F_\varepsilon^n(x_0, z_0)$ with initial conditions

$$\mathbb{E}(g(x_0, z_0)) = \int_{\mathbb{T}^1} \rho(x) g(x, \bar{z}_0) dx \ , \tag{6-2}$$

where $\bar{z}_0 \in \mathbb{T}^1$, while $\rho \in \mathcal{C}^2(\mathbb{T}^1, \mathbb{R}_+)$.

Let us compare this super simplified model with the Hamiltonian system (3-2). First of all, (6-1) is in discrete time and not continuous time. This is technically much simpler, but morally not so different.

More serious is the fact that the system is not time reversible and has no Hamiltonian or symplectic structure. This makes it rather artificial, however a time reversible system with some symplectic structure could be constructed using an Anosov map on the two torus instead of an expanding map of the circle. Hence, we can consider our model as a preliminary step toward a more realistic one.

Next, (6-1) has only two variables hence the question of taking the hydrodynamic limit makes no sense.[12] However, this model is intended only to explore the possibility to control the statistical properties of (3-2) for a time longer than ε^{-2}. Of course, ultimately this must be done with some uniformity on the number of degree of freedom, but if one cannot do it with one degree of freedom it does not make much sense to think about large systems.

On the bright side, (6-1) has a conserved quantity (z, which plays the role of the local energy in (3-2)) for $\varepsilon = 0$ and, for ε small is a fast-slow system. The quantity $\varepsilon\omega$, which determines the change of the almost conserved quantity z, plays the role of the current.

The local dynamics depends on the conserved quantity z as the local hamiltonian dynamics in (3-2) depends on the local energy. The local dynamics $f(\cdot, z)$ have a strong chaotic character similar to the hypothesis that the local, unperturbed, hamiltonian flows is a contact Anosov flow (or, more generally, enjoys exponential decay of correlations), hence the partial hyperbolicity. The initial conditions are very similar as one can fix exactly the almost conserved quantity (slow variable) but must have a smooth distribution for the fast variables, similarly to (4-1).

It is well known, Baladi [2000], that our hypotheses on (6-1) imply that

[12] Nonetheless one can consider many of such systems weakly coupled, whereby reproducing a situation in which it is possible to perform the hydrodynamic limit. In the simple case in which the fast dynamics does not depend form the slow one, this has been done, obtaining indeed the heat equation, Bricmont and Kupiainen [2013].

1. for each $z \in \mathbb{T}^1$, $f(\cdot, z)$ has a unique SRB measure μ_z which is absolutely continuous with respect to Lebesgue and has density $h(\cdot, z)$
2. $h \in \mathcal{C}^2(\mathbb{T}^2, \mathbb{R}_+)$.

The above facts take care of another apparent difference between (3-2) and (6-1): the former has an explicit and natural invariant measure (Liouville). Now we know that, even though not totally apparent, the same holds for (6-1). The measures μ_z will play the role of the equilibrium measures.

However, there is a last issue: due to the reversibility and the hamiltonian structure in (3-2) the average of the current is always zero. This is the reason why the evolution of the energy happens on the time scale ε^{-2} rather than on the scale ε^{-1}. In the present simplified setting this would correspond to the conditions $\bar{\omega}(z) := \mu_z(\omega(\cdot, z)) = 0$. Unfortunately, this turns out to be a much harder problem. At the moment are available partially satisfactory results only in the case $\bar{\omega}(z) \neq 0$. More precisely, in the case in which $\bar{\omega}$ has only finitely many non-degenerate zeroes. In this case the natural time scale in which the slow variable evolves is ε^{-1}, however the problem of understanding the statistical properties of the system for arbitrarily long times remain a non trivial challenge and its study is a preliminary step to attack the, harder, case $\bar{\omega} \equiv 0$.

To describe the existing results it is convenient to introduce the continuous paths

$$z_\varepsilon(t) = z_{\lfloor \varepsilon^{-1} t \rfloor} + (\varepsilon^{-1} t - \lfloor \varepsilon^{-1} t \rfloor)(z_{\lfloor \varepsilon^{-1} t \rfloor + 1} - z_{\lfloor \varepsilon^{-1} t \rfloor}), \quad t \in [0, T].$$

The paths $z_\varepsilon \in \mathcal{C}^0([0, T], \mathbb{R})$ are random variables due to the randomness of the initial conditions. Since the $\{z_\varepsilon\}$ are uniformly Lipschitz, they belong to a compact set in $\mathcal{C}^0([0, T], \mathbb{R})$, hence they have convergent subsequences. It is possible to show that all the accumulation points $\bar{z}$ must satisfy the ODE

$$\begin{aligned}
&\dot{\bar{z}} = \bar{\omega}(\bar{z}) \\
&\bar{z}(0) = \bar{z}_0 \\
&\bar{\omega}(z) = \int_{\mathbb{T}^1} \omega(x, z) h(x, z) dx = \mu_z(\omega(\cdot, z)).
\end{aligned}$$

This type of results goes back, at least, to Anosov [1960] and Bogoliubov and Mitropolsky [1961] in the early '60.

Next, let us consider the quantity $\zeta_\varepsilon(t) = \varepsilon^{-\frac{1}{2}} [z_\varepsilon(t) - \bar{z}(t)]$. These are the fluctuations around the average. To discuss this case we need to recall that a function $\phi \in \mathcal{C}^0(\mathbb{T})$ is said to be a *(continuous) coboundary (with respect to a map $f : \mathbb{T} \to \mathbb{T}$)* if there exists $\beta \in \mathcal{C}^0(\mathbb{T})$ so that

$$\phi = \beta - \beta \circ f.$$

Two functions $\phi_1, \phi_2 \in \mathcal{C}^0(\mathbb{T})$ are said to be *cohomologous (with respect to f)* if their difference $\phi_2 - \phi_1$ is a coboundary (with respect to f). We make the non-degeneracy assumption that, for each $z \in \mathbb{T}$, the function $\omega(\cdot, z)$ is not cohomologous to a constant with respect to f_z. Note that in De Simoi and Liverani [n.d.] it is shown that this assumption is in fact generic in $\mathcal{C}^2$.

A computation using the decay of correlations of the maps $f(\cdot, z)$ yields

$$\mathbb{E}([\zeta_\varepsilon(t) - \zeta_\varepsilon(s)]^4) \leq C|t-s|^2.$$

Hence, by Kolmogorov criteria, the sequence is tight. It is possible to show that the accumulation points ζ of ζ_ε satisfy

$$\begin{aligned} d\zeta &= \bar{\omega}'(\bar{z}(t))\zeta(t)dt + \hat{\boldsymbol{\sigma}}(\bar{z}(t))dB \\ \zeta(0) &= 0 \end{aligned} \tag{6-3}$$

where B is the standard Brownian, $\hat{\boldsymbol{\sigma}} > 0$ is given by the Green-Kubo formula

$$\hat{\boldsymbol{\sigma}}(z)^2 = \mu_z\left(\hat{\omega}(\cdot, z)\hat{\omega}(\cdot, z)\right) + 2\sum_{m=1}^{\infty} \mu_z\left(\hat{\omega}(f_z^m(\cdot), z)\hat{\omega}(\cdot, z)\right),$$

and we have used the notation $f_z(x) = f(x, z)$.

As the above equation has a unique solution, this identifies the limit.

This type of results are much more recent and, in the above form, have been obtained by Dolgopyat and Kifer at the beginning of the new millenium, Dolgopyat [2005] and Y. Kifer [2004], but see De Simoi and Liverani [2015] for a pedagogical exposition.

We have thus seen that z_ε is close to $\bar{z} + \sqrt{\varepsilon}\zeta$. On the other hand it is possible to show, J. I. Kifer [1976], that $\bar{z} + \sqrt{\varepsilon}\zeta$ is close to the solution $\tilde{z}_\varepsilon$ of the stochastic differential equation

$$\begin{aligned} d\tilde{z}_\varepsilon &= \bar{\omega}(\tilde{z}_\varepsilon)dt + \sqrt{\varepsilon}\hat{\boldsymbol{\sigma}}(\tilde{z}_\varepsilon)dB \\ \tilde{z}_\varepsilon(0) &= \bar{z}_0. \end{aligned} \tag{6-4}$$

Thus the motion is described by an ODE with a small random noise of the type introduced by Hasselmann [1976] to model climate and extensively studied by Ventcel and Freidlin [1969] and Freidlin and Wentzell [2012] and J. I. Kifer [1974, 1977] and Y. Kifer [1981] in the 70's.

The above is the equivalent, in the present context, of Theorem 5.1. We can now pose for the current model the question that we would like to answer in the previously described context: what happens on time scales longer than ε^{-1}. A first result is the following:

Theorem 6.1 (De Simoi, Liverani, Poquet, and Volk [2017] Corollary 3.3)**.** *For any $\beta > 0$, $\alpha \in (0, \beta)$, $\varepsilon \in (0, \varepsilon_0)$, and $t \in [\varepsilon^{1/2000}, \varepsilon^{-\alpha}]$, there exists $C_\beta > 0$ and a coupling $\mathbb{P}_c$ between $z_\varepsilon(t)$ and $\tilde{z}_\varepsilon(t)$, such that:*

$$\mathbb{P}_c(|z_\varepsilon - \tilde{z}_\varepsilon(t)| \geq \varepsilon) \leq C_\beta \varepsilon^{1/2-\beta}.$$

The above result is based on a drastic sharpening of (6-3), which amounts to a local central limit theorem with error term for the diffusion limit.

Theorem 6.2 (De Simoi and Liverani [n.d., Theorem 2.7])**.** *For any $T > 0$, there exists $\varepsilon_0 > 0$ so that the following holds. For any $\beta > 0$, compact interval $I \subset \mathbb{R}$, $|I| \leq 1$, real numbers $\kappa > 0$, $\varepsilon \in (0, \varepsilon_0)$, $t \in [\varepsilon^{1/2000}, T]$, we have:*

$$\frac{\mathbb{E}(\zeta_\varepsilon(t, \cdot) \in \varepsilon^{1/2} I + \kappa)}{\sqrt{\varepsilon}} = \operatorname{Leb} I \left[\frac{e^{-\kappa^2/2\boldsymbol{\sigma}_t^2(\bar{z}_0)}}{\boldsymbol{\sigma}_t(\bar{z}_0)\sqrt{2\pi}}\right] + \mathcal{O}(\varepsilon^{1/2-\beta}).$$

where the variance $\boldsymbol{\sigma}_t^2(z)$ reads

$$\boldsymbol{\sigma}_t^2(z) = \int_0^t e^{2\int_s^t \bar{\omega}'(\bar{z}(r,z))\mathrm{d}r}\, \hat{\boldsymbol{\sigma}}^2(\bar{z}(s, z))\mathrm{d}s.$$

Theorem 6.1 says that the deterministic dynamics remains close to the stochastic one for a time of order almost $\varepsilon^{-\frac{3}{2}}$. Since (6-4) reaches at lest a metastable state in a time of order $\varepsilon^{-1} \ln \varepsilon^{-1}$, Freidlin and Wentzell [2012], it follows that also the deterministic system must reach similar states. This control should be sufficient to start an investigation of the even longer time properties of the system.

It turns out that current techniques to study the statistical properties of partially hyperbolic systems depend on the positivity or negativity of the central Lyapunov exponent. So at the moment it is unclear in which generality the program can be completed. Here we present the best available result, but before stating it is necessary to introduce some notation.

We say that a Lipschitz path h of length T is *admissible* if for any $s \in [0, T]$, $\boldsymbol{\partial} h(s) \subset \operatorname{int} \Omega(h(s))$,[13] where, for $z \in \mathbb{T}$, we define the (non-empty, convex and compact) set

$$\Omega(z) = \{\mu(\omega(\cdot, z)) \mid \mu \text{ is a } f_z\text{-invariant probability}\}.$$

The last condition is:

[13] For each $s \in [0, T]$, $\boldsymbol{\partial} h(s)$ is the *Clarke generalized derivative* of h as the set-valued function:

$$\boldsymbol{\partial} h(s) = \operatorname{hull}\{\lim_{k\to\infty} h'(s_k) \,:\, s_k \to s\}.$$

The set $\boldsymbol{\partial} h(s)$ is compact and non-empty (see Clarke, Ledyaev, Stern, and Wolenski [1998, Proposition 2.1.5]) and so is its graph.

- there exists $i \in \{1, \cdots, n_Z\}$ so that for any $z \in \mathbb{T}$, there exists an admissible $(z, z_{i,-})$-path. We can always assume, without loss of generality, that $i = 1$.

Observe that the above condition is trivially satisfied if $n_Z = 1$.

Theorem 6.3 (De Simoi and Liverani [2016]Main Theorem)**.** *The map F_ε admits a unique SRB measure μ_ε. This measure enjoys exponential decay of correlations for Hölder observables. More precisely: there exist $C_1, C_2, C_3, C_4 > 0$ (independent of ε) such that, for any $\alpha \in (0, 3]$ and $\beta \in (0, 1]$, any two functions $A \in \mathcal{C}^\alpha(\mathbb{T}^2)$ and $B \in \mathcal{C}^\beta(\mathbb{T}^2)$:*

$$|\mathrm{Leb}(A \cdot B \circ F_\varepsilon^n) - \mathrm{Leb}(A)\mu_\varepsilon(B)| \leq C_1 \sup_z \|A(\cdot, z)\|_\alpha \sup_x \|B(x, \cdot)\|_\beta e^{-\alpha\beta c_\varepsilon n},$$

where

$$c_\varepsilon = \begin{cases} C_2\varepsilon / \log \varepsilon^{-1} & \text{if } n_Z = 1, \\ C_3 \exp(-C_4\varepsilon^{-1}) & \text{otherwise.} \end{cases}$$

The proof of the above results is based on the standard pair technique Dolgopyat [2005] and Theorem 6.2, but also on a considerable sharpening of the large deviation results previously obtained for uniformly hyperbolic dynamical systems, notably Y. Kifer [2009], see De Simoi and Liverani [n.d.] for details.

Also note the extremely slow decay of correlations in the case in which more than one sink is present. This is a well known phenomena: *metastability*. A phenomena widely studied in stochastic equations like (6-4), see Freidlin and Wentzell [2012] and Olivieri and Vares [2005], but seen here for the first time in a purely deterministic setting.

7 Final considerations

The previous sections show on the one hand that even establishing partial results for a very simplified system entails a tremendous amount of work. On the other hand we have shown that the research program of studying the dynamics of (3-2) is not totally hopeless, especially if the arguments could be substantially simplified.

In particular, we have put forward the general philosophy of proving that a deterministic dynamics behaves like a stochastic one for a time long enough for the stochastic dynamics to exhibit some asymptotic property. This allows to deduce that the deterministic system shares such asymptotic behaviour. This approach seems to be a powerful point of view that can be applied to many other contexts.

References

D. V. Anosov (1960). "Averaging in systems of ordinary differential equations with rapidly oscillating solutions". *Izv. Akad. Nauk SSSR Ser. Mat.* 24, pp. 721–742. MR: 0126592 (cit. on p. 2679).

Viviane Baladi (2000). *Positive transfer operators and decay of correlations*. Vol. 16. Advanced Series in Nonlinear Dynamics. World Scientific Publishing Co., Inc., River Edge, NJ, pp. x+314. MR: 1793194 (cit. on p. 2678).

Viviane Baladi, Mark F. Demers, and Carlangelo Liverani (2018). "Exponential decay of correlations for finite horizon Sinai billiard flows". *Invent. Math.* 211.1, pp. 39–177. MR: 3742756 (cit. on p. 2677).

P. Bálint, P. Nándori, D. Szász, and I. P. Tóth (n.d.). "Equidistribution for standard pairs in planar dispersing billiard flows". arXiv: 1707.03068v1 (cit. on p. 2677).

Péter Bálint, Thomas Gilbert, Péter Nándori, Domokos Szász, and Imre Péter Tóth (2017). "On the limiting Markov process of energy exchanges in a rarely interacting ball-piston gas". *J. Stat. Phys.* 166.3-4, pp. 903–925. MR: 3607595 (cit. on p. 2677).

Giada Basile, Cédric Bernardin, Milton Jara, Tomasz Komorowski, and Stefano Olla (2016). "Thermal conductivity in harmonic lattices with random collisions". In: *Thermal transport in low dimensions*. Vol. 921. Lecture Notes in Phys. Springer, [Cham], pp. 215–237. MR: 3524226 (cit. on p. 2668).

Giada Basile and Stefano Olla (2014). "Energy diffusion in harmonic system with conservative noise". *J. Stat. Phys.* 155.6, pp. 1126–1142. MR: 3207732 (cit. on p. 2668).

G. Benettin, H. Christodoulidi, and A. Ponno (2013). "The Fermi-Pasta-Ulam problem and its underlying integrable dynamics". *J. Stat. Phys.* 152.2, pp. 195–212. MR: 3082647 (cit. on p. 2671).

G. Benettin, R. Livi, and A. Ponno (2009). "The Fermi-Pasta-Ulam problem: scaling laws vs. initial conditions". *J. Stat. Phys.* 135.5-6, pp. 873–893. MR: 2548597 (cit. on p. 2671).

Cédric Bernardin, François Huveneers, Joel L. Lebowitz, Carlangelo Liverani, and Stefano Olla (2015). "Green-Kubo formula for weakly coupled systems with noise". *Comm. Math. Phys.* 334.3, pp. 1377–1412. MR: 3312438 (cit. on p. 2674).

N. N. Bogoliubov and Y. A. Mitropolsky (1961). *Asymptotic methods in the theory of nonlinear oscillations*. Translated from the second revised Russian edition. International Monographs on Advanced Mathematics and Physics. Hindustan Publishing Corp., Delhi, Gordon and Breach Science Publishers, New York, pp. x+537. MR: 0141845 (cit. on p. 2679).

Christian Bonatti, Lorenzo J. Díaz, and Marcelo Viana (2005). *Dynamics beyond uniform hyperbolicity*. Vol. 102. Encyclopaedia of Mathematical Sciences. A global geometric and probabilistic perspective, Mathematical Physics, III. Springer-Verlag, Berlin, pp. xviii+384. MR: 2105774 (cit. on p. 2671).

F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet (2000). "Fourier's law: a challenge to theorists". In: *Mathematical physics 2000*. Imp. Coll. Press, London, pp. 128–150. MR: 1773043 (cit. on p. 2663).

Jean Bricmont and Antti Kupiainen (2007). "Towards a derivation of Fourier's law for coupled anharmonic oscillators". *Comm. Math. Phys.* 274.3, pp. 555–626. MR: 2328905 (cit. on p. 2672).

– (2013). "Diffusion in energy conserving coupled maps". *Comm. Math. Phys.* 321.2, pp. 311–369. MR: 3063914 (cit. on p. 2678).

Leonid Bunimovich, Carlangelo Liverani, Alessandro Pellegrinotti, and Yurii Suhov (1992). "Ergodic systems of n balls in a billiard table". *Comm. Math. Phys.* 146.2, pp. 357–396. MR: 1165188 (cit. on pp. 2668, 2675).

F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski (1998). *Nonsmooth analysis and control theory*. Vol. 178. Graduate Texts in Mathematics. Springer-Verlag, New York, pp. xiv+276. MR: 1488695 (cit. on p. 2681).

P. Collet and J.-P. Eckmann (2009). "A model of heat conduction". *Comm. Math. Phys.* 287.3, pp. 1015–1038. MR: 2486671 (cit. on p. 2672).

Thierry Dauxois, Michel Peyrard, and Stefano Ruffo (2005). "The Fermi-Pasta-Ulam 'numerical experiment': history and pedagogical perspectives". *European J. Phys.* 26.5, S3–S11. MR: 2167379 (cit. on p. 2671).

A. De Masi, N. Ianiro, A. Pellegrinotti, and E. Presutti (1984). "A survey of the hydrodynamical behavior of many-particle systems". In: *Nonequilibrium phenomena, II*. Stud. Statist. Mech., XI. North-Holland, Amsterdam, pp. 123–294. MR: 757003 (cit. on p. 2666).

Jacopo De Simoi and Carlangelo Liverani (n.d.). "Limit Theorems for Fast-slow partially hyperbolic systems". Preprint, To appear in Inventiones Mathematicae. arXiv: 1408.5453 (cit. on pp. 2680–2682).

– (2015). "The martingale approach after Varadhan and Dolgopyat". In: *Hyperbolic dynamics, fluctuations and large deviations*. Vol. 89. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, pp. 311–339. MR: 3309102 (cit. on p. 2680).

– (2016). "Statistical properties of mostly contracting fast-slow partially hyperbolic systems". *Invent. Math.* 206.1, pp. 147–227. MR: 3556527 (cit. on p. 2682).

Jacopo De Simoi, Carlangelo Liverani, Christophe Poquet, and Denis Volk (2017). "Fast-slow partially hyperbolic systems versus Freidlin-Wentzell random systems". *J. Stat. Phys.* 166.3-4, pp. 650–679. MR: 3607585 (cit. on p. 2681).

Dmitry Dolgopyat (2005). "Averaging and invariant measures". *Mosc. Math. J.* 5.3, pp. 537–576, 742. MR: 2241812 (cit. on pp. 2680, 2682).

Dmitry Dolgopyat, Gerhard Keller, and Carlangelo Liverani (2008). "Random walk in Markovian environment". *Ann. Probab.* 36.5, pp. 1676–1710. MR: 2440920 (cit. on p. 2672).

Dmitry Dolgopyat and Carlangelo Liverani (2008). "Random walk in deterministically changing environment". *ALEA Lat. Am. J. Probab. Math. Stat.* 4, pp. 89–116. MR: 2413089 (cit. on p. 2672).

– (2011). "Energy transfer in a fast-slow Hamiltonian system". *Comm. Math. Phys.* 308.1, pp. 201–225. MR: 2842975 (cit. on p. 2673).

Dmitry Dolgopyat and Péter Nándori (n.d.). "The first encounter of two billiard particles of small radius". Preprint. arXiv: 1603.07590 (cit. on p. 2677).

– (2016). "Nonequilibrium density profiles in Lorentz tubes with thermostated boundaries". *Comm. Pure Appl. Math.* 69.4, pp. 649–692. MR: 3465085 (cit. on p. 2677).

J.-P. Eckmann and M. Hairer (2001). "Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise". *Comm. Math. Phys.* 219.3, pp. 523–565. MR: 1838749 (cit. on p. 2671).

J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet (1999). "Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures". *Comm. Math. Phys.* 201.3, pp. 657–697. MR: 1685893 (cit. on p. 2671).

J.-P. Eckmann and L.-S. Young (2004). "Temperature profiles in Hamiltonian heat conduction". *Europhys. Lett.* 68.6, pp. 790–796. MR: 2136848 (cit. on p. 2672).

– (2006). "Nonequilibrium energy profiles for a class of 1-D models". *Comm. Math. Phys.* 262.1, pp. 237–267. MR: 2200889 (cit. on p. 2672).

Albert Einstein (1905). "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". *Annalen der physik* 17.8, pp. 549–560 (cit. on p. 2662).

E Fermi, J Pasta, and S Ulam (1965). "Los Alamos Report No. LA-1940, 1955". In: *Collected Papers of Enrico Fermi*. Vol. 2. University of Chicago Press Chicago, pp. 977–978 (cit. on p. 2671).

Mark I. Freidlin and Alexander D. Wentzell (2012). *Random perturbations of dynamical systems*. Third. Vol. 260. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the 1979 Russian original by Joseph Szücs. Springer, Heidelberg, pp. xxviii+458. MR: 2953753 (cit. on pp. 2680–2682).

József Fritz, Carlangelo Liverani, and Stefano Olla (1997). "Reversibility in infinite Hamiltonian systems with conservative noise". *Comm. Math. Phys.* 189.2, pp. 481–496. MR: 1480029 (cit. on p. 2668).

G. Gallavotti and E. G. D. Cohen (1995). "Dynamical ensembles in stationary states". *J. Statist. Phys.* 80.5-6, pp. 931–970. MR: 1349772 (cit. on p. 2669).

Pierre Gaspard and Thomas Gilbert (2008). "Heat conduction and Fourier's law by consecutive local mixing and thermalization". *Phys. Rev. Lett.* 101.2, p. 020601 (cit. on pp. 2668, 2677).

A. Grigo, K. Khanin, and D. Szász (2012). "Mixing rates of particle systems with energy exchange". *Nonlinearity* 25.8, pp. 2349–2376. MR: 2956578 (cit. on p. 2677).

M. Z. Guo, G. C. Papanicolaou, and S. R. S. Varadhan (1988). "Nonlinear diffusion limit for a system with nearest neighbor interactions". *Comm. Math. Phys.* 118.1, pp. 31–59. MR: 954674 (cit. on pp. 2666, 2670).

Klaus Hasselmann (1976). "Stochastic climate models part I. Theory". *Tellus* 28.6, pp. 473–485 (cit. on p. 2680).

M. W. Hirsch, C. C. Pugh, and M. Shub (1977). *Invariant manifolds*. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, pp. ii+149. MR: 0501173 (cit. on p. 2677).

T. J. Hunt and R. S. MacKay (2003). "Anosov parameter values for the triple linkage and a physical system with a uniformly chaotic attractor". *Nonlinearity* 16.4, pp. 1499–1510. MR: 1986308 (cit. on p. 2673).

Milton Jara, Tomasz Komorowski, and Stefano Olla (2015). "Superdiffusion of energy in a chain of harmonic oscillators with noise". *Comm. Math. Phys.* 339.2, pp. 407–453. MR: 3370610 (cit. on p. 2668).

Ju. I. Kifer (1974). "Some results concerning small random perturbations of dynamical systems". *Teor. Verojatnost. i Primenen.* 19, pp. 514–532. MR: 0353451 (cit. on p. 2680).

– (1976). "On the asymptotic behavior of transition densities of processes with small diffusion". *Teor. Verojatnost. i Primenen.* 21.3, pp. 527–536. MR: 0431398 (cit. on p. 2680).

– (1977). "Small random perturbations of hyperbolic limit sets". *Uspehi Mat. Nauk* 32.1 (193), pp. 193–194. MR: 0445623 (cit. on p. 2680).

Yuri Kifer (1981). "The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point". *Israel J. Math.* 40.1, pp. 74–96. MR: 636908 (cit. on p. 2680).

– (2004). "Averaging principle for fully coupled dynamical systems and large deviations". *Ergodic Theory Dynam. Systems* 24.3, pp. 847–871. MR: 2062922 (cit. on p. 2680).

– (2009). "Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging". *Mem. Amer. Math. Soc.* 201.944, pp. viii+129. MR: 2547839 (cit. on p. 2682).

C. Kipnis, C. Marchioro, and E. Presutti (1982). "Heat flow in an exactly solvable model". *J. Statist. Phys.* 27.1, pp. 65–74. MR: 656869 (cit. on p. 2666).

C. Kipnis and S. R. S. Varadhan (1986). "Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions". *Comm. Math. Phys.* 104.1, pp. 1–19. MR: 834478 (cit. on p. 2669).

Tomasz Komorowski and Stefano Olla (2016). "Ballistic and superdiffusive scales in the marcroscopic evolution of a chain of oscillators". *Nonlinearity* 29.3, pp. 962–999. MR: 3465990 (cit. on p. 2668).

– (2017). "Diffusive propagation of energy in a non-acoustic chain". *Arch. Ration. Mech. Anal.* 223.1, pp. 95–139. MR: 3590370 (cit. on p. 2668).

Oscar E. Lanford III, Joel L. Lebowitz, and Elliott H. Lieb (1977). "Time evolution of infinite anharmonic systems". *J. Statist. Phys.* 16.6, pp. 453–461. MR: 0452388 (cit. on p. 2667).

Raphaël Lefevere and Lorenzo Zambotti (2010). "Hot scatterers and tracers for the transfer of heat in collisional dynamics". *J. Stat. Phys.* 139.4, pp. 686–713. MR: 2638933 (cit. on p. 2672).

Carlangelo Liverani (2004). "On contact Anosov flows". *Ann. of Math. (2)* 159.3, pp. 1275–1312. MR: 2113022 (cit. on p. 2672).

Carlangelo Liverani and Stefano Olla (1996). "Ergodicity in infinite Hamiltonian systems with conservative noise". *Probab. Theory Related Fields* 106.3, pp. 401–445. MR: 1418847 (cit. on p. 2668).

– (2012). "Toward the Fourier law for a weakly interacting anharmonic crystal". *J. Amer. Math. Soc.* 25.2, pp. 555–583. MR: 2869027 (cit. on pp. 2668, 2674).

Carlangelo Liverani, Stefano Olla, and Makiko Sasada (n.d.). "Diffusive scaling in energy ginzburg-Landau dynamics". Preprint. arXiv: 1509.06116 (cit. on p. 2674).

Enzo Olivieri and Maria Eulália Vares (2005). *Large deviations and metastability*. Vol. 100. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, pp. xvi+512. MR: 2123364 (cit. on p. 2682).

S. Olla, S. R. S. Varadhan, and H.-T. Yau (1993). "Hydrodynamical limit for a Hamiltonian system with weak noise". *Comm. Math. Phys.* 155.3, pp. 523–560. MR: 1231642 (cit. on pp. 2666, 2668).

Stefano Olla and Makiko Sasada (2013). "Macroscopic energy diffusion for a chain of anharmonic oscillators". *Probab. Theory Related Fields* 157.3-4, pp. 721–775. MR: 3129802 (cit. on pp. 2668, 2674).

Z Rieder, JL Lebowitz, and E Lieb (1967). "Properties of a harmonic crystal in a stationary nonequilibrium state". *Journal of Mathematical Physics* 8.5, pp. 1073–1078 (cit. on p. 2667).

David Ruelle (2012). "A mechanical model for Fourier's law of heat conduction". *Comm. Math. Phys.* 311.3, pp. 755–768. MR: 2909762 (cit. on p. 2672).

Makiko Sasada (2015). "Spectral gap for stochastic energy exchange model with nonuniformly positive rate function". *Ann. Probab.* 43.4, pp. 1663–1711. MR: 3353812 (cit. on p. 2677).

Herbert Spohn (1991). *Large scale dynamics of interacting particles*. Texts and monographs in Physics. Springer Science & Business Media (cit. on p. 2666).

Daniel W. Stroock and S. R. Srinivasa Varadhan (2006). *Multidimensional diffusion processes*. Classics in Mathematics. Reprint of the 1997 edition. Springer-Verlag, Berlin, pp. xii+338. MR: 2190038 (cit. on p. 2673).

S. R. S. Varadhan (1993). "Nonlinear diffusion limit for a system with nearest neighbor interactions. II". In: *Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*. Vol. 283. Pitman Res. Notes Math. Ser. Longman Sci. Tech., Harlow, pp. 75–128. MR: 1354152 (cit. on pp. 2666, 2670).

A. D. Ventcel and M. I. Freidlin (1969). "Small random perturbations of a dynamical system with stable equilibrium position". *Dokl. Akad. Nauk SSSR* 187, pp. 506–509. MR: 0249795 (cit. on p. 2680).

Tatiana Yarmola (2014). "Sub-exponential mixing of open systems with particle-disk interactions". *J. Stat. Phys.* 156.3, pp. 473–492. MR: 3217533 (cit. on p. 2672).

Xenophon Zotos (2002). "Ballistic Transport in Classical and Quantum Integrable Systems". *Journal of Low Temperature Physics* 126 (3-4), pp. 1185–1194 (cit. on p. 2667).

Received 2017-12-06.

Carlangelo Liverani
Dipartimento di Matematica
II Università di Roma (Tor Vergata)
Via della Ricerca Scientifica
00133 Roma
Italy
liverani@mat.uniroma2.it

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2687–2704)

BOGOLIUBOV EXCITATION SPECTRUM FOR BOSE–EINSTEIN CONDENSATES

BENJAMIN SCHLEIN

Abstract

We consider interacting Bose gases trapped in a box $\Lambda = [0;1]^3$ in the Gross–Pitaevskii limit. Assuming the potential to be weak enough, we establish the validity of Bogoliubov's prediction for the ground state energy and the low-energy excitation spectrum. These notes are based on a joint work with C. Boccato, C. Brennecke and S. Cenatiempo.

1 Introduction

In the last two decades, since the first experimental realisations of Bose-Einstein condensates Anderson, Ensher, Matthews, Wieman, and Cornell [1995] and Davis, Mewes, Andrews, Van Druten, Durfee, Kurn, and Ketterle [1995], the study of bosonic systems at low temperature has been a very active field of research in physics (experimental and theoretical) and also in mathematics.

Trapped Bose gases observed in typical experiments are well described as quantum systems of N particles, interacting through a repulsive two-body potential with scattering length of the order N^{-1}; this asymptotic regime is commonly known as the Gross-Pitaevskii limit. If particles are confined in a box $\Lambda = [0;1]^3$ and we impose periodic boundary conditions, the Bose gas in the Gross-Pitaevskii regime is described by the Hamilton operator

$$H_N = \sum_{j=1}^{N} -\Delta_{x_j} + \kappa \sum_{i<j}^{N} N^2 V(N(x_i - x_j)) \,. \tag{1-1}$$

According to the bosonic statistics, (1-1) acts on the Hilbert space $L^2(\Lambda)^{\otimes_s N}$, the subspace of $L^2(\Lambda^N)$ consisting of all functions that are invariant with respect to permutations of the

The author gratefully acknowledges support from the NCCR SwissMAP and from the Swiss National Foundation of Science through the SNF Grant "Dynamical and energetic properties of Bose-Einstein condensates".
MSC2010: primary 81Q10; secondary 82B10, 35P05, 47A75.

N particles. In (1-1), V is assumed to be non-negative, spherically symmetric, compactly supported and sufficiently regular (in fact, the condition $V \in L^3(\mathbb{R}^3)$ will suffice) and $\kappa > 0$ is a coupling constant that will be later supposed to be small enough (but fixed, independent of N). We denote by a_0 the scattering length of κV, which is defined by the requirement that the solution of the zero-energy scattering equation

$$\left[-\Delta + \frac{\kappa}{2} V(x)\right] f(x) = 0 \tag{1-2}$$

with the boundary condition $f(x) \to 1$ as $|x| \to \infty$, is given by $f(x) = 1 - a_0/|x|$, for $|x|$ large enough (outside the range of V). Equivalently, we can determine the scattering length through

$$8\pi a_0 = \kappa \int V(x) f(x) dx \,. \tag{1-3}$$

It follows from the results of E. Lieb and Yngvason [1998] and of E. H. Lieb, Seiringer, and Yngvason [2000] that the ground state energy E_N of (1-1) is such that

$$\lim_{N\to\infty} \frac{E_N}{N} = 4\pi a_0 \,. \tag{1-4}$$

Furthermore, the work of E. H. Lieb and R. Seiringer [2002], recently revised also in Nam, Rougerie, and Seiringer [2016], implies that the ground state of (1-1) exhibits Bose-Einstein condensation. In other words, if $\psi_N \in L^2(\Lambda)^{\otimes_s N}$ is the normalized ground state of (1-1) and γ_N denotes the one-particle reduced density associated with ψ_N, which is defined as the non-negative trace class operator on $L^2(\Lambda)$ with the integral kernel

$$\gamma_N(x; y) = \int dx_2 \dots dx_N \, \psi_N(x, x_2, \dots, x_N) \, \overline{\psi_N}(y, x_2, \dots, x_N)$$

then, as $N \to \infty$,

$$\gamma_N \to |\varphi_0\rangle\langle\varphi_0| \tag{1-5}$$

where $\varphi_0(x) = 1$ for all $x \in \Lambda$ is the zero-momentum mode. The convergence (1-5) (which holds in any reasonable topology, for example with respect to the trace-class norm) means that, in the ground state of (1-1), all particles are described by φ_0, up to a fraction vanishing in the limit of large N. One should however stress the fact that (1-5) does not imply that the product state $\varphi_0^{\otimes N}$ is a good approximation for the ground state of (1-1). In fact, a simple computation shows that

$$\langle \varphi_0^{\otimes N}, H_N \varphi_0^{\otimes N} \rangle = \frac{(N-1)}{2} \kappa \widehat{V}(0) \tag{1-6}$$

which is not compatible with (1-4). The point, which will often recur in these notes, is that, because of the singular interaction, the ground state of (1-1) (and, in fact, all low-energy states, as we will see below) develops a short scale correlation structure, varying on the length scale N^{-1} (and therefore disappearing in the limit $N \to \infty$), which is responsible for lowering the energy from (1-6) to (1-4) (from (1-3) it is clear that $8\pi a_0 < \kappa \widehat{V}(0)$).

Equation (1-4) establishes E_N to leading order. Our goal in these notes is to obtain more precise information about the ground state energy and about low-energy excitations of (1-1), determining them up to errors that vanish in the limit $N \to \infty$.

Theorem 1.1. *Let $V \in L^3(\mathbb{R}^3)$ be non-negative, spherically symmetric and compactly supported. Let the coupling constant $\kappa > 0$ be small enough. Then we have*

$$\begin{aligned} E_N &= 4\pi(N-1)a_N \\ &\quad - \frac{1}{2}\sum_{p\in\Lambda_+^*}\left[p^2 + 8\pi a_0 - \sqrt{|p|^4 + 16\pi a_0 p^2} - \frac{(8\pi a_0)^2}{2p^2}\right] + \mathcal{O}(N^{-1/4}) \end{aligned} \tag{1-7}$$

where $\Lambda_+^* = 2\pi\mathbb{Z}^3 \backslash \{0\}$, and

$$\begin{aligned} 8\pi a_N &= \kappa\widehat{V}(0) + \\ &+ \sum_{k=1}^{\infty} \frac{(-1)^k \kappa^{k+1}}{(2N)^k} \sum_{p_1,\dots,p_k\in\Lambda_+^*} \frac{\widehat{V}(p_1/N)}{p_1^2} \left(\prod_{i=1}^{k-1} \frac{\widehat{V}((p_i - p_{i+1})/N)}{p_{i+1}^2}\right) \widehat{V}(p_k/N) . \end{aligned} \tag{1-8}$$

Furthermore, the spectrum of $H_N - E_N$ below a threshold $\zeta > 0$ consists of eigenvalues given, in the limit $N \to \infty$, by

$$\sum_{p\in\Lambda_+^*} n_p \sqrt{|p|^4 + 16\pi a_0 p^2} + \mathcal{O}(N^{-1/4}(1+\zeta^3)) . \tag{1-9}$$

Here $n_p \in \mathbb{N}$ for all $p \in \Lambda_+^*$ and $n_p \neq 0$ for finitely many $p \in \Lambda_+^*$ only.

Remarks:

1) Taylor expanding the square root to third order, it is easy to check that the sum in (1-7) is absolutely convergent and therefore that it gives a contribution of order one to the ground state energy E_N.

2) The k-th term in (1-8) is bounded by $C^k\kappa^k$, for some constant $C > 0$. Hence, the series is absolutely convergent and bounded uniformly in N, if $\kappa > 0$ is small enough.

3) The expression (1-8) for $8\pi a_N$ can be compared with the Born series for the unscaled scattering length a_0, given by

$$8\pi a_0 = \kappa \widehat{V}(0) + \sum_{k=2}^{\infty} \frac{(-1)^k \kappa^{k+1}}{2^k (2\pi)^{3k}} \int_{\mathbb{R}^{3k}} dp_1 \dots dp_k \, \frac{\widehat{V}(p_1)}{p_1^2} \left(\prod_{i=1}^{k-1} \frac{\widehat{V}(p_i - p_{i+1})}{p_{i+1}^2} \right) \widehat{V}(p_k). \tag{1-10}$$

In particular, it is possible to show that the difference $4\pi(a_N - a_0)N$ remains bounded, of order one, in the limit of large N. Notice, however, that it does not seem to tend to zero; this means that, in (1-7), we cannot replace a_N with a_0. In other words, at the level of precision of (1-7), the ground state energy is sensitive to the finite size of the box and it cannot be simply expressed in terms of the infinite volume scattering length a_0.

The results of Theorem 1.1, in particular the expression (1-9) for the excitation spectrum, have already been predicted by Bogolubov [1947]. In his work, Bogoliubov rewrote the Hamilton operator (1-1) as

$$H_N = \sum_{p \in \Lambda^*} p^2 a_p^* a_p + \frac{\kappa}{2N} \sum_{p,q,r \in \Lambda^*} \widehat{V}(r/N) a_{p+r}^* a_q^* a_p a_{q+r} \tag{1-11}$$

using the formalism of second quantization. For every momentum $p \in \Lambda^* = 2\pi\mathbb{Z}^3$, a_p^*, a_p are the usual creation and annihilation operators defined on the bosonic Fock space $\mathcal{F} = \bigoplus_{n \geq 0} L^2(\Lambda)^{\otimes_s n}$ and satisfying canonical commutation relations $[a_p, a_q^*] = \delta_{pq}$, $[a_p, a_q] = [a_p^*, a_q^*] = 0$. Since low-energy states exhibit Bose-Einstein condensation, we expect the operator $a_0^* a_0$ measuring the number of particles in the zero-momentum state φ_0 to be of the order N and therefore much larger than the commutator $[a_0, a_0^*] = 1$. Motivated by this observation, Bogoliubov decided to replace, in (1-11), all creation and annihilation operators a_0^*, a_0 by $\sqrt{N}$ and then to neglect all resulting terms with more than two creation and annihilation operators associated with momenta different than zero. With this approximation, Bogoliubov derived an Hamilton operator quadratic in creation and annihilation operators a_p^*, a_p with $p \neq 0$ that he could diagonalize explicitly. Finally he argued, following a hint of Landau, that certain expressions that appeared in his formulas for the ground state energy and for the excitation spectrum were just first and second order Born approximations of the scattering length, and thus he replaced them with a_0; with this final substitution he obtained essentially results equivalent to those stated in Theorem 1.1.

From the point of view of mathematical physics, the validity of the Bogoliubov approximation has been first established by E. H. Lieb and Solovej [2002] in the computation of the ground state energy of the one-component charged Bose gas. It was then proved by

Giuliani and Seiringer [2009] in their derivation of the Lee-Huang-Yang formula for the ground state energy of a Bose gas in a combined weak coupling and high density regime, and by Seiringer [2011], Grech and Seiringer [2013], Lewin, Nam, Serfaty, and Solovej [2015], Dereziński and Napiórkowski [2014], Pizzo [n.d.] in their analysis of the low-energy spectrum of Bose gases in the mean field limit. More recently, the validity of Bogoliubov prediction was established in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(c)] for systems of N bosons interacting through singular potential, described by the Hamiltonian (written like (1-11) in second quantized form)

$$H_N^\beta = \sum_{p \in \Lambda^*} p^2 a_p^* a_p + \frac{\kappa}{2N} \sum_{p,q,r \in \Lambda^*} \widehat{V}(r/N^\beta) a_{p+r}^* a_q^* a_p a_{q+r} \tag{1-12}$$

for a parameter $\beta \in (0;1)$. Notice that (1-12) interpolates between the mean-field regime that is recovered for $\beta = 0$ and the Gross-Pitaevskii limit, which corresponds to $\beta = 1$.

In the rest of these notes, we are going to sketch the main ideas going into the proof of Theorem 1.1; for more details, see Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(a)].

2 Excitation Hamiltonian

The first step in the proof of Theorem 1.1 consists in factoring out the condensate to focus on its orthogonal excitations. We use here an idea from Lewin, Nam, Serfaty, and Solovej [2015]. Since we expect that, for low-energy states, most particles occupy the zero-momentum mode $\varphi_0(x) = 1$ for all $x \in \Lambda$ (for the ground state, this follows from (1-5)), we write an arbitrary N-particle wave function $\psi \in L^2(\Lambda)^{\otimes_s N}$ as

$$\psi_N = \alpha_0 \, \varphi_0^{\otimes N} + \alpha_1 \otimes_s \varphi_0^{\otimes (N-1)} + \alpha_2 \otimes_s \varphi_0^{\otimes (N-2)} + \cdots + \alpha_N$$

where $\alpha_j \in L^2_\perp(\Lambda)^{\otimes_s j}$, for all $j = 0, 1, \ldots, N$. Here, $L^2_\perp(\Lambda)$ denotes the orthogonal complement of the one-dimensional subspace spanned by φ_0 in $L^2(\Lambda)$. It is easy to check that the choice of $\alpha_0, \ldots, \alpha_N$ is unique, and that $\sum_{j=0}^N \|\alpha_j\|^2 = \|\psi_N\|^2$. Hence, with the notation

$$\mathcal{F}_+^{\leq N} = \bigoplus_{j=0}^{N} L^2_\perp(\Lambda)^{\otimes_s j}$$

for the truncated Fock space constructed over $L^2_\perp(\Lambda)$, we can define a unitary map

$$\begin{aligned} U_N : \; L^2(\Lambda)^{\otimes_s N} &\to \mathcal{F}_+^{\leq N} \\ \psi_N &\to \{\alpha_0, \ldots, \alpha_N\}. \end{aligned}$$

The map U_N allows us to focus on the orthogonal excitations of the condensate that are described in the Hilbert space $\mathcal{F}_+^{\leq N}$. Conjugating the Hamiltonian (1-1) with the unitary map U_N, we define an excitation Hamiltonian $\mathcal{L}_N = U_N H_N U_N^* : \mathcal{F}_+^{\leq N} \to \mathcal{F}_+^{\leq N}$. With the notation $\mathcal{N}_+$ for the number of particles operator on $\mathcal{F}_+^{\leq N}$, we find

$$\begin{aligned} U_N a_p^* a_q U_N^* &= a_p^* a_q \, , \\ U_N a_p^* a_0 U_N^* &= a_p^* \sqrt{N - \mathcal{N}_+} \, , \\ U_N a_0^* a_p U_N^* &= \sqrt{N - \mathcal{N}_+} a_p \, , \\ U_N a_0^* a_0 U_N^* &= (N - \mathcal{N}_+) \, . \end{aligned} \tag{2-1}$$

Applying these rules to the Hamiltonian (1-1) written in second quantized form as in (1-11), we arrive at

$$\mathcal{L}_N = \mathcal{L}_N^{(0)} + \mathcal{L}_N^{(2)} + \mathcal{L}_N^{(3)} + \mathcal{L}_N^{(4)}, \tag{2-2}$$

with (recall the notation $\Lambda_+^* = 2\pi \mathbb{Z}^3 \backslash \{0\}$)

$$\begin{aligned} \mathcal{L}_N^{(0)} &= \frac{N-1}{2N} \kappa \widehat{V}(0)(N - \mathcal{N}_+) + \frac{\kappa \widehat{V}(0)}{2N} \mathcal{N}_+ (N - \mathcal{N}_+) \, , \\ \mathcal{L}_N^{(2)} &= \sum_{p \in \Lambda_+^*} p^2 a_p^* a_p + \sum_{p \in \Lambda_+^*} \kappa \widehat{V}(p/N) a_p^* a_p \left(\frac{N - \mathcal{N}_+}{N} \right) \\ &\quad + \frac{\kappa}{2} \sum_{p \in \Lambda_+^*} \widehat{V}(p/N) \left[a_p^* a_{-p}^* \sqrt{\frac{N - 1 - \mathcal{N}_+}{N} \frac{N - \mathcal{N}_+}{N}} + \text{h.c.} \right] , \\ \mathcal{L}_N^{(3)} &= \frac{\kappa}{\sqrt{N}} \sum_{p,q \in \Lambda_+^* : p+q \neq 0} \widehat{V}(p/N) \left[a_{p+q}^* a_{-p}^* a_q \sqrt{\frac{N - \mathcal{N}_+}{N}} + \text{h.c.} \right] , \\ \mathcal{L}_N^{(4)} &= \frac{\kappa}{2N} \sum_{p,q \in \Lambda_+^*, r \in \Lambda^* : r \neq -p,-q} \widehat{V}(r/N) a_{p+r}^* a_q^* a_p a_{q+r} \, , \end{aligned} \tag{2-3}$$

where h.c. indicates the hermitian conjugate operator and where, in the notation $\mathcal{L}_N^{(j)}$, the label $j \in \{0, 2, 3, 4\}$ refers to the number of creation and annihilation operators.

Conjugation with U_N extracts contributions from the quartic interaction in (1-11) and moves them into the constant and the quadratic parts $\mathcal{L}_N^{(0)}$ and $\mathcal{L}_N^{(2)}$ of the excitation Hamiltonian. In the mean-field case considered in Seiringer [2011], Grech and Seiringer [2013], Lewin, Nam, Serfaty, and Solovej [2015], Dereziński and Napiórkowski [2014], and Pizzo [n.d.] (corresponding to the (1-12) with $\beta = 0$), one can show that, after application of U_N, the cubic and quartic terms $\mathcal{L}_N^{(3)}$ and $\mathcal{L}_N^{(4)}$ are negligible on low-energy states, in the limit

$N \to \infty$. In this case, the low-lying excitation spectrum can therefore be determined diagonalizing the quadratic operator $\mathcal{L}_N^{(2)}$. This is not the case in the Gross-Pitaevskii regime considered here. Applying the unitary map U_N we factor out the condensate but we do not remove the short scale correlation structure which, as explained after (1-6), still carries an energy of order N. As a consequence, in the Gross-Pitaevskii regime, cubic and quartic terms in $\mathcal{L}_N$ are not negligible on low-energy states.

Notice that conjugation with U_N can be interpreted as a rigorous version of the substitution proposed by Bogoliubov of all creation and annihilation operators a_0^*, a_0 associated with zero momentum with factors of $\sqrt{N}$. The fact that $\mathcal{L}_N^{(3)}$ and $\mathcal{L}_N^{(4)}$ are not negligible means, therefore, that in the Gross-Pitaevskii regime the Bogoliubov approximation cannot be justified. But then, why did Bogoliubov obtained the correct expressions for the low-energy spectrum, the same expressions appearing in Theorem 1.1? The point is that, when at the end of his computation Bogoliubov replaced, following the hint of Landau, first and second Born approximations with the full scattering length a_0, he exactly made up for the (non-negligible) contributions that are hidden in $\mathcal{L}_N^{(3)}$ and $\mathcal{L}_N^{(4)}$ and that he neglected with his approximation.

It is clear that to obtain a rigorous proof of Theorem 1.1 we cannot neglect cubic and quartic parts of the excitation Hamiltionian. Instead, to extract the important contributions from $\mathcal{L}_N^{(3)}$ and $\mathcal{L}^{(4)}$, we need to conjugate $\mathcal{L}_N$ with another unitary map, a map that implements correlations among particles.

3 Generalized Bogoliubov Transformations

A strategy to implement correlations has been introduced in Benedikter, de Oliveira, and Schlein [2015], a paper devoted to the study of the dynamics in the Gross-Pitaevskii regime, for approximately coherent initial data in the bosonic Fock space. In that paper, correlations were produced by unitary conjugation with a Bogoliubov transformation of the form

$$\widetilde{T}(\eta) = \exp\left[\frac{1}{2}\sum_{p\in\Lambda_+^*} \eta_p (a_p^* a_{-p}^* - a_p a_{-p})\right] \tag{3-1}$$

for an appropriate real function $\eta \in \ell^2(\Lambda_+^*)$ (in fact, in Benedikter, de Oliveira, and Schlein [ibid.] the problem is not translation invariant and therefore slightly more complicated transformations were considered). Bogoliubov transformations are very convenient because their action on creation and annihilation operators is explicitly given by

$$\widetilde{T}^*(\eta)\, a_q\, \widetilde{T}(\eta) = \cosh(\eta_q)\, a_q + \sinh(\eta_q)\, a_{-q}^* \,. \tag{3-2}$$

Unfortunately, Bogoliubov transformations of the form (3-1) do not preserve the number of particles and therefore they do not leave the excitation Hilbert space $\mathcal{F}_+^{\leq N}$ invariant. To solve this problem, we follow Brennecke and Schlein [n.d.] and we introduce, on $\mathcal{F}_+^{\leq N}$, modified creation and annihilation operators defined, for any $p \in \Lambda_+^*$, by

$$b_p^* = a_p^* \sqrt{\frac{N - \mathcal{N}_+}{N}}, \qquad \text{and} \qquad b_p = \sqrt{\frac{N - \mathcal{N}_+}{N}}\, a_p \,.$$

Observing that, from (2-1),

$$U_N^* \, b_p^* \, U_N = a_p^* \frac{a_0}{\sqrt{N}}, \qquad U_N^* \, b_p \, U_N = \frac{a_0^*}{\sqrt{N}} a_p \,, \tag{3-3}$$

we conclude that the modified creation operator b_p^* creates a particle with momentum p and, at the same time, it annihilates a particle from the condensate (i.e. a particle with momentum $p = 0$) while b_p annihilate a particle with momentum p and creates a particle in the condensate. In other words, b_p^* creates and b_p annihilates an excitation with momentum p, preserving however the total number of particles. This is the reason why modified creation and annihilation operators leave the excitation Hilbert space $\mathcal{F}_+^{\leq N}$ invariant, in contrast with the standard creation and annihilation operators.

Using the modified field operators we can now introduce generalized Bogoliubov transformations by defining, in analogy to (3-1),

$$T(\eta) = \exp\left[\frac{1}{2} \sum_{p \in \Lambda_+^*} \eta_p (b_p^* b_{-p}^* - b_p b_{-p})\right]. \tag{3-4}$$

By construction, $T(\eta) : \mathcal{F}_+^{\leq N} \to \mathcal{F}_+^{\leq N}$. The price we have to pay for replacing the original Bogoliubov transformations (3-1) with their generalization (3-4) is the fact that there is no explicit formula like (3-2) describing the action of $T(\eta)$ on creation and annihilation operators (because modified creation and annihilation operators do not satisfy canonical commutation relations). Still, when we consider states exhibiting Bose-Einstein condensation where $a_0, a_0^* \simeq \sqrt{N}$, we may expect from (3-3) that $b_p \simeq a_p$ and $b_p^* \simeq a_p^*$ and therefore that (3-2) is approximately correct, even if we replace $\widetilde{T}(\eta)$ by $T(\eta)$. It is possible to quantify this last statement through the introduction of remainder operators. For $p \in \Lambda_+^*$, we define d_p, d_p^* by

$$\begin{aligned} T^*(\eta)\, b_p \, T(\eta) &= \cosh(\eta_p)\, b_p + \sinh(\eta_p)\, b_{-p}^* + d_p, \\ T(\eta)\, b_p^* \, T(\eta) &= \cosh(\eta_p)\, b_p^* + \sinh(\eta_p)\, b_{-p} + d_p^* . \end{aligned} \tag{3-5}$$

Then it is possible to prove that, if $\eta \in \ell^2(\Lambda_+^*)$ with $\|\eta\|_2$ small enough,

$$\begin{aligned} \|d_p^* \xi\| &\leq \frac{C}{N} \|(\mathcal{N}_+ + 1)^{3/2} \xi\|, \\ \|d_p \xi\| &\leq \frac{C}{N} \|(\mathcal{N}_+ + 1)^{3/2} \xi\|, \end{aligned} \tag{3-6}$$

for all $\xi \in \mathcal{F}_+^{\leq N}$. On states exhibiting condensation, the operator $\mathcal{N}_+$ is small; in this case (3-6) can be use to show that the remainder operators d_p, d_p^* are small (we gain a factor N^{-1}). The bounds (3-6) (and some more refined version) are discussed in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(a), Section 7]; their proof is based on Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(b), Lemma 2.5] which is a translation to momentum space of Brennecke and Schlein [n.d., Lemma 3.2].

To implement correlations, the choice of the coefficients η_p in (3-4) must be related with the solution of the zero-energy scattering equation (1-2). More precisely, since we are working on the finite box $\Lambda = [0;1]^3$, we consider the Neumann problem

$$\left[-\Delta + \frac{\kappa}{2} V(x) \right] f_\ell(x) = \lambda_\ell f_\ell(x) \tag{3-7}$$

on the ball $|x| \leq N\ell$ with the normalization $f_\ell(x) = 1$ on the boundary $|x| = N\ell$. We find that the smallest Neumann eigenvalue λ_ℓ is such that

$$\lambda_\ell = \frac{3a_0}{N^3 \ell^3} \left[1 + \mathcal{O}\left(\frac{a_0}{N\ell} \right) \right]$$

and that $f_\ell = 1 - w_\ell$, where

$$0 \leq w_\ell(x) \leq \frac{C\kappa}{|x|+1}, \qquad |\nabla w_\ell(x)| \leq \frac{C\kappa}{|x|^2+1} \tag{3-8}$$

for all $|x| \leq N\ell$ (this confirms the intuition that f_ℓ is a small modification of the solution of the zero energy scattering equation (1-2) which is given, for large $|x|$, by $1 - a_0/|x|$). By scaling, we find that

$$\left[-\Delta + \frac{\kappa N^2}{2} V(N.) \right] f_\ell(N.) = \lambda_\ell N^2 f_\ell(N.)$$

on the ball $|x| \leq \ell$. Fixing $\ell < 1/2$ (independently of N), we can extend $f_\ell(Nx) = 1$ and also $w_\ell(Nx) = 1 - f_\ell(Nx) = 0$ for all $x \in \Lambda$ with $|x| > \ell$. Hence, the maps $x \to w_\ell(Nx)$ and $x \to f_\ell(Nx)$ can be expressed as Fourier series with coefficients $N^{-3}\widehat{w}_\ell(p/N)$ and, respectively, $\delta_{p,0} - N^{-3}\widehat{w}_\ell(p/N)$, for all $p \in \Lambda^*$. Here

$$\widehat{w}_\ell(z) = \frac{1}{(2\pi)^3} \int dx\, e^{-ix\cdot z} w_\ell(x)$$

is the Fourier transform of w_ℓ (as a compactly supported function on $\mathbb{R}^3$). For $p \in \Lambda^*$, we define

$$\eta_p = -\frac{1}{N^2}\widehat{w}_\ell(p/N)\,. \tag{3-9}$$

From (3-8), it is easy to check that

$$|\eta_p| \le C\kappa|p|^{-2} \tag{3-10}$$

for all $p \in \Lambda_+^*$. It follows that $\eta \in \ell^2(\Lambda_+^*)$, with $\|\eta\|_2 \le C$, uniformly in N. On the other hand, it is important to notice that (3-10) does not provide enough decay in momentum to estimate the H^1-norm of η. Since the decay of $\widehat{w}_\ell(p/N)$ kicks in for $|p| \gtrsim N$, we obtain that

$$\|\eta\|_{H^1}^2 = \sum_{p\in\Lambda_+^*}(1+p^2)|\eta_p|^2 \simeq CN\,. \tag{3-11}$$

From (3-5) and using the notation $\mathcal{K} = \sum_{p\in\Lambda_+^*} p^2 a_p^* a_p$ for the kinetic energy operator, it is easy to check that

$$\begin{aligned} T^*(\eta)\,\mathcal{N}_+\,T(\eta) &\simeq \mathcal{N}_+ + \|\eta\|_2^2\,,\\ T^*(\eta)\,\mathcal{K}\,T(\eta) &\simeq \mathcal{K} + \|\eta\|_{H^1}^2\,. \end{aligned} \tag{3-12}$$

The uniform bound for $\|\eta\|_2$ and the estimate (3-11) for the H^1-norm of η imply, therefore, that conjugation with $T(\eta)$ only creates finitely many excitations of the condensates, but also that these excitations carry a macroscopic energy, of order N (in (3-12) we only consider the change of the kinetic energy but also the change of the potential energy is of comparable size, leading to a net gain of order N). One can hope, therefore, that conjugating with $T(\eta)$ we can preserve condensation and, at the same time, decrease the energy to make up for the difference between (1-6) and the true ground state energy (1-7).

4 Renormalized Excitation Hamiltonian

We introduce the renormalized excitation Hamiltonian

$$\mathcal{G}_N = T^*(\eta)\mathcal{L}_N T(\eta) = T^*(\eta)U_N H_N U_N^* T(\eta) : \mathcal{F}_+^{\le N} \to \mathcal{F}_+^{\le N} \tag{4-1}$$

with η defined as in (3-9). The next proposition was proven in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(b)].

Proposition 4.1. *Let $V \in L^3(\mathbb{R}^3)$ be non-negative, spherically symmetric and compactly supported. Let the coupling constant $\kappa > 0$ be small enough. Let $\mathcal{G}_N$ be defined as in (4-1). Then we can write*

$$\mathcal{G}_N = 4\pi a_0 N + \mathcal{H}_N + \delta_{\mathcal{G}_N} \tag{4-2}$$

where

$$\mathcal{H}_N = \sum_{p\in\Lambda_+^*} p^2 a_p^* a_p + \frac{\kappa}{2N} \sum_{\substack{p,q\in\Lambda_+^*, r\in\Lambda^*:\\ r\neq -p,-q}} \widehat{V}(r/N) a_{p+r}^* a_q^* a_p a_{q+r}$$

is the restriction of (1-11) to $\mathcal{F}_+^{\leq N}$ and where the remainder operator $\delta_{\mathcal{G}_N}$ is such that, for all $\alpha > 0$ there exists $C > 0$ with

$$\pm\, \delta_{\mathcal{G}_N} \leq \alpha \mathcal{H}_N + C\kappa(\mathcal{N}_+ + 1) \tag{4-3}$$

as an operator inequality on $\mathcal{F}_+^{\leq N}$.

To prove (4-2) we apply (3-5) to the operators $\mathcal{L}_N^{(j)}$, $j = 0, 2, 3, 4$ in (2-3). It is clear that we will generate terms that are not normally ordered (for this simplified discussion, ignore the remainders d_p, d_p^*). To restore normal order, we generate terms of lower order in creation and annihilation operators. Specifically, conjugating $\mathcal{L}_N^{(2)}$ we generate new constant terms while conjugating $\mathcal{L}_N^{(4)}$ we generate new quadratic and new constant contributions. The choice (3-9) of η guarantees that, on the one hand, the combination of old and new constant terms reproduces, up to an error of order one, the correct ground state energy (1-7) and, on the other hand, that there is a cancellation among quadratic terms that allows us to bound everything in terms of $\mathcal{H}_N$ and $\mathcal{N}_+$ (as indicated in (4-3)).

Noticing that, on $\mathcal{F}_+^{\leq N}$, the kinetic energy operator $\mathcal{K}$ is gapped, we find $\mathcal{N}_+ \leq C\mathcal{K} \leq C\mathcal{H}_N$. The bound (4-3) implies therefore that, if $\kappa > 0$ is small enough,

$$C\mathcal{N}_+ - C \leq \frac{1}{2}\mathcal{H}_N - C \leq \mathcal{G}_N - 4\pi a_0 N \leq C(\mathcal{H}_N + 1)\,. \tag{4-4}$$

Hence, if the N-particle wave function $\psi_N \in L^2(\Lambda)^{\otimes_s N}$ is such that $\langle \psi_N, H_N \psi_N\rangle \leq 4\pi a_0 N + \zeta$, then we can write $\psi_N = U_N^* T(\eta)\xi_N$, where the excitation vector $\xi_N = T^*(\eta) U_N \psi_N \in \mathcal{F}_+^{\leq N}$ is such that

$$\langle \xi_N, \mathcal{N}_+ \xi_N\rangle \leq C\langle \xi_N, \mathcal{H}_N \xi_N\rangle \leq C(\zeta + 1)\,. \tag{4-5}$$

It is interesting to remark that (4-5) implies Bose-Einstein condensation in the sense of (1-5), since

(4-6)
$$\begin{aligned}1-\langle\varphi_0,\gamma_N\varphi_0\rangle &= 1-\frac{1}{N}\langle\psi_N,a^*(\varphi_0)a(\varphi_0)\psi_N\rangle = \frac{1}{N}\langle U_N\psi_N,\mathcal{N}_+U_N\psi_N\rangle\\ &= \frac{1}{N}\langle\xi_N,T^*(\eta)\mathcal{N}_+T(\eta)\xi_N\rangle \le \frac{C}{N}\langle\xi_N,(\mathcal{N}_++1)\xi_N\rangle \le \frac{C(\zeta+1)}{N}\end{aligned}$$

where we used the rules (2-1) and the bounds (3-12) and (4-5) (it is then easy to check that (4-6) implies $\gamma_N \to |\varphi_0\rangle\langle\varphi_0|$ first of all in the Hilbert-Schmidt topology but then also with respect to the trace norm). Eq. (4-6) improves (1-5) (in the case of small κ) by giving a precise and optimal bound on the rate of the convergence of the one-particle density matrix.

We can derive stronger bounds on the excitation vector ξ_N associated with a normalized N-particle wave function $\psi_N \in L^2(\Lambda)^{\otimes_s N}$ if, instead of imposing the condition $\langle\psi_N,H_N\psi_N\rangle \le 4\pi a_0 N+\zeta$, we require ψ_N to belong to the spectral subspace of H_N associated with energies below $4\pi a_0 N+\zeta$. The proof of the next lemma can be found in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(a), Section 4].

Lemma 4.2. *Let $V \in L^3(\mathbb{R}^3)$ be non-negative, spherically symmetric and compactly supported. Let the coupling constant $\kappa > 0$ be small enough. Let $\psi_N \in L^2(\Lambda)^{\otimes_s N}$ be normalized and such that $\psi_N = \mathbf{1}_{(-\infty;E_N+\zeta]}(H_N)\psi_N$ where $\mathbf{1}_I$ indicates the characteristic function of the interval $I \subset \mathbb{R}$. Then $\psi_N = U_N^*T(\eta)\xi_N$, where the excitation vector $\xi_N = T^*(\eta)U_N\psi_N \in \mathcal{F}_+^{\le N}$ is such that*

$$\big\langle\xi_N,\big[(\mathcal{N}_++1)^3+(\mathcal{N}_++1)(\mathcal{H}_N+1)\big]\xi_N\big\rangle \le C(1+\zeta^3) \tag{4-7}$$

uniformly in N.

With Lemma 4.2 we can go back to the renormalized excitation Hamiltonian and we can show that several terms contributing to $\mathcal{G}_N$ are negligible, in the limit of large N, on low-energy states. The result is the next proposition, whose proof is given in Boccato, Brennecke, Cenatiempo, and Schlein [ibid., Section 7].

Proposition 4.3. *Let $V \in L^3(\mathbb{R}^3)$ be non-negative, spherically symmetric and compactly supported. Let the coupling constant $\kappa > 0$ be small enough. Let $\mathcal{G}_N$ be defined as in (4-1). Then we can write*

$$\mathcal{G}_N = C_{\mathcal{G}_N}+\mathcal{Q}_{\mathcal{G}_N}+\mathcal{C}_N+\mathcal{V}_N+\mathcal{E}_{\mathcal{G}_N} \tag{4-8}$$

where, using the notation $\sigma_p = \sinh(\eta_p)$ *and* $\gamma_p = \cosh(\eta_p)$*,*

$$\mathcal{C}_N = \frac{\kappa}{\sqrt{N}} \sum_{\substack{p,q\in\Lambda_+^*:\\ q\neq -p}} \widehat{V}(p/N)\left[b^*_{p+q}b^*_{-p}(\gamma_q b_q + \sigma_q b^*_{-q}) + h.c.\right],$$

(4-9)

$$\mathcal{V}_N = \frac{\kappa}{2N} \sum_{\substack{p,q\in\Lambda_+^*, r\in\Lambda^*:\\ r\neq -p,-q}} \widehat{V}(r/N)\, a^*_{p+r}a^*_q a_p a_{q+r}\,,$$

and where

$$\begin{aligned} C_{\mathcal{G}_N} = \frac{(N-1)}{2}\kappa\widehat{V}(0) &+ \sum_{p\in\Lambda_+^*} p^2\sigma_p^2 + \kappa\widehat{V}(p/N)\left(\sigma_p\gamma_p + \sigma_p^2\right) \\ &+ \frac{\kappa}{2N}\sum_{p,q\in\Lambda_+^*}\widehat{V}((p-q)/N)\sigma_q\gamma_q\sigma_p\gamma_p \\ &+ \frac{1}{N}\sum_{p\in\Lambda^*}\left[p^2\eta_p^2 + \frac{\kappa}{2N}\left(\widehat{V}(\cdot/N)*\eta\right)_p\eta_p\right] \\ &- \frac{1}{N}\sum_{q\in\Lambda^*}\kappa\widehat{V}(q/N)\eta_q\sum_{p\in\Lambda_+^*}\sigma_p^2 \end{aligned}$$

and

$$\mathcal{Q}_{\mathcal{G}_N} = \sum_{p\in\Lambda_+^*}\Phi_p\, b_p^* b_p + \frac{1}{2}\sum_{p\in\Lambda_+^*}\Gamma_p\left(b_p^*b_{-p}^* + b_p^* b_{-p}^*\right)$$

with

$$\begin{aligned} \Phi_p =\ & (\sigma_p^2+\gamma_p^2)\,p^2 + \kappa\widehat{V}(p/N)\,(\gamma_p+\sigma_p)^2 + \frac{2\kappa}{N}\gamma_p\sigma_p\sum_{q\in\Lambda^*}\widehat{V}((p-q)/N)\eta_q \\ & - (\gamma_p^2+\sigma_p^2)\frac{\kappa}{N}\sum_{q\in\Lambda^*}\widehat{V}(q/N)\eta_q\,, \\ \Gamma_p =\ & 2p^2\sigma_p\gamma_p + \kappa\widehat{V}(p/N)(\gamma_p+\sigma_p)^2 + (\gamma_p^2+\sigma_p^2)\frac{\kappa}{N}\sum_{q\in\Lambda^*}\widehat{V}((p-q)/N)\eta_q \\ & - 2\gamma_p\sigma_p\frac{\kappa}{N}\sum_{q\in\Lambda^*}\widehat{V}(q/N)\eta_q\,. \end{aligned}$$

Moreover, we have

(4-10)
$$\pm\,\mathcal{E}_{\mathcal{G}_N} \leq \frac{C}{N^{1/4}}\left[(\mathcal{N}_+ + 1)^3 + (\mathcal{N}_+ + 1)(\mathcal{H}_N + 1)\right].$$

On the r.h.s. of (4-8) we have a constant and a quadratic term that can be easily diagonalized by means of a generalized Bogoliubov transformation. From (4-10) it follows that the error term $\mathcal{E}_{\mathcal{G}_N}$ is negligible on low-energy states. There are, however, still two terms, the cubic term $\mathcal{C}_N$ and the quartic term $\mathcal{V}_N$ in (4-9), whose contribution to the spectrum cannot be easily determined and that are not negligible. This is the main difference between the Gross-Pitaevskii regime that we are considering here and regimes described by the Hamilton operator (1-12), with parameter $0 < \beta < 1$, that were considered in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(c)]. For $\beta < 1$, the expectation for example of the quartic interaction can be bounded by

$$\begin{aligned}\langle \xi, \mathcal{V}_N \xi \rangle &\leq \frac{\kappa}{2N} \sum_{p,q,r \in \Lambda_+^*} |\widehat{V}(r/N^\beta)| \, \|a_{p+r} a_q \xi\| \, \|a_{q+r} a_p \xi\| \\ &\leq \frac{C}{N} \sum_{p,q,r \in \Lambda_+^*} \frac{|\widehat{V}(r/N^\beta)|}{(q+r)^2} (p+r)^2 \|a_{p+r} a_q \xi\|^2 \leq C N^{\beta - 1} \langle \xi, \mathcal{N}_+ \mathcal{K} \xi \rangle\end{aligned}$$

where we used the estimate

$$\sup_{q \in \Lambda_+^*} \sum_{r \in \Lambda_+^*} \frac{|\widehat{V}(r/N^\beta)|}{(q+r)^2} \leq C N^\beta \, .$$

Hence, for all $\beta < 1$, the quartic and, similarly, also the cubic terms on the r.h.s. of (4-8) are negligible in the limit $N \to \infty$ and can be included in the error term $\mathcal{E}_{\mathcal{G}_N}$. This means that, for $\beta < 1$, we can read off the spectrum of $\mathcal{G}_N$ (and therefore, of the initial Hamiltonian H_N), diagonalizing the quadratic operator on the r.h.s. of (4-8). This is not the case for $\beta = 1$.

5 Cubic Conjugation

It is not surprising that there are still important contributions hidden in the cubic and quartic terms on the r.h.s. of (4-8). Already from Erdős, Schlein, and Yau [2008] and, more recently, from Napiórkowski, Reuvers, and Solovej [n.d.], it follows that Bogoliubov states, i.e. in our setting states of the form $U_N^* T(\mu)\Omega$ for some $\mu \in \ell^2(\Lambda_+^*)$, can only approximate the ground state energy up to an error of order one, even after optimizing the choice of the function μ. To go beyond this resolution, we need to conjugate $\mathcal{G}_N$ with a more complicated unitary operator. Since Bogoliubov transformations are the exponential of quadratic expressions in creation and annihilation operators, the natural guess is to use the exponential of a antisymmetric cubic phase. In fact, a similar approach was introduced by Yau and Yin [2009] to obtain a precise upper bound for the ground state energy of a dilute

Bose gas in the thermodynamic limit, correct up to second order, in agreement with the Lee-Huang-Yang formula. In our setting, we consider the operator

$$S(\eta) = e^{A(\eta)} = \exp\left(\frac{1}{\sqrt{N}}\sum_{r\in P_H, v\in P_L}\eta_r\left[b^*_{r+v}b^*_{-r}(\gamma_v b_v + \sigma_v b^*_{-v}) - \text{h.c.}\right]\right) \tag{5-1}$$

where, as above $\sigma_v = \sinh(\eta_p), \gamma_v = \cosh(\eta_p)$, and where we used the notation $P_H = \{p \in \Lambda^*_+ : |p| > \sqrt{N}\}$ and $P_L = \{p \in \Lambda^*_+ : |p| \le \sqrt{N}\}$. Here, $\eta \in \ell^2(\Lambda^*_+)$ is the same function defined in (3-9) entering the definition of the Bogoliubov transformation $T(\eta)$. With the operator (5-1), we can define a new, twice renormalized, excitation Hamiltonian

$$\mathcal{J}_N = S^*(\eta)\mathcal{G}_N S(\eta) = S^*(\eta)T^*(\eta)U_N H_N U_N^* T(\eta) S(\eta) : \mathcal{F}_+^{\le N} \to \mathcal{F}_+^{\le N}\,. \tag{5-2}$$

To study the operator $\mathcal{J}_N$, we start from the decomposition (4-8) of $\mathcal{G}_N$ and we analyze how conjugation with $S(\eta)$ acts on the different terms. The first remark is that, when we conjugate with $S(\eta)$, the growth of the number of particles and of the energy remains bounded, independently of N. More precisely, we show in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(a), Section 4] that

$$\begin{aligned} S^*(\eta)\,(\mathcal{N}_+ + 1)^m\, S(\eta) &\le C\,(\mathcal{N}_+ + 1)^m\,, \\ S^*(\eta)\,(\mathcal{H}_N + 1)(\mathcal{N}_+ + 1)\, S(\eta) &\le C\,(\mathcal{H}_N + 1)(\mathcal{N}_+ + 1)\,. \end{aligned} \tag{5-3}$$

In particular, Eq. (5-3) implies that the error term $\mathcal{E}_{\mathcal{G}_N}$ on the r.h.s. of (4-8) remains negligible, after conjugation with $S(\eta)$. To conjugate the quadratic operator $\mathcal{Q}_{\mathcal{G}_N}$ with $S(\eta)$, we observe first that

$$\pm\,[\mathcal{Q}_{\mathcal{G}_N}, A(\eta)] \le \frac{C}{\sqrt{N}}(\mathcal{N}_+ + 1)^2\,.$$

This bound, combined with the expansion

$$S^*(\eta)\mathcal{Q}_N S(\eta) = \mathcal{Q}_N + \int_0^1 ds\, e^{-sA(\eta)}\,[\mathcal{Q}_N, A(\eta)]\,e^{sA(\eta)}$$

and with the first estimate in (5-3), implies that

$$S^*(\eta)\mathcal{Q}_N S(\eta) = \mathcal{Q}_N + \mathcal{E}_1 \tag{5-4}$$

where the error operator $\mathcal{E}_1$ is such that $\pm\mathcal{E}_1 \le CN^{-1/2}(\mathcal{N}_+ + 1)^2$. To conjugate the cubic term $\mathcal{C}_N$ on the r.h.s. of (4-8), we compute

$$[\mathcal{C}_N, A(\eta)] = \Theta + \widetilde{\mathcal{E}}_2 \tag{5-5}$$

where

$$\Theta = \frac{2}{N} \sum_{r \in P_H, v \in P_L} \kappa \big(\widehat{V}(r/N) + \widehat{V}((r+v)/N) \big) \eta_r \times \Big[\sigma_v^2 + (\gamma_v^2 + \sigma_v^2)\, b_v^* b_v + \gamma_v \sigma_v \left(b_v b_{-v} + b_v^* b_{-v}^* \right) \Big]$$

and

$$\pm \widetilde{\mathcal{E}}_2 \leq C N^{-1/2} \left[(\mathcal{N}_+ + 1)^3 + (\mathcal{K}_N + 1)(\mathcal{N}_+ + 1) \right].$$

The term Θ on the r.h.s. of (5-5) is not small, but it is such that

$$\pm \left[\Theta, A(\eta) \right] \leq C N^{-1/2} (\mathcal{N}_+ + 1)^2 .$$

Hence, expanding to second order, we conclude that

(5-6)
$$S^*(\eta) \mathcal{C}_N S(\eta) = \mathcal{C}_N + \frac{2}{N} \sum_{r \in P_H, v \in P_L} \kappa \big(\widehat{V}(r/N) + \widehat{V}((r+v)/N) \big) \eta_r \times \Big[\sigma_v^2 + (\gamma_v^2 + \sigma_v^2)\, b_v^* b_v + \gamma_v \sigma_v \left(b_v b_{-v} + b_v^* b_{-v}^* \right) \Big] + \mathcal{E}_2$$

where

$$\pm \mathcal{E}_2 \leq C N^{-1/2} \left[(\mathcal{N}_+ + 1)^3 + (\mathcal{K}_N + 1)(\mathcal{N}_+ + 1) \right].$$

To compute the action of $S(\eta)$ on the quartic interaction $\mathcal{V}_N$, we proceed similarly (but here we have to expand one contribution up to third order). We obtain that

$$\begin{aligned} S^*(\eta) \mathcal{V}_N S(\eta) = \mathcal{V}_N &- \frac{1}{\sqrt{N}} \sum_{\substack{r \in P_H \\ v \in P_L}} \kappa \widehat{V}(r/N) \Big[b_{r+v}^* b_{-r}^* \left(\gamma_v b_v + \sigma_v b_{-v}^* \right) + \text{h.c.} \Big] \\ &- \frac{1}{N} \sum_{\substack{r \in P_H \\ v \in P_L}} \kappa \big(\widehat{V}(r/N) + \widehat{V}((r+v)/N) \big) \eta_r \\ &\qquad \Big[\sigma_v^2 + (\gamma_v^2 + \sigma_v^2)\, b_v^* b_v + \gamma_v \sigma_v \left(b_v b_{-v} + b_v^* b_{-v}^* \right) \Big] \\ &+ \mathcal{E}_3 \end{aligned} \tag{5-7}$$

where $\pm \mathcal{E}_3 \leq C N^{-1/4} [(\mathcal{N}_+ + 1)^3 + (\mathcal{K}_N + 1)(\mathcal{N}_+ + 1)]$. The proof of (5-4), (5-6) and (5-7) can be found in Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(a), Sect. 8]. Combining these results with Proposition 4.3, we arrive at the following proposition.

Proposition 5.1. *Let $V \in L^3(\mathbb{R}^3)$ be non-negative, spherically symmetric and compactly supported. Let the coupling constant $\kappa > 0$ be small enough. Let $\mathcal{J}_N$ be defined as in (5-2). Then we can write*

$$\mathcal{J}_N = C_{\mathcal{J}_N} + \mathcal{Q}_{\mathcal{J}_N} + \mathcal{V}_N + \mathcal{E}_{\mathcal{J}_N} \tag{5-8}$$

with

$$\begin{aligned} C_{\mathcal{J}_N} := \frac{(N-1)}{2}\kappa \widehat{V}(0) &+ \sum_{p\in\Lambda_+^*}\Big[p^2\sigma_p^2 + \kappa\widehat{V}(p/N)\sigma_p\gamma_p + \kappa\big(\widehat{V}(\cdot/N) * \widehat{f}_{N,\ell}\big)_p\sigma_p^2\Big] \\ &+ \frac{\kappa}{2N}\sum_{p,q\in\Lambda_+^*}\widehat{V}((p-q)/N)\sigma_q\gamma_q\sigma_p\gamma_p \\ &+ \frac{1}{N}\sum_{p\in\Lambda^*}\Big[p^2\eta_p^2 + \frac{\kappa}{2N}\big(\widehat{V}(\cdot/N) * \eta\big)_p\eta_p\Big] \end{aligned}$$

and the quadratic term

$$\mathcal{Q}_{\mathcal{J}_N} = \sum_{p\in\Lambda_+^*} F_p\, b_p^* b_p + \frac{1}{2}\sum_{p\in\Lambda_+^*} G_p\left(b_p^* b_{-p}^* + b_p b_{-p}\right) \tag{5-9}$$

with

$$\begin{aligned} F_p &= p^2(\sigma_p^2 + \gamma_p^2) + \kappa\big(\widehat{V}(\cdot/N) * \widehat{f}_{N,\ell}\big)_p(\gamma_p + \sigma_p)^2; \\ G_p &= 2p^2\sigma_p\gamma_p + \kappa\big(\widehat{V}(\cdot/N) * \widehat{f}_{N,\ell}\big)_p(\gamma_p + \sigma_p)^2 \end{aligned} \tag{5-10}$$

where $\widehat{f}_{N,\ell}$ is the Fourier series of $x \to f_{N,\ell}(x) = f_\ell(Nx)$. Moreover, the error operator $\mathcal{E}_{\mathcal{J}_N}$ is such that, on $\mathcal{F}_+^{\le N}$,

$$\pm\,\mathcal{E}_{\mathcal{J}_N} \le CN^{-1/4}\Big[(\mathcal{H}_N + 1)(\mathcal{N}_+ + 1) + (\mathcal{N}_+ + 1)^3\Big]. \tag{5-11}$$

6 Diagonalization and Excitation Spectrum

Comparing (5-8) with the decomposition (4-8) for $\mathcal{G}_N$, we notice the absence of the cubic term $\mathcal{C}_N$, achieved through conjugation with the cubic phase $S(\eta)$. The quartic term $\mathcal{V}_N$ still appears on the r.h.s. of (5-8) but it is non-negative and therefore we do not worry about it. To read off the excitation spectrum, we conjugate $\mathcal{J}_N$ with a last Bogoliubov transformation that diagonalize the quadratic operator $\mathcal{Q}_{\mathcal{J}_N}$. For $p \in \Lambda_+^*$, we define

$\tau_p \in \mathbb{R}$ such that (remark that, for $\kappa > 0$ small enough, it is clear that the coefficients F_p, G_p defined in (5-10) satisfy $|G_p/F_p| < 1$)

$$\tanh(2\tau_p) = -\frac{G_p}{F_p}\,. \tag{6-1}$$

Using the coefficients τ_p, we define

$$\mathfrak{M}_N = T^*(\tau)\mathcal{G}_N T(\tau) = T^*(\tau)S^*(\eta)T^*(\eta)U_N H_N U_N^* T(\eta)S(\eta)T(\tau) : \mathcal{F}_+^{\leq N} \to \mathcal{F}_+^{\leq N}.$$

From (5-10) and with the definition (3-9) of η, we find that $\tau \in H^1(\Lambda_+^*)$ with norm bounded uniformly in N. It follows from (3-12) that conjugation with $T(\tau)$ can only increase number of particles and energy by bounded quantities. This makes it easy to control the action of $T(\tau)$. We find (see Boccato, Brennecke, Cenatiempo, and Schlein [n.d.(a), Section 5] for more details) that

(6-2)

$$\mathfrak{M}_N = 4\pi(N-1)a_N + \frac{1}{2}\sum_{p\in\Lambda_+^*}\left[-p^2 - 8\pi a_0 + \sqrt{p^4 + 16\pi a_0 p^2} + \frac{(8\pi a_0)^2}{2p^2}\right] + \sum_{p\in\Lambda_+^*}\sqrt{|p|^4 + 16\pi a_0 p^2}\; a_p^* a_p + \mathcal{V}_N + \mathcal{E}_{\mathfrak{M}_N}$$

where

$$\pm\mathcal{E}_{\mathfrak{M}_N} \leq CN^{-1/4}[(\mathcal{H}_N + 1)(\mathcal{N}_+ + 1) + (\mathcal{N}_+ + 1)^3]\,.$$

Theorem 1.1 now follows making use of the min-max principle to compare the eigenvalues of $\mathfrak{M}_N$ (which coincide with those of the initial Hamiltonian (1-1)) with those of the quadratic operator

$$\mathfrak{D}_N = 4\pi(N-1)a_N + \frac{1}{2}\sum_{p\in\Lambda_+^*}\left[-p^2 - 8\pi a_0 + \sqrt{p^4 + 16\pi a_0 p^2}\right] + \sum_{p\in\Lambda_+^*}\sqrt{|p|^4 + 16\pi a_0\, p^2}\, a_p^* a_p$$

applying the a-priori bound (4-7) to control the contribution of the error term $\mathcal{E}_{\mathfrak{M}_N}$. The quartic interaction $\mathcal{V}_N$ can be neglected in the lower bounds because of its positivity. To prove that it can be neglected also in the proof of the necessary upper bounds, it is enough to observe that, on the range of the spectral projection $\mathbf{1}_{(-\infty;\zeta]}(\mathfrak{D}_N)$ (spanned by finitely many eigenvectors of $\mathfrak{D}_N$), we have (see Boccato, Brennecke, Cenatiempo, and Schlein [ibid., Section 6])

$$\mathcal{V}_N \leq CN^{-1}(\zeta + 1)^{7/2}\,.$$

References

Mike H Anderson, Jason R Ensher, Michael R Matthews, Carl E Wieman, and Eric A Cornell (1995). "Observation of Bose-Einstein condensation in a dilute atomic vapor". *science* 269, pp. 198–201 (cit. on p. 2687).

Niels Benedikter, Gustavo de Oliveira, and Benjamin Schlein (2015). "Quantitative derivation of the Gross-Pitaevskii equation". *Comm. Pure Appl. Math.* 68.8, pp. 1399–1482. MR: 3366749 (cit. on p. 2693).

Chiara Boccato, Christian Brennecke, Serena Cenatiempo, and Benjamin Schlein (n.d.[a]). "Bogoliubov excitation spectrum for Bose-Einstein condensates". Paper in preparation (cit. on pp. 2691, 2695, 2698, 2701, 2702, 2704).

– (n.d.[b]). "The excitation spectrum of Bose gases interacting through singular potentials". arXiv: 1704.04819 (cit. on pp. 2695, 2696).

– (n.d.[c]). "The excitation spectrum of Bose gases interacting through singular potentials". arXiv: 1704.04819 (cit. on pp. 2691, 2700).

N. Bogolubov (1947). "On the theory of superfluidity". *Acad. Sci. USSR. J. Phys.* 11, pp. 23–32. MR: 0022177 (cit. on p. 2690).

Christian Brennecke and Benjamin Schlein (n.d.). "Gross-Pitaevskii Dynamics for Bose-Einstein Condensates". arXiv: 1702.05625 (cit. on pp. 2694, 2695).

Kendall B Davis, M-O Mewes, Michael R Andrews, NJ Van Druten, DS Durfee, DM Kurn, and Wolfgang Ketterle (1995). "Bose-Einstein condensation in a gas of sodium atoms". *Physical review letters* 75.22, pp. 3969–3973 (cit. on p. 2687).

Jan Dereziński and Marcin Napiórkowski (2014). "Excitation spectrum of interacting bosons in the mean-field infinite-volume limit". *Ann. Henri Poincaré* 15.12, pp. 2409–2439. MR: 3272826 (cit. on pp. 2691, 2692).

E. H. Lieb and R. Seiringer (2002). "Proof of Bose-Einstein condensation for dilute trapped gases". *Phys. Rev. Lett.* 88, pp. 170–409 (cit. on p. 2688).

László Erdős, Benjamin Schlein, and Horng-Tzer Yau (2008). "Ground-state energy of a low-density Bose gas: A second-order upper bound". *Physical Review A* 78.5, p. 053627 (cit. on p. 2700).

Alessandro Giuliani and Robert Seiringer (2009). "The ground state energy of the weakly interacting Bose gas at high density". *J. Stat. Phys.* 135.5-6, pp. 915–934. MR: 2548599 (cit. on p. 2691).

Philip Grech and Robert Seiringer (2013). "The excitation spectrum for weakly interacting bosons in a trap". *Comm. Math. Phys.* 322.2, pp. 559–591. MR: 3077925 (cit. on pp. 2691, 2692).

Mathieu Lewin, Phan Thành Nam, Sylvia Serfaty, and Jan Philip Solovej (2015). "Bogoliubov spectrum of interacting Bose gases". *Comm. Pure Appl. Math.* 68.3, pp. 413–471. MR: 3310520 (cit. on pp. 2691, 2692).

EH Lieb and J Yngvason (1998). "Ground State Energy of the Low Density Bose Gas". *Phys. Rev. Lett.* 80, pp. 2504–2507 (cit. on p. 2688).

Elliott H. Lieb and Jan Philip Solovej (2002). "Erratum: "Ground state energy of the one-component charged Bose gas" [Comm. Math. Phys. **217** (2001), no. 1, 127–163; MR1815028 (2002c:82008)]". *Comm. Math. Phys.* 225.1, pp. 219–221. MR: 1877316 (cit. on p. 2690).

Elliott H Lieb, Robert Seiringer, and Jakob Yngvason (2000). "Bosons in a trap: A rigorous derivation of the Gross-Pitaevskii energy functional". *Physical Review A* 61.4, p. 043602 (cit. on p. 2688).

Phan Thành Nam, Nicolas Rougerie, and Robert Seiringer (2016). "Ground states of large bosonic systems: the Gross-Pitaevskii limit revisited". *Anal. PDE* 9.2, pp. 459–485. MR: 3513141 (cit. on p. 2688).

Marcin Napiórkowski, Robin Reuvers, and Jan Philip Solovej (n.d.). "The Bogoliubov free energy functional I. Existence of minimizers and phase diagram". arXiv: 1511.05935 (cit. on p. 2700).

Alessandro Pizzo (n.d.). "Bose particles in a box III. A convergent expansion of the ground state of the Hamiltonian in the mean field limiting regime". arXiv: 1511.07026 (cit. on pp. 2691, 2692).

Robert Seiringer (2011). "The excitation spectrum for weakly interacting bosons". *Comm. Math. Phys.* 306.2, pp. 565–578. MR: 2824481 (cit. on pp. 2691, 2692).

Horng-Tzer Yau and Jun Yin (2009). "The second order upper bound for the ground energy of a Bose gas". *J. Stat. Phys.* 136.3, pp. 453–503. MR: 2529681 (cit. on p. 2700).

Received 2017-11-30.

Benjamin Schlein
Institute of Mathematics, University of Zurich
Winterthurerstrasse 190
8057 Zurich
Switzerland
benjamin.schlein@math.uzh.ch

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2705–2726)

TRANSFER OPERATOR APPROACH TO 1D RANDOM BAND MATRICES

Mariya Shcherbina and Tatyana Shcherbina

Abstract

We discuss an application of the transfer operator approach to the analysis of the different spectral characteristics of 1d random band matrices (correlation functions of characteristic polynomials, density of states, spectral correlation functions). We show that when the bandwidth W crosses the threshold $W = N^{1/2}$, the model has a kind of phase transition (crossover), whose nature can be explained by the spectral properties of the transfer operator.

1 Introduction

Random band matrices (RBM) represent quantum systems on a large box in $\mathbb{Z}^d$ with random transition amplitudes effective up to distances of order W, which is called a bandwidth. They are natural intermediate models to study eigenvalue statistics and quantum propagation in disordered systems as they interpolate between Wigner matrices and random Schrödinger operators: Wigner matrix ensembles represent mean-field models without spatial structure, where the quantum transition rates between any two sites are i.i.d. random variables; in contrast, random Schrödinger operator has only a random diagonal potential in addition to the deterministic Laplacian on a box in $\mathbb{Z}^d$.

In the simplest 1d case RBM H is a Hermitian or real symmetric $N \times N$ matrix with independent (up to the symmetry condition) entries H_{ij} such that

$$\mathbb{E}\{H_{ij}\} = 0, \quad \mathbb{E}\{|H_{ij}|^2\} = (2W)^{-1}\mathbf{1}_{|i-j|\le W},$$

i.e. H is a Hermitian matrix which has only $2W + 1$ non zero diagonals whose entries are i.i.d. random variables (up to the symmetry) and the sum of the variances of entries in each line is 1.

Supported in part by NSF grant DMS-1700009.
MSC2010: 60B20.

In a more general case H is a Hermitian random $N \times N$ matrix, whose entries H_{jk} are independent (up to the symmetry) complex random variables with mean zero and variances scaled as

$$\mathbb{E}\{|H_{jk}|^2\} = \frac{1}{W^d} J\Big(\frac{|j-k|}{W}\Big). \tag{1.1}$$

Here Λ is a box in $\mathbb{Z}^d$, $|\Lambda| = N$, and $J : \mathbb{R}^d \to \mathbb{R}_+$ is a function having the compact support or decaying sufficiently fast at infinity and normalized in such a way that

$$W^{-d} \sum_{k \in \Lambda} J(|k|/W) = 1,$$

and the bandwidth $W \gg 1$ is a large parameter.

The density of states ρ of a general class of RBM with $W \gg 1$ is given by the well-known Wigner semicircle law (see Bogachev, Molchanov, and Pastur [1991] and Molchanov, Pastur, and Khorunzhiĭ [1992]):

$$\rho(E) = (2\pi)^{-1}\sqrt{4 - E^2}, \quad E \in [-2, 2]. \tag{1.2}$$

As it was mentioned above, a substantial interest to random band matrices is caused by the fact that they have a non-trivial spatial structure like random Schrödinger matrices (in contrast to classical random matrix ensembles), and furthermore RBM and random Schrödinger matrices are expected to have some similar qualitative properties (for more details on these conjectures see Spencer [2012]). For instance, RBM can be used to model the celebrated Anderson metal-insulator phase transition in $d \geq 3$. Moreover, the crossover for RBM can be investigated even in $d = 1$ by varying the bandwidth W.

The key physical parameter of RBM is the localization length ℓ_ψ, which describes the length scale of the eigenvector $\psi(E)$ corresponding to the energy $E \in (-2, 2)$. The system is called delocalized if for all E in the bulk of spectrum ℓ_ψ is comparable with the system size, $\ell_\psi \sim N$, and it is called localized otherwise. Delocalized systems correspond to electric conductors, and localized systems are insulators.

In the case of 1d RBM there is a fundamental conjecture stating that for every eigenfunction $\psi(E)$ in the bulk of the spectrum ℓ_ψ is of order W^2 (see Casati, Molinari, and Izrailev [1990] and Fyodorov and Mirlin [1991]). In $d = 2$, the localization length is expected to be exponentially large in W, in $d \geq 3$ it is expected to be macroscopic, $\ell_\psi \sim N$, i.e. system is delocalized.

Notice that the global eigenvalue statistics for 1d RBM such as density of states does not feel any difference between the regime $W \gg \sqrt{N}$ and $W \ll \sqrt{N}$ (see (1.2)). Same situation with the central limit theorem for the linear eigenvalue statistics which was proved in M. Shcherbina [2015] for any $W \gg 1$ (see also Li and Soshnikov [2013] for CLT in the regime $W \gg \sqrt{N}$). However, the questions of the localization length are closely related to the universality conjecture of the bulk *local* regime of the random matrix theory. The bulk local regime deals with the behaviour of eigenvalues of $N \times N$ random matrices on the intervals whose length is of the order $O(N^{-1})$. According to the Wigner – Dyson universality conjecture, this local behaviour does not depend on the matrix probability law (ensemble) and is determined only by the symmetry type of matrices (real symmetric, Hermitian, or quaternion real in the case of real eigenvalues and orthogonal, unitary or symplectic in the case of eigenvalues on the unit circle). In terms of eigenvalue statistics the conjecture about the localization length of RBM in $d = 1$ means that 1d RBM in the bulk of the spectrum changes the spectral local behaviour of random operator type with Poisson local eigenvalue statistics (for $W \ll \sqrt{N}$) to the local spectral behaviour of the GUE/GOE type (for $W \gg \sqrt{N}$). In particular, it means that if we consider the second correlation function R_2 defined by the equality

$$\mathbb{E}\Big\{ \sum_{j_1 \neq j_2} \varphi(\lambda_{j_1}, \lambda_{j_2}) \Big\} = \int_{\mathbb{R}^2} \varphi(\lambda_1, \lambda_2) R_2(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2, \tag{1.3}$$

where $\{\lambda_j\}$ are eigenvalues of a random matrix, the function $\varphi : \mathbb{R}^2 \to \mathbb{C}$ is bounded, continuous and symmetric in its arguments, and the summation is over all pairs of distinct integers $j_1, j_2 \in \Lambda$, then in the delocalization region $W \gg \sqrt{N}$

$$(N\rho(E))^{-2} R_2 \left(E + \frac{\xi_1}{\rho(E)\, N}, E + \frac{\xi_2}{\rho(E)\, N} \right) \longrightarrow 1 - \frac{\sin^2(\pi(\xi_1 - \xi_2))}{\pi^2(\xi_1 - \xi_2)^2}, \tag{1.4}$$

while in the localization region

$$(N\rho(E))^{-2} R_2 \left(E + \frac{\xi_1}{\rho(E)\, N}, E + \frac{\xi_2}{\rho(E)\, N} \right) \longrightarrow 1. \tag{1.5}$$

The conjecture on the crossover in RBM with $W \sim \sqrt{N}$ is supported by physical derivation due to Fyodorov and Mirlin (see Fyodorov and Mirlin [1991]) based on supersymmetric formalism, and also by the so-called Thouless scaling. However, there are only a few partial results on the mathematical level of rigour. At the present time only some upper and lower bounds for ℓ_ψ for the general class of 1d RBM are proved rigorously. It is known from the paper Schenker [2009] that $\ell_\psi \leq W^8$. Recently this bound was improved in Peled, Schenker, Shamis, and Sodin [2016] to W^7. On the other side, for the general Wigner matrices (i.e. $W = N$) the bulk universality has been proved in Erdős, Yau, and Yin [2012] and Tao and Vu [2011], which gives $\ell_\psi \geq W$. By a development of the Erdős-Yau approach, there were also obtained some other results, where the localization length is controlled: $\ell_\psi \geq W^{7/6}$ in Erdős and Knowles [2011] and $\ell_\psi \geq W^{5/4}$

in Erdős, Knowles, Yau, and Yin [2013]. GUE/GOE gap distributions for $W \sim N$ was proved recently in Bourgade, Erdős, Yau, and Yin [2017].

The study of the eigenfunctions decay is closely related to properties of the Green function $(H - E - i\varepsilon)^{-1}$ with a small ε. For instance, if $|(H - E - i\varepsilon)^{-1}_{ii}|^2$ (without expectation) is bounded for all i and all $E \in (-2, 2)$, then each normalized eigenvector ψ of H is delocalized on the scale ε^{-1} in a sense that

$$\max_i |\psi_i|^2 \le C\varepsilon$$

and so ψ is supported on at least ε^{-1} sites. In particular, if $|(H - E - i\varepsilon)^{-1}_{ii}|^2$ can be controlled down to the scale $\varepsilon \sim 1/N$, then the system is in the complete delocalized regime. Moreover, in view of the bound

$$\mathbb{E}\{|(H - E - i\varepsilon)^{-1}_{jk}|^2\} \sim C\varepsilon^{-1} e^{-\|j-k\|/\ell}$$

which holds for the localized regime, the problem of localization/delocalization reduces to controlling

$$\mathbb{E}\{|(H - E - i\varepsilon)^{-1}_{jk}|^2\}$$

for $\varepsilon \sim 1/N$. As will be shown below, similar estimates of $\mathbb{E}\{|\mathrm{Tr}\,(H - E - i\varepsilon)^{-1}|^2\}$ for $\varepsilon \sim 1/N$ are required to work with the correlation functions of RBM.

Despite many attempts, such control has not been achieved so far. The standard approaches of Erdős, Yau, and Yin [2012] and Erdős, Knowles, Yau, and Yin [2013] do not seem to work for $\varepsilon \le W^{-1}$, and so cannot give an information about the strong form of delocalization (i.e. for *all* eigenfunctions). Classical moment methods, even with a delicate renormalization approach Sodin [2011], could not break the barrier $\varepsilon \sim W^{-1}$ either.

Another method which allows to work with random operators with non-trivial spatial structures and breaks that barrier, is supersymmetry techniques (SUSY). It is based on the representation of the determinant as an integral (formal) over the Grassmann variables. Combining this representation with the representation of the inverse determinant as an integral over the Gaussian complex field, SUSY allows to obtain the integral representation for the main spectral characteristics such as averaged density of states and correlation functions, as well as for $\mathbb{E}\{G_{jk}(E+i\varepsilon)\}$, $\mathbb{E}\{|G_{jk}(E+i\varepsilon)|^2\}$, etc. For instance, according to the properties of the Stieljes transform, the second correlation function can be rewritten in the form

(1.6)

$$\begin{aligned} R_2(\lambda_1,\lambda_2) =&(\pi N)^{-2}\lim_{\varepsilon\to 0}\mathbb{E}\{\Im\,\mathrm{Tr}\,(H-\lambda_1-i\varepsilon)^{-1}\Im\,\mathrm{Tr}\,(H-\lambda_2-i\varepsilon)^{-1}\}\\ =&(2i\pi N)^{-2}\lim_{\varepsilon\to 0}\mathbb{E}\Big\{\Big(\mathrm{Tr}\,(H-\lambda_1-i\varepsilon)^{-1}-\mathrm{Tr}\,(H-\lambda_1+i\varepsilon)^{-1}\Big)\\ &\times\Big(\mathrm{Tr}\,(H-\lambda_2-i\varepsilon)^{-1}-\mathrm{Tr}\,(H-\lambda_2+i\varepsilon)^{-1}\Big)\Big\}, \end{aligned}$$

and since

(1.7)
$$\mathbb{E}\{\mathrm{Tr}\,(H-z_1)^{-1}\mathrm{Tr}\,(H-z_2)^{-1}\} = \frac{d^2}{dz_1'dz_2'}\mathbb{E}\Big\{\frac{\det(H-z_1)\det(H-z_2))}{\det(H-z_1')\det(H-z_2'))}\Big\}\Big|_{z'=z},$$

R_2 can be represented as a sum of derivatives of the expectation of ratio of four determinants. Besides, it is expected that if we set

$$\begin{aligned} z_1 &= E + i\varepsilon/N + \xi_1/N\rho(E), \quad z_2 = E + i\varepsilon/N + \xi_2/N\rho(E), \\ z_1' &= E + i\varepsilon/N + \xi_1'/N\rho(E), \quad z_2' = E + i\varepsilon/N + \xi_2'/N\rho(E), \end{aligned} \tag{1.8}$$

then the r.h.s. of (1.7) before taking derivatives is an analytic function in $\xi_1, \xi_2, \xi_1', \xi_2'$. Thus, to study the second correlation function, it suffices to study the ratio of four determinants, which we call the second "generalized" correlation functions

$$\begin{aligned} \mathcal{R}_2^{+-}(z_1, z_1'; z_2, z_2') &= \mathbb{E}\Big\{\frac{\det(H-z_1)\det(H-\overline{z}_2)}{\det(H-z_1')\det(H-\overline{z}_2')}\Big\}, \\ \mathcal{R}_2^{++}(z_1, z_1'; z_2, z_2') &= \mathbb{E}\Big\{\frac{\det(H-z_1)\det(H-z_2)}{\det(H-z_1')\det(H-z_2')}\Big\}. \end{aligned} \tag{1.9}$$

Similarly the derivative of the first "generalized" correlation function

$$\mathcal{R}_1(z_1, z_1') := \mathbb{E}\Big\{\frac{\det(H-z_1')}{\det(H-z_1)}\Big\}$$

gives the Stieltjes transform of of the density of states (the first correlation function).

Instead of eigenvalue correlation functions one can consider more simple objects which are the correlation functions of characteristic polynomials:

$$\mathcal{R}_0(\lambda_1, \lambda_2) = \mathbb{E}\Big\{\det(H-\lambda_1)\det(H-\lambda_2)\Big\}, \quad \lambda_{1,2} = E \pm \xi/N\rho(E). \tag{1.10}$$

Characteristic polynomials are the objects of independent interest because of their connections to the number theory, quantum chaos, integrable systems, combinatorics, representation theory and others. But in our context the main point is that from the SUSY point of view correlation functions of characteristic polynomials correspond to the so-called fermion-fermion (Grassmann) sector of the supersymmetric full model describing the usual correlation functions (since they represent two determinants in the numerator of (1.9)). They are especially convenient for the SUSY approach and were successfully studied by the techniques for many ensembles (see Brézin and Hikami [2000], Brézin and Hikami [2001], T. Shcherbina [2011], T. Shcherbina [2013], etc.). In addition, although $\mathcal{R}_0(\lambda_1, \lambda_2)$ is not a local object, it is also expected to be universal in some sense.

Moreover, correlation functions of characteristic polynomials are expected to exhibit a crossover which is similar to that of local eigenvalue statistics. In particular, for 1d RBM they are expected to have the same local behaviour as for GUE for $W \gg \sqrt{N}$, and the different behaviour for $W \ll \sqrt{N}$. Besides, the analysis of $\mathcal{R}_0(\lambda_1, \lambda_2)$ is much less involved than that for $\mathcal{R}_2^{+-}(z_1, z_1'; z_2, z_2')$, but on the other hand, this analysis allows to understand the nature of the crossover in RBM when W crosses the threshold $W \sim \sqrt{N}$.

The derivation of SUSY integral representation is basically an algebraic step, and usually it can be done by the standard algebraic manipulations. SUSY is widely used in the physics literature, but the rigorous analysis of the obtained integral representation is a real mathematical challenge. Usually it is quite difficult, and it requires a powerful analytic and statistical mechanics techniques, such as a saddle point analysis, transfer operators, cluster expansions, renormalization group methods, etc. However, it can be done rigorously for some special class of RBM.

There exist especially convenient classes of RBM, where the control of SUSY integral representation becomes more accessible. One of them was introduced in Disertori, Pinson, and Spencer [2002]: it is (1.1) with Gaussian elements with variance

$$\mathbb{E}\{|H_{jk}|^2\} = \left(-W^2\Delta + 1\right)^{-1}_{jk}, \tag{1.11}$$

where $\triangle$ is the discrete Laplacian on Λ with Neumann boundary conditions: for the case $d = 1$,

$$(-\Delta f)_j = \begin{cases} -f_{j-1} + 2f_j - f_{j+1}, & j \neq 1, n, \\ -f_{j-1} + f_j - f_{j+1}, & j = 1, n \end{cases} \tag{1.12}$$

with $f_0 = f_{n+1} = 0$. It is easy to see that in 1d case $J_{jk} \approx C_1 W^{-1} \exp\{-C_2|j-k|/W\}$, and so the variance of matrix elements is exponentially small when $|j - k| \gg W$.

Another class of convenient models are the Gaussian block RBM which are the special class of Wegner's orbital models (see Wegner [1979]). Gaussian block RBM are $N \times N$ Hermitian block matrices composed from n^2 blocks of the size $W \times W$ ($N = nW$). Only 3 block diagonals are non zero:

$$H = \begin{pmatrix} A_1 & B_1 & 0 & 0 & 0 & \dots & 0 \\ B_1^* & A_2 & B_2 & 0 & 0 & \dots & 0 \\ 0 & B_2^* & A_3 & B_3 & 0 & \dots & 0 \\ . & . & B_3^* & . & . & . & . \\ . & . & . & . & . & A_{n-1} & B_{n-1} \\ 0 & . & . & . & 0 & B_{n-1}^* & A_n \end{pmatrix}.$$

Here $A_1, \dots A_n$ are independent $W \times W$ GUE-matrices with i.i.d. (up to the symmetry) Gaussian entries with variance $(1 - 2\alpha)/W$, $\quad \alpha < \frac{1}{4}$, and $B_1, \dots B_{n-1}$ are independent $W \times W$ Ginibre matrices with i.i.d. Gaussian entries with variance α/W.

More precisely, H is Hermitian matrices with complex zero-mean random Gaussian entries $H_{jk,\alpha\beta}$, where $j,k \in \Lambda \subset \mathbb{Z}^d$ (they parameterize the lattice sites) and $\alpha, \gamma = 1, \dots, W$ (they parametrize the orbitals on each site), such that

$$\langle H_{j_1k_1,\alpha_1\gamma_1} H_{j_2k_2,\alpha_2\gamma_2}\rangle = \delta_{j_1k_2}\delta_{j_2k_1}\delta_{\alpha_1\gamma_2}\delta_{\gamma_1\alpha_2} J_{j_1k_1} \tag{1.13}$$

with

$$J = 1/W + \alpha\Delta/W, \tag{1.14}$$

where $W \gg 1$ and Δ is the discrete Laplacian on Λ (as in (1.11)). The probability law of H can be written in the form

$$P_N(dH) = \exp\Big\{-\frac{1}{2}\sum_{j,k\in\Lambda}\sum_{\alpha,\gamma=1}^{W}\frac{|H_{jk,\alpha\gamma}|^2}{J_{jk}}\Big\}dH. \tag{1.15}$$

This model is one of the possible realizations of the Gaussian RBM, for example for $d = 1$ they correspond to the band matrices with the bandwidth $2W + 1$. Let us remark that for this model $N = nW$, hence the crossover is expected for $n \sim W$.

The main advantage of both models (1.11) and (1.13) – (1.14) is that the main spectral characteristics such as density of states, R_2, $\mathbb{E}\{|G_{jk}(E + i\varepsilon)|^2\}$ for these models can be expressed via SUSY as the averages of certain observables of *nearest-neighbour* statistical mechanics models on Λ, which makes the model easier. For instance, the detailed information about the averaged density of states Gaussian RBM (1.11) in dimension 3 including local semicircle low at arbitrary short scales and smoothness in energy (in the limit of infinite volume and fixed large band width W) was obtained in Disertori, Pinson, and Spencer [2002]. The techniques of that paper were used in Disertori and Lager [2017] to obtain the same result in 2d. The rigorous application of SUSY to the Gaussian block RBM (1.13) – (1.14) was developed in T. Shcherbina [2014b], where the universality of the bulk local regime for $n = const$ was proved. Combining this approach with Green's function comparison strategy it has been proved in Bao and Erdős [2017] that $\ell \geq W^{7/6}$ (in a strong sense) for the block band matrices with rather general element's distribution.

The nearest-neighbour structure of the model also allows to combine the SUSY techniques with a transfer operator approach.

2 The idea of the transfer operator approach

The supersymmetric transfer operator formalism was first suggested by Efetov (see Efetov [1997]) and on a heuristic level it was adapted specifically for RBM in Fyodorov and Mirlin [1994] (see also references therein). The rigorous application of the method to the density of states and correlation function of characteristic polynomials was done in M. Shcherbina and T. Shcherbina [2016], M. Shcherbina and T. Shcherbina [2017], M. Shcherbina and T. Shcherbina [2018], T. Shcherbina [n.d.]. The approach is based on the fact that many nearest-neighbour statistical mechanics problems in 1d can be formulated in terms of properties of some integral operator K that is called a transfer operator. More precisely, the discussion above yields that for 1d RBM of the form (1.11) or (1.13) – (1.14) the SUSY techniques helps to find a scalar kernel $\mathcal{K}_0(X_1, X_2)$ and matrix kernels $\mathcal{K}_1(X_1, X_2)$, $\mathcal{K}_2(X_1, X_2)$ (containing $z_{1,2}, z'_{1,2}$ as parameters) such that

(2.1)

$$\begin{aligned}
\mathcal{R}_0(\lambda_1, \lambda_2) &= C_N \int g_0(X_1)\mathcal{K}_0(X_1, X_2)\ldots\mathcal{K}_0(X_{n-1}, X_n) f_0(X_n) \prod dX_i, \\
\mathcal{R}_1(z_1, z_1') &= W^2 \int g_1(X_1)\mathcal{K}_1(X_1, X_2)\ldots\mathcal{K}_1(X_{n-1}, X_n) f_1(X_n) \prod dX_i, \\
\mathcal{R}_2(z_1, z_1'; z_2, z_2') &= W^4 \int g_2(X_1)\mathcal{K}_2(X_1, X_2)\ldots\mathcal{K}_2(X_{n-1}, X_n) f_2(X_n) \prod dX_i,
\end{aligned}$$

where $\{X_j\}$ are Hermitian 2×2 matrices for the cases of $\mathcal{R}_0$, 2×2 matrices whose entries depend on 2 spacial variables $x_{1j}, y_{1j} \in \mathbb{R}$ for the cases $\mathcal{R}_1$, and for the case of $\mathcal{R}_2$ $\{X_j\}$ are vectors of dimensionality 70, whose components depend on 4 spacial variables $x_{1j}, x_{2j}, y_{1j}, y_{2j} \in \mathbb{R}$, unitary 2×2 matrix U_j, and hyperbolic 2×2 matrix S_j, dX_j means the standard measure on Herm(2) for $\mathcal{R}_0$, $dX_j = dx_{j1}dy_{j1}$ for $\mathcal{R}_1$, and for $\mathcal{R}_2$ dX_j means the integration over $dx_{1j}dx_{2j}dy_{1j}dy_{2j}dU_jdS_j$ with dU, dS being the corresponding Haar measures.

Remark, that for the model (1.11) $n = N$, while for the block band matrix (1.13) – (1.15) n is a number of blocks on the main diagonal.

The idea of the transfer operator approach is very simple and natural. Let $\mathcal{K}(X, Y)$ be the matrix kernel of the compact integral operator in $\oplus_{i=1}^{p} L_2[X, d\mu(X)]$. Then

$$\begin{aligned}
&\int g(X_1)\mathcal{K}(X_1, X_2)\ldots\mathcal{K}(X_{n-1}, X_n) f(X_n) \prod d\mu(X_i) = (\mathcal{K}^{n-1} f, \bar{g}) \\
&= \sum_{j=0}^{\infty} \lambda_j^{n-1}(\mathcal{K}) c_j, \quad with \quad c_j = (f, \psi_j)(g, \tilde{\psi}_j).
\end{aligned} \tag{2.2}$$

Here $\{\lambda_j(\mathcal{K})\}_{j=0}^{\infty}$ are the eigenvalues of $\mathcal{K}$ ($|\lambda_0| \geq |\lambda_1| \geq \ldots$), ψ_j are corresponding eigenvectors, and $\tilde{\psi}_j$ are the eigenvectors of $\mathcal{K}^*$. Hence, to study the correlation function, it is sufficient to study the eigenvalues and eigenfunctions of the integral operator with a kernel $\mathcal{K}(X, Y)$.

The main difficulties here are the complicated structure and non self-adjointness of the corresponding transfer operators.

In fact, since the analysis of eigenvectors of non self-adjoint operators is rather involved, it is simpler to work with the resolvent analog of (2.2)

$$\mathcal{R}_\alpha = (\mathcal{K}_\alpha^{n-1} f, \bar{g}) = -\frac{1}{2\pi i}\oint_{\mathcal{L}} z^{n-1}(\mathcal{G}_\alpha(z) f, \bar{g})dz, \qquad \mathcal{G}_\alpha(z) = (\mathcal{K}_\alpha - z)^{-1}, \tag{2.3}$$

where $\alpha = 0, 1, 2$ and $\mathcal{L}$ is any closed contour which contains all eigenvalues of $\mathcal{K}_\alpha$. For any α if we set

$$\lambda_* = \lambda_0(\mathcal{K}_\alpha), \quad (\lambda_* \sim 1),$$

then it suffices to choose $\mathcal{L}$ as $\mathcal{L}_0 = \{z : |z| = |\lambda_*|(1 + O(n^{-1}))\}$. However, it is more convenient to choose $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$, where $\mathcal{L}_2 = \{z : |z| = |\lambda_*|(1 - \log^2 n/n)\}$, and $\mathcal{L}_1$ is some contour in the domain between $\mathcal{L}_0$ and $\mathcal{L}_2$ which contains all eigenvalues of $\mathcal{K}_\alpha$ outside of $\mathcal{L}_2$. Then

$$(\mathcal{K}_\alpha^{n-1} f, \bar{g}) = -\frac{1}{2\pi i}\oint_{\mathcal{L}_1} z^{n-1}(\mathcal{G}_\alpha(z) f, \bar{g})dz - \frac{1}{2\pi i}\oint_{\mathcal{L}_2} z^{n-1}(\mathcal{G}_\alpha(z) f, \bar{g})dz,$$

and if we have a reasonable bound for $\|\mathcal{G}_\alpha(z)\|$ ($z \in \mathcal{L}_2$), then the second integral is small comparing with $|\lambda_*|^{n-1}$, since

$$|z|^{n-1} \le |\lambda_*|^{n-1} e^{-\log^2 n}.$$

Hence, it is natural to expect that the integral over $\mathcal{L}_1$ gives the main contribution to $\mathcal{R}_\alpha$.

Definition 2.1. *We shall say that the operator $\mathcal{A}_{n,W}$ is equivalent to $\mathcal{B}_{n,W}$ ($\mathcal{A}_{n,W} \sim \mathcal{B}_{n,W}$), if for some certain contour $\mathcal{L}_1$ (the choice of $\mathcal{L}_1$ depends on the problem)*

$$((\mathcal{A}_{n,W} - z)^{-1} f, \bar{g}) = ((\mathcal{B}_{n,W} - z)^{-1} f, \bar{g})(1 + o(1)), \quad n, W \to \infty,$$

with f, g of (2.2).

The idea is to find some $\mathcal{K}_{*\alpha} \sim \mathcal{K}_\alpha$ whose spectral analysis we are ready to perform.

3 Mechanism of the crossover for $\mathcal{R}_0$

As it was mentioned in Section 1, the simplest object which allows to understand the crossover's mechanism for the 1d RBM (1.11) is the correlation function of characteristic polynomials $\mathcal{R}_0$. Using SUSY and the idea of the transfer operator approach, one can write $\mathcal{R}_0$ (see M. Shcherbina and T. Shcherbina [2017]) as

$$\mathcal{R}_0\Big(E + \frac{\xi}{N\rho(E)}, E - \frac{\xi}{N\rho(E)}\Big) = C_N \cdot W^{-4n} \det{}^{-2} J \cdot (K_{0\xi}^{n-1}\mathcal{F}_\xi, \bar{\mathcal{F}}_\xi), \tag{3.1}$$

where $(\cdot,\cdot)$ is a standard inner product in $L_2(\mathrm{Herm}(2), dX)$ (i.e., 2×2 Hermitian matrices), with respect to the measure

$$dX_j = d(X_j)_{11} d(X_j)_{22} d\Re(X_j)_{12} d\Im(X_j)_{12},$$

C_N is some ξ-independent constant, $K_{0\xi} : \mathcal{H} \to \mathcal{H}$ is the operator with the kernels

$$K_\xi(X,Y) = \frac{W^4}{2\pi^2}\mathcal{F}_\xi(X)\,\exp\Big\{-\frac{W^2}{2}\mathrm{Tr}\,(X-Y)^2\Big\}\,\mathcal{F}_\xi(Y). \tag{3.2}$$

where $\hat{\xi} = \mathrm{diag}\,\{\xi,-\xi\}$, $\Lambda_0 = E\cdot I_2$, and $\mathcal{F}_\xi(X)$ is the operator of multiplication by

$$\mathcal{F}_\xi(X) = \mathcal{F}(X)\cdot\exp\Big\{-\frac{i}{2n\rho(E)}\,\mathrm{Tr}\,X\hat{\xi}\Big\} \tag{3.3}$$

with

$$\mathcal{F}(X) = \exp\Big\{-\frac{1}{4}\,\mathrm{Tr}\,\Big(X+\frac{i\Lambda_0}{2}\Big)^2 + \frac{1}{2}\,\mathrm{Tr}\,\log\big(X - i\Lambda_0/2\big) - C_+\Big\}$$

and some specific C_+. Notice that the stationary points of $\mathcal{F}$ are

$$a_+ = -a_- = \sqrt{1-E^2/4} = \pi\rho(E). \tag{3.4}$$

The first step is to show that if we introduce the projection $P_\pm$ onto the $W^{-1/2}\log W$-neighbourhood of the "surface" $X_*(U) = UDU^*$ with $D = \mathrm{diag}\,\{a_+,a_-\}$ and $U \in \mathring{U}(2) := U(2)/U(1)\times U(1)$, then in the sense of Definition 2.1

$$K_{0\xi} \sim P_\pm K_{0\xi} P_\pm. \tag{3.5}$$

Then, to study the operators $P_\pm K_{0\xi} P_\pm$, we use the "polar coordinates". Namely, introduce

$$t = (x_1-y_1)(x_2-y_2), \quad p(x,y) = \frac{\pi}{2}(x-y)^2, \tag{3.6}$$

and denote by dU the integration with respect to the Haar measure on the group $\mathring{U}(2)$. Consider the space $L_2[\mathbb{R}^2, p]\times L_2[\mathring{U}(2), dU]$. The inner product and the action of an integral operator in this space are

(3.7)

$$(f,g)_p = \int f(x,y)\bar{g}(x,y)p(x,y)\,dx\,dy;$$

$$(Mf)(x_1,y_1,U_1) = \int M(x_1,y_1,U_1;x_2,y_2,U_2)\,f(x_2,y_2,U_2)\,p(x_2,y_2)dx_2dy_2dU_2.$$

Changing the variables

$$X = U^*\Lambda U, \quad \Lambda = \mathrm{diag}\{x_1, x_2\}, \quad x_1 > x_2, \quad U \in \mathring{U}(2),$$

we obtain that $K_{0\xi}$ can be represented as an integral operator in $L_2[\mathbb{R}^2, p] \times L_2[\mathring{U}(2), dU]$ defined by the kernel

$$\mathcal{K}_{0\xi}(X, Y) \to \mathcal{K}_{0\xi}(x_1, y_1, U_1; x_2, y_2, U_2), \tag{3.8}$$

where

$$\mathcal{K}_{0\xi}(x_1, y_1, U_1; x_2, y_2, U_2) = \\ = t^{-1} A_1(x_1, x_2) A_2(y_1, y_2) K_{*0\xi}(t, U_1, U_2)(1 + O(n^{-1}W^{-1/2}));$$

$$A_{1,2}(x_1, x_2) = (2\pi)^{-1/2} e^{-W^2(x_1-x_2)^2/2} e^{f_{1,2}(x_1)+f_{1,2}(x_2)} \tag{3.9}$$

$$K_{*0\xi}(t, U_1, U_2) := \\ = W^2 t \cdot e^{tW^2 \mathrm{Tr}\, U_1 U_2^* L (U_1 U_2^*)^* L/4 - tW^2/2} e^{-i\xi\pi(\nu(U_1)+\nu(U_2))/n} \tag{3.10}$$

$$\nu(U) = \mathrm{Tr}\, U^* L U L/2, \qquad L = \mathrm{diag}\{1, -1\},$$

and t is defined in (3.6). The concrete form of $f_{1,2}$ in (3.9) is not important for us now. It is important that they are analytic functions with stationary points $a_\pm$ (see (3.4)). The analysis of the resolvent of A_1 and A_2 allows us to show that only eigenfunctions localized in the $W^{-1/2} \log W$ neighbourhood of a_+ and a_- give essential contribution in (2.2). More precisely, the resolvent analysis of $A_{1,2}$ allows to prove (3.5). Further resolvent analysis gives

$$\begin{aligned} &P_\pm \mathcal{K}_{0\xi} P_\pm \sim \mathcal{K}_{*\xi} \otimes \mathcal{A}, \\ &\mathcal{K}_{*\xi}(U_1, U_2) := K_{*0\xi}(t^*, U_1, U_2) \quad with \quad t^* = (a_+ - a_-)^2 = 4\pi^2 \rho(E)^2, \\ &\mathcal{A}(x_1, x_2, y_1, y_2) = A_1(x_1, x_2) A_2(y_1, y_2). \end{aligned} \tag{3.11}$$

Then from (2.3) and Definition 2.1 it is easy to obtain

$$\mathcal{R}_\xi = C_n(\mathcal{K}_{*\xi}^{n-1} \otimes \mathcal{A}^{n-1} f, \bar{g})(1 + o(1)) = (\mathcal{K}_{*\xi}^{n-1} f_0, f_0)(\mathcal{A}^{n-1} f_1, \bar{g}_1)(1 + o(1)),$$

where we used that both f, g asymptotically can be replaced by $f_0(U) \otimes f_1(x, y)$ with

$$f_0 \equiv 1. \tag{3.12}$$

If we introduce

$$D_2 = \Re_0(E, E), \tag{3.13}$$

then the above consideration yields

$$D_2^{-1}\Re_0\Big(E + \frac{\xi}{N\rho(E)}, E - \frac{\xi}{N\rho(E)}\Big) = \frac{(\mathcal{K}_{*\xi}^{n-1} f_0, f_0)}{(\mathcal{K}_{*0}^{n-1} f_0, f_0)}(1 + o(1)). \tag{3.14}$$

A good news here is that the operator $\mathcal{K}_{*0}$ is self-adjoint and his kernel depends only on $|(U_1U_2^*)_{12}|^2$. By Vilenkin [1968], his eigenfunctions are associated Legendre polynomials P_k^j. Moreover since $\mathcal{K}_{*0}$ is reduced by the space $\mathcal{E}_0 \subset L_2(U(2))$ of functions which depends only on $|U_{12}|^2$, and $f_0 \in \mathcal{E}_0$, we can restrict our spectral analysis to $\mathcal{E}_0$. In this space eigenfunctions of $\mathcal{K}_{*0}$ are Legendre polynomials P_j and it is easy to check that correspondent eigenvalues have the form

$$\lambda_j = 1 - j(j+1)/t^*W^2 + O((j(j+1)/W^2)^2), \quad j = 0, 1 \dots \tag{3.15}$$

with t^* of (3.11). Moreover, it follows from (3.10) that

$$\mathcal{K}_{*\xi} = \mathcal{K}_{*0} - 2n^{-1}\pi i\xi\hat{\nu} + o(n^{-1}),$$

where $\hat{\nu}$ is the operator of multiplication by ν of (3.10). Thus, the eigenvalues of $\mathcal{K}_{*\xi}$ are in the n^{-1}-neighbourhood of λ_j. This implies that for $W^{-2} \gg n^{-1} = N^{-1}$

$$|\lambda_1(\mathcal{K}_{*\xi})| \le 1 - O(W^{-2}), \quad \lambda_0 = 1 - 2n^{-1}\pi i\xi(\nu f_0, f_0) + o(n^{-1})$$

Since

$$(\nu f_0, f_0) = 0,$$

we obtain that the numerator and the denominator of (3.14) tends to 1 in this regime.

To study the regime $W^{-2} = Cn^{-1} = CN^{-1}$, observe that the Laplace operator Δ_U on $U(2)$ is also reduced by $\mathcal{E}_0$ and has the same eigenfunctions as $\mathcal{K}_{*0}$ with eigenvalues

$$\lambda_j^* = j(j+1).$$

Hence, we can write $\mathcal{K}_{*\xi}$ as

$$\mathcal{K}_{*\xi} \sim 1 - n^{-1}(C\Delta_U - 2i\xi\pi\nu) \Rightarrow (\mathcal{K}_{*\xi}^{n-1} f_0, f_0) \to (e^{-C\Delta_U + 2i\xi\pi\hat{\nu}} f_0, f_0),$$

where

$$\Delta_U = -\frac{d}{dx}x(1-x)\frac{d}{dx}, \quad x = |U_{12}|^2. \tag{3.16}$$

And in the regime $W^{-2} \ll n^{-1} = N^{-1}$ we have $\mathcal{K}_{*0}^{n-1} \to I$ in the strong vector topology, hence

$$\mathcal{K}_{*\xi} \sim 1 - n^{-1} 2i\xi\pi\nu \Rightarrow (\mathcal{K}_{*\xi}^{n-1} f_0, f_0) \to (e^{-2i\xi\pi\hat{\nu}} f_0, f_0)$$

and the numerator of (3.14) is given by the multiplication of f_0 by $e^{-2i\xi\pi\hat{\nu}}$, which gives the same form as for the correlation function of the Wigner model.

The last result was proved in T. Shcherbina [2014a] with a different method:

Theorem 3.1 (**T. Shcherbina [ibid.]**). *For the 1d RBM of (1.11) with $W^2 = N^{1+\theta}$, where $0 < \theta \le 1$, we have*

$$\lim_{n\to\infty} D_2^{-1} \mathfrak{R}_0\Big(E + \frac{\xi}{N\rho(E)}, E - \frac{\xi}{N\rho(E)}\Big) = \frac{\sin(2\pi\xi)}{2\pi\xi}, \tag{3.17}$$

i.e. the limit coincides with that for GUE. The limit is uniform in ξ varying in any compact set $C \subset \mathbb{R}$. Here $\rho(x)$ and $\mathfrak{R}_0$ are defined in (1.2) and (1.10), $E \in (-2, 2)$.

The regime $W^{-2} \gg N^{-1}$ was studied in M. Shcherbina and T. Shcherbina [2017]:

Theorem 3.2. *For the 1d RBM of (1.11) with $1 \ll W \le \sqrt{N/C_* \log N}$ for sufficiently big C_*, we have*

$$\lim_{n\to\infty} D_2^{-1} \mathfrak{R}_0\Big(E + \frac{\xi}{N\rho(E)}, E - \frac{\xi}{N\rho(E)}\Big) = 1,$$

where the limit is uniform in ξ varying in any compact set $C \subset \mathbb{R}$. Here $E \in (-2, 2)$, and $\rho(x)$, $\mathfrak{R}_0$, and D_2 are defined in (1.2), (1.10), and (3.13).

Remark 3.1. *Although the result is formulated for $\xi_1 = -\xi_2 = \xi$ in (1.8), one can prove Theorem 3.2 for $\xi_1, \xi_2 \in [-C, C] \subset \mathbb{R}$ by the same arguments with minor revisions. The only difference is a little bit more complicated expressions for D_2 and K_ξ.*

The regime $W^{-2} = C_* N^{-1}$ is studied in T. Shcherbina [n.d.]:

Theorem 3.3. *For the 1d RBM of (1.11) with $N = C_* W^2$, we have*

$$\lim_{n\to\infty} D_2^{-1} \mathfrak{R}_0\Big(E + \frac{\xi}{N\rho(E)}, E - \frac{\xi}{N\rho(E)}\Big) = (e^{-C\Delta_U - 2\pi i\xi\hat{\nu}} f_0, f_0),$$

where $C = 1/t^* C_*$ with t^* of (3.11), and the limit is uniform in ξ varying in any compact subset of $\mathbb{R}$. Here $E \in (-2, 2)$.

4 Analysis of $\mathfrak{R}_1$

In the case of $\mathfrak{R}_1$ the transfer operator $\mathcal{K}_1$ of (2.2) has the form

$$\mathcal{K}_1 = A_1(x_1, x_2)A_2(y_1, y_2)\hat{Q}, \quad \hat{Q} := \begin{pmatrix} 1 + L(\bar{x}, \bar{y})/W^2 & -1/W^2 \\ -L(\bar{x}, \bar{y}) & 1 \end{pmatrix} \tag{4.1}$$

with some explicit function L which does not depend on W and whose concrete form is not important for us now. Operators $A_{1,2}$ (the same as for $\mathfrak{R}_0$) contain a large parameter W in the exponent, hence only $W^{-1/2}$- neighbourhood of the stationary point gives the main contribution. The spectral analysis of A_1 gives us that

$$\begin{aligned} & A_1 \sim e^{\xi g_+(E)/N} A_+, \quad A_2 \sim A_+, \\ & A_+(x, y) = (2\pi)^{-1/2} W^2 e^{-W^2(x-y)/2 + c_+(x^2+y^2)/2}, \quad c_+ = 1 + a_+^{-2}, \\ & g_+(E) = (-E + i\sqrt{4 - E^2})/2. \end{aligned} \tag{4.2}$$

Then since

$$\lambda_j(A_+) = \Big(1 + \frac{2\alpha_+}{W} + \frac{c_+}{W^2}\Big)^{-1/2-j}, \tag{4.3}$$

$$\alpha_+ = \sqrt{\frac{c_+}{2}}\Big(1 + \frac{c_+}{2W^2}\Big)^{1/2}, \tag{4.4}$$

we obtain that the spectral gap for $A_{1,2}$ is of the order $W^{-1} \gg N^{-1}$, hence one could expect that A_1^{N-1} converges in the strong vector topology to the projection

$$A_{1,2}^{N-1} \to \lambda_0^{N-1}(A_1)\psi_0 \otimes \psi_0^*$$

where

$$A_1\psi_0 = \lambda_0(A_1)\psi_0, \quad A_1^*\psi_0^* = \overline{\lambda_0(A_1)}\psi_0^*.$$

The entry Q_{12} here is small hence the main order of our operator contains the Jordan cell. A simple computation shows that if we just replace in (4.1) $A_{1,2}$ by A_+ and Q_{12} by 0, then the answer will be wrong. Hence one should apply more refine analysis. An important point of such analysis is an application of the "gauge" transformation of $\mathcal{K}_1$ with matrix T

$$\begin{aligned} & \mathcal{K}_1 \to \mathcal{K}_{1T} = T\mathcal{K}_1 T^{-1} = A_1 A_2 \hat{S}, \quad \hat{S} = T\hat{Q}T^{-1}; \\ & T = \begin{pmatrix} 0 & W^{-1/2} \\ W^{1/2} & 0 \end{pmatrix}, \quad \hat{S} = \begin{pmatrix} 1 & -L/W \\ -1/W & 1 + L/W^2 \end{pmatrix}. \end{aligned} \tag{4.5}$$

With this transformation it can be shown that for any W

$$\lambda_0(\mathcal{K}_{1T}) = e^{\xi g_+(E)/N}(1 + O(n^{-2})), \quad |\lambda_1(\mathcal{K}_{1T})| \leq 1 - c/W, \qquad c > 0.$$

Hence for any $1 \ll W \ll N$ we get that $(\mathcal{K}_{1T})^{N-1}$ converges in the strong vector topology to the projection (non-orthogonal) on the eigenvector, corresponding to $\lambda_0(\mathcal{K}_{1T})$. This gives

Theorem 4.1. *Let H be 1d Gaussian RBM defined in (1.11) with $N \geq C_0 W \log W$, and let $|E| \leq 4\sqrt{2}/3 \approx 1.88$.*

$$\mathcal{R}_1(E + \xi/N, E) \to e^{\xi g(E)}, \qquad \Big|\frac{\partial}{\partial \xi}\mathcal{R}_1(E + \xi/N, E)\Big|_{\xi=0} - g_+(E)\Big| \leq C/W.$$

The second relation implies that

$$|\bar{\rho}_N(E) - \rho(E)| \leq C/W, \tag{4.6}$$

where $\bar{\rho}_N(E) = R_1(E)$ is the first correlation function, and $\rho(E)$ is defined in (1.2).

Remark 4.1. *The statement is expected to be true for all $|E| < 2$. The condition $|E| \leq 4\sqrt{2}/3 \approx 1.88$ is technical, and it can be removed by the proper deformation of the integration contour in the integral representation.*

Theorem 4.1 yields, in particular, that for $g_N(E + i\varepsilon)$ (the Stieltjes transform of the first correlation function $\bar{\rho}_N(E)$) and $g(E + i\varepsilon)$ (the Stieltjes transform of $\rho(E)$) we have

$$|\bar{g}_N(E + i\varepsilon) - g(E + i\varepsilon)| \leq C/W \tag{4.7}$$

uniformly in any arbitrary small $\varepsilon \geq 0$. As it was mentioned above, similar asymptotics (with correction C/W^2) for RBM of (1.11) in 3d was obtained in Disertori, Pinson, and Spencer [2002] and in 2d was obtained in Disertori and Lager [2017] (by the same techniques), however their method cannot be directly applied to 1d case since it essentially uses the Fourier analysis which is different in 1d. All other previous results about the density of states for RBM deal with $\varepsilon \gg W^{-1}$ or bigger (for fixed $\varepsilon > 0$ the asymptotics (4.7) follows from the results of Bogachev, Molchanov, and Pastur [1991]; Erdős and Knowles [2011] gives (4.7) with $\varepsilon \gg W^{-1/3}$; Sodin [2011] yields (4.7) for 1d RBM with Bernoulli elements distribution for $\varepsilon \geq W^{-0.99}$, and Erdős, Yau, and Yin [2012] proves similar to (4.7) asymptotics with correction $1/(W\varepsilon)^{1/2}$ for $\varepsilon \gg 1/W$). On the other hand, the methods of Erdős and Knowles [2011], Erdős, Yau, and Yin [2012] allow to control $N^{-1}\mathrm{Tr}\,(E + i\varepsilon - H_N)^{-1}$ and $(E + i\varepsilon - H_N)^{-1}_{xy}$ for $\varepsilon \gg W^{-1}$ without expectation, which gives some information about the localization length. This cannot be obtained from Theorem 4.1, since it requires estimates on $\mathbb{E}\{|(E + i\varepsilon - H_N)^{-1}_{xy}|^2\}$.

5 Analysis of $\mathfrak{R}_2$ for the block RBM

5.1 Sigma-model approximation for $\mathfrak{R}_2$ for the block RBM. We start from the analysis of so-called sigma-model approximation for the model (1.13) – (1.14). Sigma-model approximation is often used by physicists to study complicated statistical mechanics systems. In such approximation spins take values in some symmetric space (± 1 for Ising model, S^1 for the rotator, S^2 for the classical Heisenberg model, etc.). It is expected that sigma-models have all the qualitative physics of more complicated models with the same symmetry (for more details see, e.g., Spencer [2012]). The sigma-model approximation for RBM was introduced by Efetov (see Efetov [1997]), and the spins there are 4×4 matrices with both complex and Grassmann entries (this approximation was studied in Fyodorov and Mirlin [1991], Fyodorov and Mirlin [1994]). Let us mention also the paper Disertori, Spencer, and Zirnbauer [n.d.], where the average conductance for 1d Efetov's sigma-model for RBM was computed.

In the subsection we present rigorous results on the derivation of the sigma-model approximation for 1d RBM and the analysis of the model in the delocalization regime. The results are published in M. Shcherbina and T. Shcherbina [2018].

To derive a sigma-model approximation for the model (1.13) – (1.14), we take α in (1.14) $\alpha = \beta/W$, i.e. put

$$J = 1/W + \beta\Delta/W^2, \quad \beta > 0, \tag{5.1}$$

fix β and n, and consider the limit $W \to \infty$, for the generalized correlation functions

$$\begin{aligned} \mathfrak{R}^{+-}_{Wn\beta}(E, \varepsilon, \xi) &= \mathbf{E}\Big\{\frac{\det(H - z_1)\det(H - \overline{z}_2)}{\det(H - z_1')\det(H - \overline{z}_2')}\Big\}, \\ \mathfrak{R}^{++}_{Wn\beta}(E, \varepsilon, \xi) &= \mathbf{E}\Big\{\frac{\det(H - z_1)\det(H - z_2)}{\det(H - z_1')\det(H - z_2')}\Big\} \end{aligned} \tag{5.2}$$

for $\xi = (\xi_1, \xi_2, \xi_1', \xi_2')$.

Theorem 5.1. *Given $\mathfrak{R}^{+-}_{Wn\beta}$ of (5.2) ,(1.13) and (5.1), with any dimension d, any fixed β, $|\Lambda|$, $\varepsilon > 0$, and $\xi = (\xi_1, \bar{\xi}_2, \xi_1', \bar{\xi}_2') \in \mathbb{C}^4$ ($|\Im \xi_j| < \varepsilon \cdot \rho(E)/2$) we have, as $W \to \infty$:*

$$\mathfrak{R}^{+-}_{Wn\beta}(E, \varepsilon, \xi) \to \mathfrak{R}^{+-}_{n\beta}(E, \varepsilon, \xi), \quad \frac{\partial^2 \mathfrak{R}^{+-}_{Wn\beta}}{\partial \xi_1' \partial \xi_2'}(E, \varepsilon, \xi) \to \frac{\partial^2 \mathfrak{R}^{+-}_{n\beta}}{\partial \xi_1' \partial \xi_2'}(E, \varepsilon, \xi), \tag{5.3}$$

where $\mathfrak{R}^{+-}_{n\beta}(E, \varepsilon, \xi) = C_{E,\xi} \int \exp\Big\{\frac{\tilde{\beta}}{4} \sum \operatorname{Str} Q_j Q_{j-1} - \frac{c_0}{2|\Lambda|} \sum \operatorname{Str} Q_j \Lambda_{\xi,\varepsilon}\Big\} dQ,$

$\tilde{\beta} = (2\pi\rho(E))^2\beta$, $U_j \in \mathring{U}(2)$, $S_j \in \mathring{U}(1,1) = U(1,1)/U(1) \times U(1)$,

$$C_{E,\xi} = e^{E(\xi_1+\xi_2-\xi_1'-\xi_2')/2\rho(E)}, \quad \rho(E) = (2\pi)^{-1}\sqrt{4-E^2},$$

and Q_j are 4×4 supermatrices with commuting diagonal and anticommuting off-diagonal 2×2 blocks

$$Q_j = \begin{pmatrix} U_j^* & 0 \\ 0 & S_j^{-1} \end{pmatrix} \begin{pmatrix} (I+2\hat{\rho}_j\hat{\tau}_j)L & 2\hat{\tau}_j \\ 2\hat{\rho}_j & -(I-2\hat{\rho}_j\hat{\tau}_j)L \end{pmatrix} \begin{pmatrix} U_j & 0 \\ 0 & S_j \end{pmatrix}, \tag{5.4}$$

$$dQ = \prod dQ_j, \quad dQ_j = (1-2n_{j,1}n_{j,2})\, d\rho_{j,1} d\tau_{j,1}\, d\rho_{j,2} d\tau_{j,2}\, dU_j\, dS_j$$

with

$$n_{j,1} = \rho_{j,1}\tau_{j,1}, \quad n_{j,2} = \rho_{j,2}\tau_{j,2},$$
$$\hat{\rho}_j = \operatorname{diag}\{\rho_{j1}, \rho_{j2}\}, \quad \hat{\tau}_j = \operatorname{diag}\{\tau_{j1}, \rho_{j2}\}, \quad L = \operatorname{diag}\{1,-1\}$$

Here $\rho_{j,l}$, $\tau_{j,l}$, $l = 1,2$ are anticommuting Grassmann variables,

$$\operatorname{Str}\begin{pmatrix} A & \sigma \\ \eta & B \end{pmatrix} = \operatorname{Tr} A - \operatorname{Tr} B,$$

and

$$\Lambda_{\xi,\varepsilon} = \operatorname{diag}\{\varepsilon - i\xi_1/\rho(E), -\varepsilon - i\xi_2/\rho(E), \varepsilon - i\xi_1'/\rho(E), -\varepsilon - i\xi_2'/\rho(E)\}.$$

Theorem 5.2. *Given $\mathfrak{R}^{++}_{Wn\beta}$ of (5.2) ,(1.13) and (5.1), with any dimension d, any fixed β, $|\Lambda|$, $\varepsilon > 0$, and $\xi = (\xi_1, \xi_2, \xi_1', \xi_2') \in \mathbb{C}^4$ ($|\Im\xi_j| < \varepsilon \cdot \rho(E)/2$) we have, as $W \to \infty$:*

(5.5)

$$\mathfrak{R}^{++}_{Wn\beta}(E,\varepsilon,\xi) \to e^{ia_+(\xi_1'+\xi_2'-\xi_1-\xi_2)/\rho(E)},$$
$$\frac{\partial^2 \mathfrak{R}^{++}_{Wn\beta}}{\partial\xi_1'\partial\xi_2'}(E,\varepsilon,\xi) \to -a_+^2/\rho^2(E)\cdot e^{ia_+(\xi_1'+\xi_2'-\xi_1-\xi_2)/\rho(E)}, \; a_+ = (iE+\sqrt{4-E^2})/2.$$

Note that $Q_j^2 = I$ for Q_j of (5.4) and so the integral in the r.h.s of (5.3) is a sigma-model approximation similar to Efetov's one (see Efetov [1997]).

The kernel of the transfer operator for $\mathfrak{R}_2^{(\sigma)}$ has a form

$$\mathcal{K}_2^{(\sigma)} = \hat{F}\hat{Q}\hat{F}$$

where $\hat{F}$ and $\hat{Q}$ are 6×6 matrix kernels, such that $\hat{F}_{\mu\nu}$ are the operators of the multiplication by some function of U, S and $\hat{Q}_{\mu\nu} = \hat{Q}_{\mu\nu}(U_1U_2^*, S_1S_2^{-1})$ are the "difference" operators.

After some asymptotic analysis $\mathcal{K}_2^{(\sigma)}$ and some "gauge" transformation similar to (4.5) we obtain that $T\mathcal{K}_2^{(\sigma)}T$ can be replaced by the 4×4 "effective" matrix kernel

$$T\mathcal{K}_2^{(\sigma)}T \sim \tilde{F}\hat{K}_0\tilde{F}, \tag{5.6}$$

$$\hat{K}_0 = \begin{pmatrix} K & \tilde{K}_1 & \tilde{K}_2 & \tilde{K}_3 \\ 0 & K & 0 & \tilde{K}_2 \\ 0 & 0 & K & \tilde{K}_1 \\ 0 & 0 & 0 & K \end{pmatrix}, \quad \tilde{F} = F\begin{pmatrix} 1 & \tilde{F}_1 & \tilde{F}_2 & \tilde{F}_1\tilde{F}_2 \\ 0 & 1 & 0 & \tilde{F}_2 \\ 0 & 0 & 1 & \tilde{F}_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $K = K_U \otimes K_S$,

$$K_U(U_1, U_2) \sim \beta e^{-\beta|(U_1U_2^*)_{12}|^2}, \;\; K_S(S_1, S_2) \sim \beta e^{-\beta|(S_1S_2^{-1})_{12}|^2},$$

$\tilde{K}_i = \tilde{K}_i(U_1U_2^*; S_1S_2^{-1})$, F is an operator of multiplication by $e^{\varphi(U,S)/2n}$, and $\tilde{F}_{1,2}$ are operators of multiplication by $n^{-1}\varphi_{1,2}(U, S)$ with some specific φ, φ_1 and φ_2. An important feature of $\tilde{K}_i$ that they satisfy the operator bound

$$|\tilde{K}_i| \le C\beta^{-1}(\Delta_U + \Delta_S),$$

where Δ_U, Δ_S are the Laplace operator on the correspondent groups (see e.g. (3.16) for the definition of Δ_U). The bounds imply that for sufficiently smooth function f $\tilde{K}_i f \sim \beta^{-1}$.

Similarly to Section 3 the idea is to show that in the regime $\beta \gg n$

$$\tilde{F}\hat{K}_0\tilde{F} \sim \tilde{F}^2.$$

Then we get

$$\begin{aligned} \mathfrak{R}^{+-}_{n\tilde{\beta}}(E, \varepsilon, \xi) &= \frac{C_E^*}{2\pi i}\oint_{\omega_A} z^{n-1}(\widehat{G}_0(z)\widehat{f}, \widehat{g})dz + o(1) = C_E^*(\widehat{F}^{2n-2}\widehat{f}, \widehat{g}) + o(1) \\ &= C_E^*\int \big(4n^2F_1F_2 - 2\big)F^{2n}dUdS + o(1), \end{aligned}$$

where

$$C_E^* = e^{-g_+(E)(\xi_1+\xi_1'-\xi_2-\xi_2')/\rho(E)}, \quad g_+(E) = (-E + i\sqrt{4 - E^2})/2. \tag{5.7}$$

This relation allows us to prove

Theorem 5.3. *If $n,\beta\to\infty$ in such a way that $\beta > Cn\log^2 n$, then for any fixed $\varepsilon > 0$ and $\xi = (\xi_1,\xi_2,\xi_1',\xi_2')\in\mathbb{C}^4$ ($|\Im\xi_j| < \varepsilon\cdot\rho(E)/2$) we have*

$$\mathfrak{R}^{+-}_{n\beta} \to C_E^*\Big(\frac{\delta_1\delta_2}{\alpha_1\alpha_2}(e^{2c_0\alpha_1}-1)-\frac{\delta_1+\delta_2}{\alpha_2}e^{2c_0\alpha_1}+e^{2c_0\alpha_1}\frac{\alpha_1}{\alpha_2}\Big), \tag{5.8}$$

$$\text{where}\quad \alpha_1 = \varepsilon - i(\xi_1-\xi_2)/2\rho(E),\quad \alpha_2 = \varepsilon - i(\xi_1'-\xi_2')/2\rho(E), \tag{5.9}$$
$$\delta_1 = i(\xi_1'-\xi_1)/2\rho(E),\quad \delta_2 = i(\xi_2-\xi_2')/2\rho(E),$$

and C_E^* is defined in (5.7).

Theorem 5.3 combined with Theorem 5.2 gives the GUE type behaviour for the spectral correlation function:

Theorem 5.4. *In the dimension $d = 1$ the behaviour of the sigma-model approximation of the second order correlation function (5.2) of (1.13), (5.1), as $\beta\gg n$, in the bulk of the spectrum coincides with those for the GUE. More precisely, if $\Lambda = [1,n]\cap\mathbb{Z}$ and H_N, $N = Wn$ are matrices (1.13) with J of (5.1), then for any $|E| < \sqrt{2}$ (1.4) holds in the limit first $W\to\infty$, and then $\beta,n\to\infty$, $\beta\ge Cn\log^2 n$.*

5.2 Analysis of $\mathfrak{R}_2$ for block RBM of (1.13)-(1.14). As it was mentioned in Section 2 in the case of $\mathfrak{R}_2$ the transfer operator $\mathcal{K}_2$ is a 70×70 matrix whose entries depend on 8 spacial variables $x_1,x_2,y_1,y_2;x_1',x_2',y_1',y_2'\in\mathbb{R}$, two unitary 2×2 matrix U,U', and two hyperbolic 2×2 matrix S,S', which acts in the direct sum of 70 Hilbert spaces $L_2(\mathbb{R}^4)\otimes L_2(\mathring{U}(2),dU)\otimes L_2(\mathring{U}(1,1),dS)$, where dU,dS are integrations with respect to the corresponding Haar measures. In general, the analysis of such operator is a very involved problem, unless there is a possibility to take into account some special features of the matrix kernel and to reduce it (in the sense of Definition 2.1) by some matrix kernel of smaller dimensionality.

In the case of $\mathcal{K}_2$ the first observation is that it can be factorised as

$$\mathcal{K}_2 = \hat{F}\hat{Q}\hat{A}\hat{F},$$

where $\hat{F}$, $\hat{Q}$ and $\hat{A}$ are 70×70 matrix kernels, such that $\hat{F}_{\mu\nu}$ are the operators of multiplication by some function of U,S,

$$\hat{Q}_{\mu\nu} = K_U K_S Q_{\mu\nu}(U(U')^*;S(S')^{-1}),$$
$$K_U = \alpha t W e^{-\alpha W t|(U(U')^*)_{12}|^2},\quad K_S = \alpha\tilde{t} W e^{-\alpha W\tilde{t}|(S(S')^{-1})_{12}|^2},$$

with $t,\tilde{t}$ defined similarly to (3.6) and functions $Q_{\mu\nu}$ which do not depend on W, and

$$\hat{A}_{\mu\nu} = A_1(x_1,x_1')A_2(y_1,y_1')A_3(x_2,x_2')A_4(y_2,y_2')\mathfrak{A}_{\mu,\nu}(\bar{x},\bar{x}',\bar{y},\bar{y}')$$

with $A_{1,2,3,4}$ being a scalar kernels similar to that for $\mathfrak{R}_0$ (see (3.9)) and functions $A_{\mu\nu}$ which do not depend on W. It is straightforward to prove that only $W^{-1/2}\log W$ -neighbourhoods of some stationary points in $\mathbb{R}^8$ give essential contributions. Further analysis shows that after some "gauge" transformation similar to (4.5) $T\mathcal{K}_2T^{-1}$ can be replaced (in the sense of Definition 2.1) by 4×4 effective kernel of the form similar to (5.6).

Remark that the analysis justifies the physics conjecture that the behaviour of the "generalized" correlation function $\mathfrak{R}_2$ for the model (1.13) – (1.14) and of its sigma-model approximation $\mathfrak{R}_2^\sigma$ of are very similar.

As a result we obtain (cf with Theorem 5.4)

Theorem 5.5. *In the dimension $d = 1$ the behaviour of the second order correlation function (1.6) of the model (1.13) – (1.14), as $W \gg n$, in the bulk of the spectrum coincides with those for the GUE. More precisely, for any $|E| < \sqrt{2}$ (1.4) holds in the limit $W, n \to \infty$ with $W/\log^2 W > Cn$.*

The theorem is the main result of the paper M. Shcherbina and T. Shcherbina [n.d.].

References

Zhigang Bao and László Erdős (2017). "Delocalization for a class of random block band matrices". *Probab. Theory Related Fields* 167.3-4, pp. 673–776. MR: 3627427 (cit. on p. 2711).

F. A. Berezin (1983). Введение в алгебру и анализ с антикоммутирующими переменными. With a preface by A. A. Kirillov. Moskov. Gos. Univ., Moscow, p. 208. MR: 732126.

L. V. Bogachev, S. A. Molchanov, and L. A. Pastur (1991). "On the density of states of random band matrices". *Mat. Zametki* 50.6, pp. 31–42, 157. MR: 1150631 (cit. on pp. 2706, 2719).

Paul Bourgade, László Erdős, Horng-Tzer Yau, and Jun Yin (2017). "Universality for a class of random band matrices". *Adv. Theor. Math. Phys.* 21.3, pp. 739–800. arXiv: 1602.02312. MR: 3695802 (cit. on p. 2708).

E. Brézin and S. Hikami (2001). "Characteristic polynomials of real symmetric random matrices". *Comm. Math. Phys.* 223.2, pp. 363–382. MR: 1864437 (cit. on p. 2709).

Edouard Brézin and Shinobu Hikami (2000). "Characteristic polynomials of random matrices". *Comm. Math. Phys.* 214.1, pp. 111–135. MR: 1794268 (cit. on p. 2709).

Giulio Casati, Luca Molinari, and Felix Izrailev (1990). "Scaling properties of band random matrices". *Phys. Rev. Lett.* 64.16, pp. 1851–1854. MR: 1046365 (cit. on p. 2706).

M. Disertori, H. Pinson, and T. Spencer (2002). "Density of states for random band matrices". *Comm. Math. Phys.* 232.1, pp. 83–124. MR: 1942858 (cit. on pp. 2710, 2711, 2719).

M. Disertori, T. Spencer, and M. R. Zirnbauer (n.d.). *Supersymmetric 1d sigma model*. Unpublished (cit. on p. 2720).

Margherita Disertori and Mareike Lager (2017). "Density of states for random band matrices in two dimensions". *Ann. Henri Poincaré* 18.7, pp. 2367–2413. MR: 3665217 (cit. on pp. 2711, 2719).

Konstantin Efetov (1997). *Supersymmetry in disorder and chaos*. Cambridge University Press, Cambridge, pp. xiv+441. MR: 1628498 (cit. on pp. 2711, 2720, 2721).

László Erdős and Antti Knowles (2011). "Quantum diffusion and eigenfunction delocalization in a random band matrix model". *Comm. Math. Phys.* 303.2, pp. 509–554. MR: 2782623 (cit. on pp. 2707, 2719).

László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin (2013). "Delocalization and diffusion profile for random band matrices". *Comm. Math. Phys.* 323.1, pp. 367–416. arXiv: 1205.5669. MR: 3085669 (cit. on p. 2708).

László Erdős, Horng-Tzer Yau, and Jun Yin (2012). "Bulk universality for generalized Wigner matrices". *Probab. Theory Related Fields* 154.1-2, pp. 341–407. arXiv: 1001.3453. MR: 2981427 (cit. on pp. 2707, 2708, 2719).

Yan V. Fyodorov and Alexander D. Mirlin (1991). "Scaling properties of localization in random band matrices: a σ-model approach". *Phys. Rev. Lett.* 67.18, pp. 2405–2409. MR: 1130103 (cit. on pp. 2706, 2707, 2720).

Yan V Fyodorov and Alexander D Mirlin (1994). "Statistical properties of eigenfunctions of random quasi 1d one-particle Hamiltonians". *International Journal of Modern Physics B* 8.27, pp. 3795–3842 (cit. on pp. 2711, 2720).

Lingyun Li and Alexander Soshnikov (2013). "Central limit theorem for linear statistics of eigenvalues of band random matrices". *Random Matrices Theory Appl.* 2.4, pp. 1350009, 50. MR: 3149439 (cit. on p. 2706).

S. A. Molchanov, L. A. Pastur, and A. M. Khorunzhiĭ (1992). "Distribution of the eigenvalues of random band matrices in the limit of their infinite order". *Teoret. Mat. Fiz.* 90.2, pp. 163–178. MR: 1182292 (cit. on p. 2706).

Ron Peled, Jeffrey Schenker, Mira Shamis, and Sasha Sodin (2016). "On the Wegner orbital model". *International Mathematics Research Notices* (cit. on p. 2707).

Jeffrey Schenker (2009). "Eigenvector localization for random band matrices with power law band width". *Comm. Math. Phys.* 290.3, pp. 1065–1097. MR: 2525652 (cit. on p. 2707).

Mariya Shcherbina (2015). "On fluctuations of eigenvalues of random band matrices". *J. Stat. Phys.* 161.1, pp. 73–90. MR: 3392508 (cit. on p. 2706).

Mariya Shcherbina and Tatyana Shcherbina (n.d.). *Transfer matrix approach to 1d random band matrices: universality regime*. In preparation (cit. on p. 2724).

– (2016). "Transfer matrix approach to 1d random band matrices: density of states". *J. Stat. Phys.* 164.6, pp. 1233–1260. MR: 3541181 (cit. on p. 2712).

Mariya Shcherbina and Tatyana Shcherbina (2017). "Characteristic polynomials for 1D random band matrices from the localization side". *Comm. Math. Phys.* 351.3, pp. 1009–1044. MR: 3623245 (cit. on pp. 2712, 2713, 2717).

– (2018). "Universality for 1d random band matrices: sigma-model approximation". *Journal of Statistical Physics*, pp. 1–38 (cit. on pp. 2712, 2720).

Tatyana Shcherbina (n.d.). *Characteristic polynomials for random band matrices near the threshold*. In preparation (cit. on pp. 2712, 2717).

– (2011). "On the correlation function of the characteristic polynomials of the Hermitian Wigner ensemble". *Comm. Math. Phys.* 308.1, pp. 1–21. MR: 2842968 (cit. on p. 2709).

– (2013). "On the correlation functions of the characteristic polynomials of the Hermitian sample covariance matrices". *Probab. Theory Related Fields* 156.1-2, pp. 449–482. MR: 3055265 (cit. on p. 2709).

– (2014a). "On the second mixed moment of the characteristic polynomials of 1D band matrices". *Comm. Math. Phys.* 328.1, pp. 45–82. MR: 3196980 (cit. on p. 2717).

– (2014b). "Universality of the local regime for the block band matrices with a finite number of blocks". *J. Stat. Phys.* 155.3, pp. 466–499. MR: 3192170 (cit. on p. 2711).

Sasha Sodin (2011). "An estimate for the average spectral measure of random band matrices". *J. Stat. Phys.* 144.1, pp. 46–59. MR: 2820034 (cit. on pp. 2708, 2719).

Thomas Spencer (2012). "SUSY statistical mechanics and random band matrices". In: *Quantum many body systems*. Vol. 2051. Lecture Notes in Math. Springer, Heidelberg, pp. 125–177. MR: 2953867 (cit. on pp. 2706, 2720).

Terence Tao and Van Vu (2011). "Random matrices: universality of local eigenvalue statistics". *Acta Math.* 206.1, pp. 127–204. MR: 2784665 (cit. on p. 2707).

N. Ja. Vilenkin (1968). *Special functions and the theory of group representations*. Translated from the Russian by V. N. Singh. Translations of Mathematical Monographs, Vol. 22. American Mathematical Society, Providence, R. I., pp. x+613. MR: 0229863 (cit. on p. 2716).

F. J. Wegner (1979). "Disordered system with n orbitals per site: $n \to \infty$ limit". *Phys. Rev. B* 19, pp. 783–792 (cit. on p. 2710).

Received 2017-12-10.

MARIYA SHCHERBINA
INSTITUTE FOR LOW TEMPERATURE PHYSICS
KHARKIV
UKRAINE

and

KARAZIN KHARKIV NATIONAL UNIVERSITY
KHARKIV
UKRAINE

Shcherbi@ilt.kharkov.ua

TATYANA SHCHERBINA
DEPARTMENT OF MATHEMATICS
PRINCETON UNIVERSITY
PRINCETON
USA
tshcherbyna@princeton.edu

Proc. Int. Cong. of Math. – 2018
Rio de Janeiro, Vol. 3 (2727–2750)

ON ‘CATEGORIES’ OF QUANTUM FIELD THEORIES

Yuji Tachikawa

Abstract

We give a rough description of the ‘categories’ formed by quantum field theories. A few recent mathematical conjectures derived from quantum field theories, some of which are now proven theorems, will be presented in this language.

1 Introduction

Studies of quantum field theories (QFTs) by physicists have led to various mathematical conjectures, some of which are investigated fruitfully by mathematicians. Existing mathematical formulations of QFTs do not, however, explain how these conjectures are arrived at in the first place. It seems to the author that more properties of QFTs as perceived by physicists can be formalized in a way that a better part of the process itself of conjuring of the conjectures become understandable to mathematicians.

For this purpose, it seems crucial to discuss not just individual QFTs but the interrelationship among them. In other words, we need to discuss the ‘categories’ formed by QFTs and possible operations in those categories. In this note, a rough description of these ‘categories’ will be given, and a few recent mathematical conjectures, some of which are now proven theorems, will be phrased in this language.*

Acknowledgments. The author thanks discussions with many fellow string theorists and helpful mathematicians. In particular, he thanks Tomoyuki Arakawa, Hiraku Nakajima and Kazuya Yonekura, who gave valuable comments on the manuscript.

The author is partially supported by JSPS KAKENHI Grant-in-Aid, Wakate-A, No.17H04837 and Kiban-S, No.16H06335 and by WPI Initiative, MEXT, Japan at IPMU, the University of Tokyo.
MSC2010: 81T99.

*This note is an abridged but updated version of a longer review article intended for mathematicians, which is available on the author’s webpage. The latest version of this manuscript will also be available on the arXiv.

2 The framework

2.1 QFTs for $\mathcal{S}$-structured manifolds. A category of QFT exists for each fixed spacetime dimension d and a structure $\mathcal{S}$ on manifolds. Here, the structure $\mathcal{S}$ can be e.g. smooth structure, Riemannian metric, conformal structure, spin structure, etc. We then denote by $\mathfrak{Q}^d_{\mathcal{S}}$ the category of QFT defined on $\mathcal{S}$-structured manifolds of dimension d. (We consider a Wick-rotated, Euclidean version of QFTs in this note.)

At the very basic level, an object $Q \in \mathfrak{Q}^d_{\mathcal{S}}$ assigns:

- a $\mathbb{C}$-vector space $\mathcal{H}_Q(N)$ called the space of states to each $(d-1)$-dimensional $\mathcal{S}$-structured manifold N without boundary,
- and the transition amplitude

  (2-1) $$Z_Q(M) : \mathcal{H}_Q(N) \to \mathcal{H}_Q(N')$$

  to each d-dimensional $\mathcal{S}$-structured manifold M with the incoming boundary N and outgoing boundary N'.

They are supposed to satisfy the standard axioms of Atiyah [1988] and Segal [2004], properly modified for the structure $\mathcal{S}$. In particular, for an empty set we demand $\mathcal{H}_Q(\varnothing) = \mathbb{C}$, and then if M is without boundary we simply have $Z_Q(M) \in \mathbb{C}$, called the partition function.

Note that a QFT Q determines a functor from a suitable bordism category to the category of vector spaces. Then $\mathfrak{Q}^d_{\mathcal{S}}$ is a category formed by those functors, but as morphisms we do not choose natural transformations between functors. We will come back to the question of morphisms in Section 2.5.

Traditionally, a QFT for smooth manifolds is called a topological QFT (despite the fact that smooth manifolds and topological manifolds can have interesting differences), a QFT for Riemannian manifolds are simply called a QFT (without adjective), a QFT for conformal structure is called a conformal field theory (CFT), etc.

If an $\mathcal{S}'$ structure on a manifold can be obtained by forgetting some data of an $\mathcal{S}$ structure, there is a functor $\mathfrak{Q}^d_{\mathcal{S}'} \to \mathfrak{Q}^d_{\mathcal{S}}$, obtained by evaluating the partition function by forgetting the additional structure on the manifold. For example, from Riemannian manifolds we can extract a smooth manifold. Correspondingly, a QFT for smooth manifolds can be considered as a QFT for Riemannian manifolds. Using the traditional language, a topological QFT is an example of a QFT.

From two objects $Q_{1,2} \in \mathfrak{Q}^d_{\mathcal{S}}$, one can form a product $Q_1 \times Q_2 \in \mathfrak{Q}^d_{\mathcal{S}}$, such that the partition function of $Q_1 \times Q_2$ is simply $Z_{Q_1 \times Q_2}(M) = Z_{Q_1}(M) Z_{Q_2}(M)$. We always have a trivial QFT triv $\in \mathfrak{Q}^d_{\mathcal{S}}$ which is the identity of this product. This makes $\mathfrak{Q}^d_{\mathcal{S}}$ a monoidal category.

2.2 Point operators of a QFT. Associated to a QFT $Q \in \mathcal{Q}^d_{\mathcal{S}}$ is a vector space $\mathcal{V}_Q$, called the space of point operators. An element of $\mathcal{V}_Q$ was traditionally just called an operator of the QFT. When $\mathcal{S}$ is the Riemannian structure, the Riemannian structure with spin structure, the conformal structure, $\mathcal{V}_Q$ has an action of the rotation group $\mathrm{SO}(d)$, its double cover $\mathrm{Spin}(d)$, or the conformal group $\mathrm{SO}(d+1,1)$, respectively. Let us denote by $(\mathcal{V}_Q)^{\mathrm{inv}}$ the subspace invariant under these groups. Given a d-dimensional manifold M and point operators $\varphi_i \in (\mathcal{V}_Q)^{\mathrm{inv}}$ the QFT associates a complex number we denote as

$$Z_Q(M;\varphi_1(x_1)\varphi_2(x_2)\cdots\varphi_n(x_n)) \in \mathbb{C} \tag{2-2}$$

for distinct points $x_i \in M$, in a way multi-linear in φ_i. This number is called the correlation function or the n-point function of the theory. We can extend this construction to arbitrary elements of $\mathcal{V}_Q$ by considering a suitable bundle over M^n.

In the traditional axiomatic quantum field theory, one considers $M = \mathbb{R}^d$ and these are the (Euclidean version of) Wightman functions. In a unitary theory we impose the reflection positivity.

When the structure $\mathcal{S}$ is the conformal structure, there is a natural isomorphism

$$\mathcal{V}_Q \simeq \mathcal{H}_Q(S^{d-1}), \tag{2-3}$$

which is called the state operator correspondence. The action of the dilatation $x \mapsto ax \in \mathbb{R}^d$ in the conformal group on $\mathcal{V}_Q$ is usually written as $a^{-\Delta}$. In a unitary theory Δ is positive semidefinite and gives a grading of $\mathcal{V}_Q$. Its eigenvalues are called the scaling dimension. Take two operators $\varphi_{1,2} \in \mathcal{V}_Q$ with scaling dimension $\Delta_{1,2}$. It can be argued that the two-point function behaves as

$$Z_Q(M;\varphi_1(x_1)\varphi_2(x_2)) \lesssim \text{const.}|x_1 - x_2|^{-(\Delta_1+\Delta_2)} \qquad \text{when } |x_1 - x_2| \to 0. \tag{2-4}$$

$\mathcal{V}_Q$ has a structure of a certain generalized kind of an algebra. When $d = 2$, it is essentially given by the axioms of the vertex operator algebras, but with both holomorphic and antiholomorphic dependence. It should not be too difficult to write a similar set of axioms for $d > 2$. The algebra structure is known to physicists under the name of operator product expansion (OPE) algebra.

When the structure $\mathcal{S}$ is the Riemannian structure we can introduce a filtration on $\mathcal{V}_Q$ by $\mathbb{R}_{\geq 0}$, still called the scaling dimension, by demanding that the above inequality holds.

In either case, it is a general feature of the n-point function that it diverges as the points approach each other. This in particular means that the elements of $\mathcal{V}_Q$ are operator-valued distributions, i.e. distributions which take values in the space of unbounded operators on $\mathcal{H}_Q(\mathbb{R}^{d-1})$. This makes their analysis rather complicated. In the standard algebraic QFT approach, see e.g. Haag [1996], one instead considers the net of algebras of bounded operators constructed out of $\mathcal{V}_Q$. It should also be possible to construct such a net from a QFT in our sense. That said, in various other applications of QFTs to mathematics, n-point functions themselves are used. For example, the Donaldson invariants are examples of n-point functions of a suitable gauge theory, as interpreted by Witten [1988]. Because of this, the author would like to keep $\mathcal{V}_Q$ as part of the data defining a QFT.

2.3 Deformations of a QFT. Although we have not completely defined what a QFT is, it should be possible to consider a family of QFTs parameterized by an arbitrary space $\mathcal{M}$. In the category $\mathcal{Q}^d_{\mathcal{S}}$ where $\mathcal{S}$ is the Riemannian structure, a fundamental fact is that given $Q_0 \in \mathcal{Q}^d_{\mathcal{S}}$, one can construct a certain universal family of QFTs parameterized by $\mathcal{M}_{\text{relevant}}$ such that

- $0 \in \mathcal{M}_{\text{relevant}}$ corresponds to Q_0
- $T_0\mathcal{M}_{\text{relevant}} \simeq$ the subspace of $(\mathcal{V}_Q)^{\text{inv}}$ whose scaling dimension is $< d$.

This family $\mathcal{M}_{\text{relevant}}$ is called the family of relevant deformations of Q_0.

The idea is that, given an element $\varphi \in$ the $\mathrm{SO}(d)$-invariant part of $\mathcal{V}_Q$, we try to define the deformed theory $Q(\lambda\varphi)$ for a small λ by the formula

$$Z_{Q(\lambda\varphi)}(M) := \text{“}Z_Q(M; e^{\int_M \lambda\varphi(x) d\mu})\text{”} \tag{2-5}$$

where the right hand side is, at least in an extremely naive level, defined by expanding in λ and writing it in terms of a sum of the n-point functions of $\varphi(x)$. The singularities in the integral need to be dealt with, and the convergence of the series needs to be proven. But physicists think that it should be possible to make sense of it when the scaling dimension Δ of φ is $< d$. The author believes that it should be possible to prove this by generalizing results already available in the study of constructive QFT.

It is also a common belief among physicists that even when $\Delta = d$, the deformation should always make sense as a formal power series in λ. This part should also be provable by generalizing results already available in the mathematical study of perturbative QFTs.These deformations with $\Delta = d$ are called marginal deformations.

It also happens that for some subset of operators with $\Delta = d$ one can have an actual family, not just in the sense of formal power series. Such deformations are called exactly marginal deformations, and are of great interest to physicists.

2.4 G-symmetric QFTs. Given a structure $\mathcal{S}$ and a group G, we can consider a new structure $\mathcal{S} \times G$ which means that the manifolds come with $\mathcal{S}$ structure together with a G-bundle with connection. Let us introduce a special notation $\mathcal{Q}^d_{\mathcal{S}}(G) := \mathcal{Q}^d_{\mathcal{S}\times G}$, an object of which is called a G-symmetric $\mathcal{S}$-structured QFT.

Given a homomorphism $\varphi : G \to G'$, we have a functor $\varphi^* : \mathbb{Q}^d_{\mathcal{S}}(G') \to \mathbb{Q}^d_{\mathcal{S}}(G)$ defined in an obvious manner. Similarly, given $Q_1 \in \mathbb{Q}^d_{\mathcal{S}}(G_1)$ and $Q_2 \in \mathbb{Q}^d_{\mathcal{S}}(G_2)$, we have $Q_1 \times Q_2 \in \mathbb{Q}^d_{\mathcal{S}}(G_1 \times G_2)$.

We see that these categories behave like categories of spaces with G action. In the latter case, we can sometimes construct from a space X with $G \times F$ action a quotient space X/G with F action, if the action of G is sufficiently mild. There is a similar construction in the categories of QFTs. Physicists call this operation the gauging of Q by G.

Namely, from a $G \times F$-symmetric QFT $Q \in \mathbb{Q}^d_{\mathcal{S}}(G \times F)$, one can sometimes gauge G and construct $Q \not{+} G \in \mathbb{Q}^d_{\mathcal{S}}(F)$. The idea is to define

$$Z_{Q\not{+}G}(M, A_F) = \text{“}\int_{\mathfrak{M}_{G,M}} Z_Q(M, A_G, A_F) d\mu\text{”} \tag{2-6}$$

where $\mathfrak{M}_{G,M}$ is the space of G-bundles with connections on the manifold M, $d\mu$ is a suitable measure on it, and $A_G \in \mathfrak{M}_{G,M}$ is a specific bundle with connection.

The problem is how to make this idea precise. When G is a finite group, or when $d = 1$ and G is compact, there is no problem, since $\mathfrak{M}_{G,M}$ is finite-dimensional and there is a suitable measure. Otherwise it is an extremely hard problem. Making it precise when $\mathcal{S}$ is the Riemannian structure, $d = 4$, $Q =$ triv, G being a compact Lie group, is a big part of one of the Millennium problems Jaffe and Witten [2006]. That said, physicists share a broad consensus on the condition on d and Q for which the gauging by a compact Lie group G makes sense. It is generally believed that there should not be a problem when $d = 2$ or 3, that it is generically impossible when $d \geq 5$, and that a simple criterion on Q is agreed upon when $d = 4$. It should also be noted that the gauged theory $Q \not{+} G$, when it exists, come in a family parameterized by what is called the gauge coupling constant, with no distinguished origin in the parameter space.

The notation $\not{+}G$ is not at all standard but was coined for the purpose of this note. This is supposed to give the impression that the gauging adds ('$+G$') the degrees of freedom of the gauge fields but at the same time it reduces the degrees of freedom by dividing ('$/G$') by the gauge group.

2.5 Submanifold operators and morphisms of a QFT. In general, associated to a QFT Q, we not only have the space of point operators $\mathcal{V}^0_Q := \mathcal{V}_Q$ discussed already, but we should also have the collection of line operators $\mathcal{V}^1_Q$, the collection of surface operators $\mathcal{V}^2_Q$, ..., up to the collection of codimension-1 operators $\mathcal{V}^{d-1}_Q$. For example, a gauge theory $Q \not{+} G$ naturally has line operators labeled by a representation R of G, such that given an embedded circle $C : S^1 \to M$ we can consider

$$Z_{Q\not{+}G}(M; R(C)) := \text{“}\int_{\mathfrak{M}_{G,M}} Z_Q(M, A) \operatorname{tr} \operatorname{Hol}_R(C) d\mu\text{”} \tag{2-7}$$

where Hol is the holonomy of the G-connection A. These are called the Wilson line operators by physicists. In this case, the set of labels of Wilson lines is given by $\mathrm{Rep}(G)$, and forms a tensor category.

From this example and others, it is reasonable to think that for a QFT Q, $\mathcal{V}_Q^0$ of point operators forms a kind of algebra, $\mathcal{V}_Q^1$ of line operators forms a kind of tensor category, $\mathcal{V}_Q^2$ of surface operators forms a 2-category of some sort, ..., $\mathcal{V}_Q^{d-1}$ of codimension-1 operators forms a $(d-1)$ category. The codimension-1 operators are somewhat special, since a codimension-1 locus N in the spacetime M can split M into two disconnected regions M_1 and M_2. Therefore, we can think of a situation where we have a QFT Q_1 on M_1, another QFT Q_2 on Q_2, and a codimension-1 operator X between the two. We consider X to be a morphism from Q_1 to Q_2: $X \in \mathrm{Hom}(Q_1, Q_2)$, and $\mathcal{V}_Q^{d-1} = \mathrm{Hom}(Q, Q)$. A codimension-2 locus can separate a codimension-1 region into two regions, supporting the morphisms $X, Y \in \mathrm{Hom}(Q_1, Q_2)$, respectively. Then such a codimension-2 operator is a morphism between morphisms, and objects in $\mathcal{V}_Q^{d-2}$ are special cases: they are morphisms between the trivial morphism in $\mathrm{Hom}(Q, Q)$. This relation goes down recursively to the case of point operators. For topological QFTs, the resulting categorical structure of the submanifold operators are discussed in the literature under the name of the fully-extended topological QFTs, see e.g. Kapustin [2010], Freed [2013], and Carqueville, Meusburger, and Schaumann [2016].

Note that, given a d'-dimensional QFT Q' and a d-dimensional QFT Q with $d' < d$, we can tautologically consider placing Q' on a dimension-d' submanifold $M' \subset M$, by defining

$$Z_Q(M; Q'(M')) := Z_Q(M) Z_{Q'}(M'). \tag{2-8}$$

In particular, take Q to be the trivial QFT $\mathrm{triv}_{\mathcal{S}}^d \in \mathfrak{Q}_{\mathcal{S}}^d$. We can place any $Q' \in \mathfrak{Q}_{\mathcal{S}}^{d-1}$ on codimension-1 subspaces. In other words, any $(d-1)$-dimensional QFT Q' is a morphism from the trivial theory in dimension d to itself:

$$\mathfrak{Q}_{\mathcal{S}}^{d-1} = \mathrm{Hom}(\mathrm{triv}_{\mathcal{S}}^d, \mathrm{triv}_{\mathcal{S}}^d) = \mathcal{V}_{\mathrm{triv}_{\mathcal{S}}^d}^{d-1}. \tag{2-9}$$

Therefore, a full understanding of the trivial theory in d dimensions in this sense entails a full understanding of all QFTs in $(d-1)$-dimensions.

2.6 Compactifications of QFTs. In the discussions above of the submanifold operators, we saw that QFTs in different spacetime dimensions are intimately related. There is also another way to relate QFTs in different dimensions. Pick a QFT $Q \in \mathfrak{Q}_{\mathcal{S}}^d$, and fix a d'-dimensional manifold M' with $\mathcal{S}$-structure. Then, we define a $(d-d')$-dimensional QFT $Q\langle M' \rangle$ by demanding

$$Z_{Q\langle M' \rangle}(M) = Z_Q(M \times M'). \tag{2-10}$$

This operation is called the compactification of Q by M' by physicists.

2.7 Anomalous and meta QFTs. So far we have been talking about what can be called ‘genuine’ QFTs Q, where the partition function $Z_Q(M)$ takes values in $\mathbb{C}$. There are, however, many ‘anomalous’ QFTs whose partition function does not take values in $\mathbb{C}$ but only in a one-dimensional $\mathbb{C}$-vector space.

To specify a d-dimensional anomalous QFT $\tilde{Q}$ with structure $\mathcal{S}$, one first needs to give a rule assigning one-dimensional vector spaces to $\mathcal{S}$-structured d-dimensional manifolds M. This can conveniently done by taking a $(d+1)$-dimensional QFT $\mathfrak{A} \in \mathbb{Q}_{\mathcal{S}}^{d+1}$ whose Hilbert space $\mathcal{H}_{\mathfrak{A}}(M)$ on any d-dimensional manifold M is one dimensional. Such a theory $\mathfrak{A}$ is called invertible, and we demand

$$Z_{\tilde{Q}}(M) \in \mathcal{H}_{\mathfrak{A}}(M). \tag{2-11}$$

The $(d+1)$-dimensional QFT $\mathfrak{A}$ is called the anomaly of the anomalous d-dimensional QFT $\tilde{Q}$. Equivalently, an anomalous d-dimensional theory $\tilde{Q}$ is a morphism from a trivial $(d+1)$-dimensional theory triv $\in \mathbb{Q}_{\mathcal{S}}^{d+1}$ to an invertible theory $\mathfrak{A} \in \mathbb{Q}_{\mathcal{S}}^{d+1}$, i.e. $\tilde{Q} \in \mathrm{Hom}(\mathrm{triv}_{\mathcal{S}}^{d+1}, \mathfrak{A})$. A genuine QFT is a special case where $\mathfrak{A}$ is also trivial.

Once we make this generalization, it is an easy step to consider also meta QFTs in d-dimensions: a meta QFT $\hat{Q}$ is such that its partition function $Z_{\hat{Q}}(M)$ takes values in a finite-dimensional Hilbert space $\mathcal{H}_{\mathcal{T}}(M)$ of a $(d+1)$-dimensional theory $\mathcal{T}$. One important example is the theory of conformal blocks of affine Lie algebras, for which $\mathcal{T}$ is the 3d Chern-Simons theory; another is the 6d $\mathcal{N}=(2,0)$ superconformal theories which will be discussed below. Meta QFTs are called relative QFTs by mathematicians Freed and Teleman [2012].

2.8 Supersymmetric QFTs. Mathematical conjectures often arose from the study of supersymmetric QFTs, in various dimensions. In the framework of this note, a supersymmetric QFT in d dimensions is a QFT for a particular structure $\mathcal{S}$ extending the Riemannian structure. Similarly, a superconformal QFT in d dimensions is a QFT for a structure $\mathcal{S}$ extending the conformal structure. For example, for $d = 4$, both supersymmetric and superconformal QFTs come in four varieties, called $\mathcal{N}=1,2,3,4$ supersymmetric QFTs and superconformal QFTs, respectively.

Unfortunately, it seems difficult to give a concise definition of what a supersymmetric structure on a manifold is, because of the following reason. Let us first consider the case of a superconformal structure. A QFT Q with a superconformal structure would have an action of a superconformal group on its space $\mathcal{V}_Q$ of point operators. The Lie algebra of a superconformal group is a super Lie algebra such that its even component contains the conformal algebra $\mathfrak{so}(d+1,1)$ and its odd component is in a spinor representation of the conformal group. The fact that the odd component is in a spinor representation is required from the spin-statistics theorem of the unitary QFT. One can also argue that the super Lie algebra in question is simple. Then it is straightforward to list all possible superconformal algebras compatible with unitarity, and one finds that the maximum possible dimension is $d = 6$ Nahm [1978]. Similarly, one finds that the maximum possible dimension for supersymmetric structures is $d = 11$.

Therefore there is no hope of formulating the supersymmetric structures in a way analogous to the Riemannian structure such that the formulation applies to arbitrary dimensions. Their existence is accidental to low dimensions in an intrinsic way.

That said, for possible dimensions, $d \leq 6$ for superconformal theories and $d \leq 11$ for supersymmetric theories, there are huge amount of literature on the physics side of the community about the superconformal/supersymmetric structures on a manifold, under the name of $\mathcal{N}$-extended supergravity in various dimensions.

3 Examples

Currently, there are many examples of QFTs which are known to physicists. Broadly speaking, there are three methods of constructions, with overlapping range of applicabilities. Let us examine them in turn.

3.1 Honest constructions. One is to construct the required data so that they satisfy the axioms. This is the only mathematically precise method at present. It should be mentioned that even in this case we do not usually understand the full set of submanifold operators.

Topological theories: Many topological QFTs have been constructed in this manner. 2d topological QFTs are famously equivalent to Frobenius algebras. 3d Chern-Simons theories for a compact group G can be rigorously constructed using the Turaev-Viro and Reshetikhin-Turaev constructions. There are 4d topological QFTs as constructed by Crane and Yetter, etc.

Conformal theories in two dimensions: In two dimensions, vertex operator algebras capture the local properties of the holomorphic side of a conformal field theory, and there are many mathematically rigorous discussions on them. Their behaviors on higher-genus Riemann surfaces are governed by their conformal blocks, which have been studied for many rational conformal field theories and also for some irrational conformal field theories. A full-fledged conformal field theory is obtained by consistently gluing the conformal blocks on the holomorphic side and on the anti-holomorphic side. This aspects have also been discussed rigorously by Schweigert, Fuchs, Runkel and their collaborators, see e.g. Schweigert, Fuchs, and Runkel [2006].

Invertible field theories: Invertible field theories are invertible objects in $\mathcal{Q}^d_{\mathcal{S}}$. From the mathematical point of view, these are the first objects one has to study in order to understand $\mathcal{Q}^d_{\mathcal{S}}$, but they got the attention of many physicists relatively recently, only in the last 10 years. Physical studies are led by condensed-matter theorists e.g. Kitaev, Wen and collaborators Kitaev [n.d.] and Chen, Gu, Liu, and Wen [2013]. A mathematical exposition for the relativistic case can be found in e.g. Freed and Hopkins [2016].

It is now known that the group of the isomorphism classes of invertible field theories in $\mathcal{Q}^d_{\mathcal{S}}(G)$, when G is a finite group, is given by $E^d_{\mathcal{S}}(BG)$, where $E^*_{\mathcal{S}}$ is a generalized cohomology theory and BG is the classifying group of G. Slightly more generally, one can consider QFTs defined on d-dimensional manifolds M with $\mathcal{S}$ structure together with a map $f : M \to X$ to a space X up to homotopy. When $\mathcal{S}$ is the smooth structure, the objects in $\mathcal{Q}^d_{\mathcal{S}}[X]$ is known as homotopical sigma models and have been studied by mathematicians, see e.g. Turaev [2010]. For any structure $\mathcal{S}$, they should form a category $\mathcal{Q}^d_{\mathcal{S}}[X]$, and $\mathcal{Q}^d_{\mathcal{S}}(G)$ for a finite group G is an example where $X = BG$. The group of the isomorphism classes of invertible field theories in $\mathcal{Q}^d_{\mathcal{S}}[X]$ should then be given by $E^d_{\mathcal{S}}(X)$.

Free theories in any dimensions: In any spacetime dimensions, for the structure $\mathcal{S}$ being the Riemannian structure with or without spin structure, the free field theories can be constructed rigorously. First is the free scalar field theories. This is a functor B_d from the category of G-vector spaces to the category $\mathcal{Q}^d(G)$ of d-dimensional QFT with Riemannian structure with G symmetry. Pick a G-vector space V. To describe $B_d(V)$, we take a d-dimensional Riemannian manifold M and a G-bundle P with connection A. We then construct the associated vector bundle $V \times_G P$, whose covariant derivative we denote by D_A. Then we have a Laplacian $\triangle_A$ constructed from D_A. Finally, the partition function $Z_{B_d(V)}(M, A)$ is defined in terms of the eigenvalues of $\triangle_A$, and the n-point functions are defined in terms of the Green function of $\triangle_A$.

Second is the free fermion theories. This is a functor F_d again from the category of G-vector spaces to the category $\tilde{\mathcal{Q}}^d(G)$ of d-dimensional possibly-anomalous QFT with Riemannian structure, spin structure and G symmetry. The partition function and the n-point functions are defined in a similar manner as above, but by tensoring with the spinor bundle of M and using the Dirac operator instead of the Laplacian. This is in general an anomalous field theory with an associated anomaly $\mathcal{A}(F_d(V)) \in \mathcal{Q}^{d+1}(G)$, whose partition function is given by the eta invariant, see Dai and Freed [1994].

3.2 Using path integrals. Another is to use the descriptions using the path integral. This might have been the most common method among physicists until recently. The rough idea goes as follows.

Traditional descriptions: To construct a d-dimensional QFT, we first pick a set of fields. As an example, we first choose the spacetime dimension d, a compact Lie group G and its representation R. We consider a G-bundle P with connection A on M, and a section φ of the vector bundle $P \times_G R$ on M. We then pick a polynomial L out of these field variables and their derivatives. As an example let us take

$$L[A,\varphi] = \frac{1}{g^2}|F|^2 + \frac{1}{2}|D_A\varphi|^2 + \mathsf{V}(\varphi) \tag{3-1}$$

where F is the curvature of the G-connection A, D_A is the covariant derivative with respect to A, and V is a G-invariant polynomial on R, usually called the potential of the system.

Then we try to specify the QFT using the field variables and L by means of the path integral. In this example, we try to specify a QFT $Q(G,R,\mathsf{V}) \in \mathbb{Q}^d$ by defining the partition function as

$$Z_{Q(G,R,\mathsf{V})}(M) := \text{“}\int_{\mathfrak{M}_{M,G,R}} e^{-\int_M \star L(A,\varphi)} d\mu\text{”} \tag{3-2}$$

where $\mathfrak{M}_{M,G,R}$ is the moduli space of G-bundles P with connections A together with the section φ, $d\mu$ is an appropriate measure on it, and $\star$ is the Hodge star on M.

Making this construction mathematically precise is an extremely difficult problem, and forms the subject of the constructive quantum field theory. Despite these problems, physicists have used this ill-defined construction to uncover many properties of QFTs. Also, physicists have put the path integral on supercomputers by discretizing the spacetime and approximating the integral by a sum, which has reproduced many experimental results to reasonable accuracy.

In our language: In the language of this note, the theory $Q(G,R,\mathsf{V})$ above is described as follows: we first consider a d-dimensional free scalar theory $B_d(R)$, which we then gauge to construct $Q_0 := B_d(R)/\!\!\!/G$. V is then an element of the space of operators of $B_d(R)/\!\!\!/G$, and can be used to deform the theory from Q_0 to $Q_0(\mathsf{V})$. The result of the deformation is $Q(G,R,\mathsf{V}) := (B_d(R)/\!\!\!/G)(\mathsf{V})$. The Standard Model of Particle Physics is also an example of this construction, obtained by gauging $\mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$ of a certain $B_4(R) \times F_4(R')$ and then by deforming it. Theoretical and experimental high-energy physicists have spent an enormous amount of efforts to pin down what is the representations R and R', and also what is the precise deformation which describes our real world. Note that this includes the brain which is reading this sentence right now.

Traditionally, most physicists only considered QFTs of the form

$$((B_d(R) \times F_d(R')) \not/ G)(\mathsf{V}) \tag{3-3}$$

for some G vector spaces R, R' and an element $\mathsf{V} \in (\mathcal{V}^0_{(B_d(R)\times F_d(R'))\not/ G})^{\mathrm{inv}}$. As such, the aim of the constructive QFT was to make this QFT construction rigorous.

In a more modern point of view, however, not all of QFTs have this form. Still, the gauging operation obtaining $Q\not/G$ from Q, or the deformation operation obtaining $Q(\mathsf{V})$ from Q, should make sense for Q, G and V satisfying appropriate conditions. Therefore, the aim of the constructive QFT should be extended to include these more generalized constructions.

3.3 Using String/M theory. The final method is to construct them using string theory or M-theory. String theory and M-theory are examples of quantum gravity theories, and fall outside of the categories of QFTs discussed in this note.

Quantum gravity theories: A quantum gravity theory is, in an extremely rough sense, a QFT where we are supposed to perform the path integral over the space of the metric, not just over the space of the connections and the sections of the associated vector bundles.

Making sense of the preceding sentence is even more difficult than making sense of the path integral of a gauge theory as above. The latter is difficult but physicists believe that it should be possible to carry it out for a large number of choices of d, G, R, V. The former is so difficult that physicists only know a finite number of sensible examples. There are a few in 10 dimensions, called string theories, and a unique one in 11 dimensions, called the M theory. They are not obtained by performing the path integral over the space of the metric. Rather, they are found accidentally. They are also all intimately related to each other.

In this sense, from mathematicians' point of view, they are even more ill-defined than QFTs. Still, simply assuming their mere existence is extremely powerful, since various QFTs can be realized and studied using string/M theory. The status might be compared with that of Weil cohomology theories and Grothendieck motives when they were first proposed: the assumption of their mere existence of these concepts allows one to explain and give a unified viewpoint on many diverse phenomena.

6d $\mathcal{N}{=}(2,0)$ theories: An important class of QFTs constructed from string theory is the 6d $\mathcal{N}{=}(2,0)$ theories. We start from a 10-dimensional string theory called the type

IIB string theory, which roughly speaking assigns the partition function $Z_{\mathrm{IIB}}(M)$ to a ten-dimensional manifold M. Now, pick a finite subgroup Γ_G of SU(2) of type $G = A_n$, D_n or $E_{6,7,8}$. We define a 6d QFT S_G as $S_G := \mathrm{IIB}\langle \mathbb{C}^2/\Gamma_G \rangle$, i.e. we define its partition function for a 6-dimensional manifold M by

$$Z_{S_G}(M) = Z_{\mathrm{IIB}}(M \times \mathbb{C}^2/\Gamma_G). \tag{3-4}$$

They are examples of 6d $\mathcal{N}=(2,0)$ superconformal meta QFTs. There are no known descriptions of these theories via path integrals. There is also a free 6d $\mathcal{N}=(2,0)$ theory, which can be considered as $S_{\mathrm{U}(1)}$. It is strongly believed that S_G for $G = \mathrm{U}(1)$, A_n, D_n or $E_{6,7,8}$ generate all 6d $\mathcal{N}=(2,0)$ theories.

4 Four-dimensional $\mathcal{N}=2$ supersymmetric theories

Now we would like to discuss the case of 4d $\mathcal{N}=2$ supersymmetric and superconformal theories in more detail. We denote the categories simply by $\mathbb{Q}$ and $\mathbb{Q}_c$. The latter is a subcategory of the former.

4.1 Basic properties. We first recall the overall structure in this particular case.

- Given a compact Lie group G over $\mathbb{C}$, there is a category $\mathbb{Q}(G)$ of 4d $\mathcal{N}=2$ theories with G symmetry.
- Given a homomorphism $\varphi : H \to G$, there is a functor $\varphi^* : \mathbb{Q}(G) \to \mathbb{Q}(H)$, satisfying expected properties.
- There is a canonical object triv $\in \mathbb{Q}(G)$ for any G, which behaves naturally under the functors given above.
- Given $Q_1 \in \mathbb{Q}(G_1)$ and $Q_2 \in \mathbb{Q}(G_2)$, we have an operation $\times$ such that $Q_1 \times Q_2 \in \mathbb{Q}(G_1 \times G_2)$.
- In particular, using the diagonal embedding $G \subset G \times G$, we see that for $Q_{1,2} \in \mathbb{Q}(G)$ we have $Q_1 \times Q_2 \in \mathbb{Q}(G)$. triv is the unit under this product operation.
- If an object $Q \in \mathbb{Q}(F \times G)$ satisfies certain properties, one can form $Q /\!\!/\!\!/ G \in \mathbb{Q}(F)$. It is known that $Q /\!\!/\!\!/ G$ is a family of $\mathcal{N}=2$ theories parameterized by a neighborhood of the origin of $(\mathbb{C}^\times)^n$, where n equals the number of simple factors of the Lie algebra $\mathfrak{g}$ of G.

Here we introduced an operation $Q /\!\!/\!\!/ G$ distinct from the operation Q / G: When $Q /\!\!/\!\!/ G$ can be formed, one can definitely also form Q / G but it is only guaranteed to be in the category of Riemannian QFTs, but not necessarily in the category of $\mathcal{N}$=2 supersymmetric theories. Rather, one needs to define

$$Q /\!\!/\!\!/ G := ([Q \times B_d(\mathfrak{g}_\mathbb{C}) \times F_d(\mathfrak{g}_\mathbb{C} \oplus \mathfrak{g}_\mathbb{C})] / G)(\mathsf{V}) \tag{4-1}$$

where V is a specially chosen deformation, to guarantee that $Q /\!\!/\!\!/ G$ is also $\mathcal{N}$=2 supersymmetric.

This is analogous to the following situation in geometry: one can consider categories $\mathfrak{X}(G)$ of Riemannian manifolds with G action. Then for an $X \in \mathfrak{X}(F \times G)$, one can often form $X/G \in \mathfrak{X}(F)$. Now, consider subcategories $\mathcal{HK}(G) \subset \mathfrak{X}(G)$ formed by hyperkähler manifolds with hyperkähler G action. For a $Y \in \mathcal{HK}(F \times G)$, we can definitely construct Y/G but this is only guaranteed to be $\in \mathfrak{X}(F)$. To get an object in $\mathcal{HK}(G)$, one needs to perform the hyperkähler quotient construction: $Y /\!\!/\!\!/ G \in \mathcal{HK}(F)$. There is a deeper relationship between $\mathcal{Q}(G)$ and $\mathcal{HK}(G)$ which will be discussed below.

Before getting there, we introduce the simplest kinds of objects in $\mathcal{Q}(G)$. Given a quaternionic vector space V with hyperkähler G action, we define a theory of free hypermultiplets based on V by the formula:

$$\mathrm{Hyp}(V) = B_d(V) \times F_d(V). \tag{4-2}$$

They are often just called hypers, and are known to be in the subcategory $\mathcal{Q}_c(G)$ of superconformal theories.

4.2 Higgs branch functor and the slicing. So far in this note we only talked about how to construct objects in the category of quantum field theories. Since this is an ill-defined category, it is of little use to serious mathematicians. There are also functors from these still-ill-defined categories to the well-defined categories. The Higgs branch functor $\mathfrak{M}_{\mathrm{Higgs}} : \mathcal{Q}(G) \to \mathcal{HK}(G)$ is one such example. We also introduce an associated concept which we call 'slicing'.

We only describe the Higgs branch functor at the level of objects, and we will not be able to discuss how morphisms are mapped to morphisms. This is due to our lack of understanding of the morphisms of $\mathcal{Q}(G)$ in the first place. The same comment applies to two other functors introduced in Section 4.5. We will see that even with this rudimentary understanding, we arrive at nontrivial statements.

The Higgs branch functor: This associates to a 4d $\mathcal{N}$=2 supersymmetric theory $Q \in \mathcal{Q}(G)$ a hyperkähaler manifold $\mathfrak{M}_{\mathrm{Higgs}}(Q) \in \mathcal{HK}(G)$, with the following basic properties:

- $\mathfrak{M}_{\text{Higgs}}(\text{Hyp}(V)) = V$,
- For $Q_1 \in \mathbb{Q}(G_1)$ and $Q_2 \in \mathbb{Q}(G_2)$ we have $\mathfrak{M}_{\text{Higgs}}(Q_1 \times Q_2) = \mathfrak{M}_{\text{Higgs}}(Q_1) \times \mathfrak{M}_{\text{Higgs}}(Q_2)$,
- For $Q \in \mathbb{Q}(F \times G)$, we have $\mathfrak{M}_{\text{Higgs}}(Q /\!\!/\!\!/ G) = \mathfrak{M}_{\text{Higgs}}(Q)/\!\!/\!\!/G$ where on the left hand side we perform the gauging and on the right hand side we perform the hyperkähler quotient.
- When Q is $\mathcal{N}$=2 superconformal, $\mathfrak{M}_{\text{Higgs}}(Q)$ is a hyperkähler cone.
- For a family Q of $\mathcal{N}$=2 superconformal theories, $\mathfrak{M}_{\text{Higgs}}(Q)$ is locally constant.

The slicing: Let us now introduce the concept of the slicing. To do this, we first recall the concept of the Slodowy slice. Consider $\mathfrak{g}_{\mathbb{C}}$ and take a nilpotent element e in it. It is known that we also have elements $h, f \in \mathfrak{g}_{\mathbb{C}}$ so that the triple (e, h, f) defines a homomorphism from $\mathfrak{su}(2)_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$ and then $\mathrm{SU}(2) \to G$. We denote the commutant of this $\mathrm{SU}(2)$ within G by G_e. We define the Slodowy slice S_e at e by the formula

$$S_e := \{e + x \mid [f, x] = 0,\ x \in \mathfrak{g}_{\mathbb{C}}\}. \tag{4-3}$$

The Slodowy slice S_e has a natural action by G_e. In the various constructions below, the results only depend on the conjugacy class of the nilpotent element e. Therefore there are essentially finite possibilities of e for a given $\mathfrak{g}$, labeled by the nilpotent orbits.

Now, given a hyperkähler space $X \in \mathcal{HK}(G)$, we can define a new hyperkähler space $X \wr e \in \mathcal{HK}(G_e)$, which is given as a complex manifold by the expression

$$X \wr e := {\mu_{\mathbb{C}}}^{-1}(S_e) \tag{4-4}$$

where $\mu_{\mathbb{C}} : X \to \mathfrak{g}_{\mathbb{C}}$ is the complex part of the moment map of the G action. It is known that we can give a hyperkähler structure. For $e = 0$, we simply have $X \wr e = X$. The notation $\wr e$ is also introduced for the sake of this exposition. It is simply chosen to vaguely suggest the letter ‘s’.

Now, for any $Q \in \mathbb{Q}(G)$, there is a QFT procedure we call the slicing of Q by e. In the physics literature it is often called the nilpotent Higgsing or the partial closure of the puncture. This results in a theory we denote by $Q \wr e \in \mathbb{Q}(G_e)$. This affects the Higgs branch in the expected way:

$$\mathfrak{M}_{\text{Higgs}}(Q \wr e) = \mathfrak{M}_{\text{Higgs}}(Q) \wr e. \tag{4-5}$$

4.3 Examples. Let us now discuss examples of 4d $\mathcal{N}$=2 theories.

4d $\mathfrak{N}$=2 gauge theories: We already introduced the hypermultiplet $\mathrm{Hyp}(V) \in \mathfrak{Q}(G \times F)$ for a quaternionic vector space V with hyperkähler action of $G \times F$. Then we can form a family of 4d $\mathfrak{N}$=2 theories

$$\mathrm{Hyp}(V)/\!\!/\!\!/G \in \mathfrak{Q}(F) \tag{4-6}$$

if the condition is right. These are called 4d $\mathfrak{N}$=2 gauge theories, and have been intensively studied by physicists. We easily see that

$$\mathfrak{M}_{\mathrm{Higgs}}(\mathrm{Hyp}(V)/\!\!/\!\!/G) = V/\!\!/\!\!/G. \tag{4-7}$$

Minahan-Nemeschansky theories: These theories are $\mathfrak{N}$=2 superconformal theories with $E_{6,7,8}$ symmetries specified by a positive integer n:

$$\mathrm{MN}(E_i, n) \in \mathfrak{Q}_c(E_i), \qquad i = 6, 7, 8, \tag{4-8}$$

constructed using a variant of the type IIB theory called the F-theory. From this construction it is known that

$$\mathfrak{M}_{\mathrm{Higgs}}(\mathrm{MN}(E_i, n)) = \mathfrak{M}^{\mathrm{inst}}_{E_i,n} \tag{4-9}$$

where the right hand side is the centered framed instanton moduli space of the group E_i on $\mathbb{R}^4$ with instanton number n.

This means that they are not an $\mathfrak{N}$=2 gauge theory of the form $\mathrm{Hyp}(V)/\!\!/\!\!/G$. If so, we would have an equality

$$\mathfrak{M}^{\mathrm{inst}}_{E_i,n} = V/\!\!/\!\!/G \tag{4-10}$$

meaning that there is an ADHM-like description for the instanton moduli spaces of exceptional groups. But this is almost surely impossible, since no such construction is known.

For $n = 1$ these theories were first studied by Minahan and Nemeschansky [1996, 1997] and are notable as one of the earliest examples of theories which are not gauge theories, although they used a different argument against having a gauge theory description. More recently, gauge theory descriptions which only manifest $\mathfrak{N}$=1 supersymmetry have been found Gadde, Razamat, and Willett [2015], but they are not useful at present to study its Higgs branch.

Class S theories: For this construction, we start from a 6d $\mathfrak{N}$=(2, 0) theory S_G, and compactify it on a two-dimensional surface $C_{g,n}$ of genus g with n punctures. It is known that by an appropriate trick the resulting 4d theory is $\mathfrak{N}$=2 supersymmetric and only depends on the complex structure of $C_{g,n}$. We then have a family of 4d $\mathfrak{N}$=2 theory $S_{G,g,n}$

parameterized by $\mathcal{M}_{g,n}$, the moduli space of Riemann surfaces of genus g with n punctures. This is the class S theory, first introduced in Gaiotto [2012] and Gaiotto, Moore, and Neitzke [2009], having the following properties:

- They are $\in \mathbb{Q}(G^n)$, where G is the simply-connected compact Lie group of type $\mathfrak{g}$, so that each factor of G is associated to a puncture of $C_{g,n}$.

- In particular, the family over $\mathcal{M}_{g,n}$ is such that when two points on $C_{g,n}$ are exchanged, two factors of G in $\mathbb{Q}(G^n)$ are exchanged.

- They are $\in \mathbb{Q}_c(G^n)$, i.e. superconformal, when $g = 0$, $n \geq 3$, or $g = 1$, $n \geq 1$, or $g \geq 2$.

- In a neighborhood of the boundary of $\mathcal{M}_{g,n}$ where the genus g surface $C_{g,n}$ degenerates to a connected sum of $C_{g',n'}$ and $C_{g'',n''}$ such that $g' + g'' = g$ and $n' + n'' = n + 2$, we have the identification that

  $$S_{G,g,n} = (S_{G,g',n'} \times S_{G,g',n''}) /\!\!/\!\!/ G. \tag{4-11}$$

  Here, the gauging operation on the right hand side is performed in the following manner. The connected sum is performed at a puncture of $C_{g',n'}$ and another of $C_{g'',n''}$. Accordingly we have a chosen subgroup G for the first puncture and a chosen subgroup G for the second puncture. We then perform the gauging with respect to the diagonal subgroup of these two. The right hand side is a family over $\mathcal{M}_{g',n'} \times \mathcal{M}_{g'',n''} \times U$ where U is a neighborhood of the origin of $\mathbb{C}^\times$, which is identified with the neighborhood of the said boundary of $\mathcal{M}_{g,n}$.

- For any G, we always have the principal embedding $\mathrm{SU}(2) \to G$ and the corresponding nilpotent element e_{prin}. The commutant is trivial, $G_{e_{\text{prin}}} = 1$. Then we have

  $$S_{G,g,n} \wr e_{\text{prin}} = S_{G,g,n-1}. \tag{4-12}$$

  Namely, by slicing a puncture of a class S theory by the principal nilpotent element e_{prin}, we can effectively remove the puncture.

From the properties listed above, it can be seen that $S_{G,g,n}$ can be constructed from $S_{G,0,3}$. For this reason a special abbreviation is introduced: $T_G := S_{G,0,3} \in \mathbb{Q}(G^3)$. From the construction, it has a natural self-equivalence permuting three factors of G.

4.4 Known overlaps among the examples. Now let us discuss some properties of the theories T_G and $S_{G,g,n}$. First, it is known that $T_{\mathrm{SU}(2)} = \mathrm{Hyp}(V \otimes_{\mathbb{C}} V \otimes_{\mathbb{C}} V)$, where V is the defining representation of $\mathrm{SU}(2)$.

Second, for other $G \neq \mathrm{SU}(2)$, no gauge theory description is known. Still, we can slice it at three nilpotent elements $e_{1,2,3} \in \mathfrak{g}_{\mathbb{C}}$ and consider the theory $T_G \wr e_1, e_2, e_3$. For a suitable choice of $e_{1,2,3}$, they are known to be equivalent to $\mathrm{Hyp}(V)$ for a suitable V. Here we only discuss one example.

Take $G = \mathfrak{su}(N)$. A nilpotent element in $\mathfrak{g}_{\mathbb{C}}$ can be conveniently described by a partition $[n_i]$ of N. We take $e = [N-1, 1]$, for which $G_e = \mathrm{U}(1)$. Then we have

$$T_{\mathrm{SU}(N)} \wr e = \mathrm{Hyp}(V \otimes \bar{W} \otimes X \oplus \bar{W} \otimes V \otimes \bar{X}) \tag{4-13}$$

where $V \simeq W \simeq \mathbb{C}^N$ have actions of $\mathrm{SU}(V)$ and $\mathrm{SU}(W)$ associated to the first and the second punctures, and X is the standard one-dimensional representation of $\mathrm{U}(1) = G_e$ associated to the third puncture sliced by e.

We also know the following equivalences:

$$\mathrm{MN}(E_6, N) = T_{\mathrm{SU}(3N)} \wr [N^3], [N^3], [N^3], \tag{4-14}$$

$$\mathrm{MN}(E_7, N) = T_{\mathrm{SU}(4N)} \wr [2N, 2N], [N^4], [N^4], \tag{4-15}$$

$$\mathrm{MN}(E_8, N) = T_{\mathrm{SU}(6N)} \wr [3N, 3N], [2N, 2N, 2N], [N^6]. \tag{4-16}$$

Note that by construction, on the right hand side are objects in $\mathfrak{Q}(\mathrm{SU}(3)\times\mathrm{SU}(3)\times\mathrm{SU}(3))$, $\mathfrak{Q}(\mathrm{SU}(2)\times\mathrm{SU}(4)\times\mathrm{SU}(4))$, and $\mathfrak{Q}(\mathrm{SU}(2)\times\mathrm{SU}(3)\times\mathrm{SU}(6))$, while on the left hand side are objects in $\mathfrak{Q}(E_6)$, $\mathfrak{Q}(E_7)$, $\mathfrak{Q}(E_8)$. To write an equality, we use the homomorphism $\mathrm{SU}(3)^3 \to (\mathrm{SU}(3)^3/\mathbb{Z}_3) \subset E_6$, etc.

4.5 Two other functors. In this section we discuss two more functors from $\mathfrak{Q}_c(G)$. One is the superconformal index functor $Z^{\mathrm{SCI}}_{p,q,t}$ and another is the vertex operator algebra functor $\mathbb{VOA}$. Applied to a family of objects in $\mathfrak{Q}_c(G)$, both give a locally constant result.

Superconformal index: The superconformal index functor $Z^{\mathrm{SCI}}_{p,q,t}$ is a functor which assigns to $\mathfrak{Q}(G)$ a virtual representation of $G \times (\mathbb{C}^\times)^3$. We describe it using $\mathbb{C}[[p,q,t]]$-valued Weyl-invariant functions on the maximal torus $T^r \subset G$, where $(p,q,t) \in (\mathbb{C}^\times)^3$ and we also use variables $z = (z_1, \ldots, z_r) \in T^r \subset G$. We use standard abbreviations $z^w = \prod_i {z_i}^{w_i}$ for a weight $w = (w_1, \ldots, w_r)$ of G.

This functor was introduced in Gadde, Rastelli, Razamat, and Yan [2013]. The essential idea was to identify a differential $d : \mathcal{V}_Q \to \mathcal{V}_Q$ which satisfies $d^2 = 0$. Then $Z^{\mathrm{SCI}}_{p,q,t}(Q)$ is the cohomology $H(\mathcal{V}_Q, d)$. The elliptic Gamma function $\Gamma_{p,q}(x)$ defined as follows will play an important role for this functor:

$$\Gamma_{p,q}(x) = \prod_{m,n\geq 0} \frac{1-x^{-1}p^{m+1}q^{n+1}}{1-xp^mq^n}. \tag{4-17}$$

The basic properties of $Z^{\mathrm{SCI}}_{p,q,t}$ are the following. First, the superconformal index of a free hypermultiplet is given by

$$Z^{\mathrm{SCI}}_{p,q,t}(\mathrm{Hyp}(V)) = \prod_{w:\text{weights of } V} \Gamma_{p,q}(t^{1/2}z^w). \tag{4-18}$$

Second, $Z^{\mathrm{SCI}}_{p,q,t}(Q_1 \times Q_2) = Z^{\mathrm{SCI}}_{p,q,t}(Q_1)Z^{\mathrm{SCI}}_{p,q,t}(Q_2)$. Third, for $Q \in \mathbb{Q}_c(F \times G)$ we have

$$Z^{\mathrm{SCI}}_{p,q,t}(Q /\!\!/\!\!/ G) = \Big(\frac{1}{\Gamma_{p,q}(t)\Gamma'_{p,q}(1)}\Big)^r \frac{1}{|W_G|}\int_{T^r} Z^{\mathrm{SCI}}_{p,q,t}(Q)\times \prod_{\alpha:\text{roots of } G} \frac{1}{\Gamma_{p,q}(z^\alpha)\Gamma_{p,q}(tz^\alpha)} \prod_{i=1}^r \frac{dz_i}{2\pi\sqrt{-1}z_i} \tag{4-19}$$

where $z \in T^r \subset G$ and $|W_G|$ is the order of the Weyl group. The measure appearing in (4-19) is an elliptic generalization of the Macdonald inner product and reduces to the standard Macdonald measure when $p = 0$ up to a trivial rescaling.

The slicing affects the superconformal index in the following manner. Given a nilpotent element $e \in \mathfrak{g}_{\mathbb{C}}$, recall that there are $f, h \in \mathfrak{g}_{\mathbb{C}}$ such that they determine a homomorphism $\mathrm{SU}(2) \to G$, and G_e is the commutant of the image in G. We then decompose $\mathfrak{g}_{\mathbb{C}}$ as

$$\mathfrak{g}_{\mathbb{C}} = \bigoplus_d V_d \otimes R_d \tag{4-20}$$

where V_d is the irreducible representation of dimension d of $\mathrm{SU}(2)$ and R_d is a representation of G_e. We then define

$$K_e(z) = \prod_d \prod_{w:\text{weights of } R_d} \Gamma_{p,q}(t^{(d+1)/2}z^w) \tag{4-21}$$

for z taking values in the maximal torus of G_e. Then we have:

$$Z^{\mathrm{SCI}}_{p,q,t}(Q \wr e)(z) = K_e(z)\left[K_0(x)^{-1}Z^{\mathrm{SCI}}_{p,q,t}(Q)(x)\right]_{x\to zt^{h/2}}. \tag{4-22}$$

Vertex operator algebra: Let us denote the category of vertex operator algebras with a homomorphism from an affine algebra $\hat{\mathfrak{g}}$ by $\mathcal{V}(G)$. The functor $\mathbb{VOA}$ associates to an object $Q \in \mathbb{Q}(G)$ a vertex operator algebra $\mathbb{VOA}(Q) \in \mathcal{V}(G)$. This functor was introduced in Beem, Lemos, Liendo, Peelaers, Rastelli, and van Rees [2015]. The essence was to show that one can locate a nice subspace of $\mathcal{V}_Q$ such that the OPE algebra structure of $\mathcal{V}_Q$ induces the structure of a vertex operator algebra on it.

Here we mostly use physicists’ notation for the vertex operator algebras. For a general introduction to vertex operator algebras, see e.g. Frenkel and Ben-Zvi [2004] and Arakawa [2018]. In the following, a VOA always stands for vertex operator super-algebra $\mathbb{V} = \oplus_n \mathbb{V}_n$, $\mathbb{V}_n = \mathbb{V}_{n,+} \oplus \mathbb{V}_{n,-}$. The part $\mathbb{V}_\pm = \oplus_n \mathbb{V}_{n,\pm}$ are called bosonic and fermionic, respectively, and the subscript n is the eigenvalue of L_0.

The basic features of this functor is as follows. First, for $Q = \mathrm{Hyp}(V) \in \mathbb{Q}_c(G)$, the corresponding $\mathbb{VOA}(Q)$ is the symplectic boson VOA $\mathbb{SB}(V)$ defined in the following manner: $\mathbb{SB}(V)$ is generated by $\mathbb{SB}(V)_{1/2,+} \simeq V$, with the operator product expansion given by

$$v(z)w(0) \simeq \frac{\langle v, w\rangle}{z} \tag{4-23}$$

for $v, w \in \mathbb{SB}(V)_{1/2,+} \simeq V$, where $\langle \cdot, \cdot \rangle$ is the symplectic pairing of the quaternionic vector space V.

Second, we have $\mathbb{VOA}(Q_1 \times Q_2) = \mathbb{VOA}(Q_1) \otimes \mathbb{VOA}(Q_2)$. Third, to describe $\mathbb{VOA}(Q /\!\!/\!\!/ G)$, we define the quotient operation in the category of vertex operator algebras. We start from an object $\mathbb{V} \in \mathcal{V}(G)$. We introduce a ghost VOA $\mathbb{BC}(G)$, generated by fermionic fields b^A in $\mathbb{BC}(G)_{1,-}$ and c_A in $\mathbb{BC}(G)_{0,-}$ for $A = 1, \ldots, \dim \mathfrak{g}$ with the OPE

$$b^A(z)c_B(w) \sim \frac{\delta^A_B}{z-w}. \tag{4-24}$$

This has a subalgebra $\hat{\mathfrak{g}}_{+2h^\vee(\mathfrak{g})}$. Denote by $J^A_{\mathbb{V}}$ and J^A_{ghost} the affine $\mathfrak{g}$ currents of $\mathbb{V}$ and $\mathbb{BC}(G)$ respectively. We define

$$j_{\mathrm{BRST}}(z) = \sum_A (c_A J^A_{\mathbb{V}}(z) + \frac{1}{2} c_A J^A_{\mathrm{ghost}}(z)). \tag{4-25}$$

Then $d = j_{\mathrm{BRST},0}$ satisfies $d^2 = 0$ if the level of $\hat{\mathfrak{g}}$ in $\mathbb{V}$ is $-2h^\vee(\mathfrak{g})$. We then take the subspace

$$\mathbb{W} \subset \mathbb{V} \otimes \mathbb{BC}(G) \tag{4-26}$$

defined by

$$\mathbb{W} = \bigcap_A \mathrm{Ker}\, b^A_0 \tag{4-27}$$

where $J^A_{\text{total}} = J^A_{\mathbb{V}} + J^A_{\text{ghost}}$. We can check that the differential d acts within $\mathbb{W}$, and finally we define

$$\mathbb{V}/G := H(\mathbb{W}, d). \tag{4-28}$$

We then have the following statement: for $Q \in \mathfrak{Q}_c(G)$, assume that $Q /\!\!/\!\!/ G \in \mathfrak{Q}_c(G)$. This implies the level of $\hat{\mathfrak{g}}$ in $\mathbb{VOA}(Q)$ is $-2h^\vee(\mathfrak{g})$, and we have $\mathbb{VOA}(Q /\!\!/\!\!/ G) = \mathbb{VOA}(Q)/G$.

Third, the slicing by a nilpotent element e of an object $V \in \mathfrak{V}(\mathfrak{g})$ is defined by the quantum Drinfeld-Sokolov reduction: $V \wr e := \mathbb{DS}(V, e)$. Then we have $\mathbb{VOA}(Q \wr e) = \mathbb{VOA}(Q) \wr e$.

4.6 Relation among the functors.

$\mathbb{VOA}$ to $\mathfrak{M}_{\text{Higgs}}$: For a vertex operator algebra $\mathbb{V} \in \mathfrak{V}(\mathfrak{g})$, one can construct the associated variety avar $\mathbb{V}$, which is a holomorphic symplectic variety with G action Arakawa [2012, 2018]. This is obtained by Spec of Zhu's C_2 algebra of the vertex algebra $\mathbb{V}$. It is believed that $\mathfrak{M}_{\text{Higgs}}(Q) = \text{avar}\,\mathbb{VOA}(Q)$ in general Beem and Rastelli [2017].

$\mathbb{VOA}$ to $Z^{\text{SCI}}_{p=0,q=t}$: For a vertex operator algebra $\mathbb{V} \in \mathfrak{V}(\mathfrak{g})$, we can define its character ch $\mathbb{V}$ as a $\mathbb{C}[[q]]$-valued function on G by the following formula:

$$G \ni z \mapsto \text{ch}\,\mathbb{V}(z) = \sum_n q^n(\text{tr}_{\mathbb{V}_{n,+}}\, z - \text{tr}_{\mathbb{V}_{n,-}}\, z). \tag{4-29}$$

It is known in general that

$$\text{ch}\,\mathbb{VOA}(Q) = Z^{\text{SCI}}_{p=0,q=t}(Q). \tag{4-30}$$

$\mathfrak{M}_{\text{Higgs}}$ to $Z^{\text{SCI}}_{p=q=0,t=\tau^2}$: For a hyperkähler cone $X \in \mathfrak{HK}_c(G)$, we can define its character ch X as a $\mathbb{C}[[\tau]]$-valued function on G in the following manner. We decompose the function ring $\mathbb{C}[X]$ into the graded pieces $\mathbb{C}[X]_n$, where the symplectic form on X is normalized to have the grade $+2$. We then define

$$G \ni z \mapsto \text{ch}\,X(z) = \sum_n \tau^n\, \text{tr}_{\mathbb{C}[X]_n}\, z. \tag{4-31}$$

For many examples including $Q = S_{G,g=0,n} \wr e_1, \dots, e_n$, it is known that

$$\text{ch}\,\mathfrak{M}_{\text{Higgs}}(Q) = Z^{\text{SCI}}_{p=q=0,t=\tau^2}(Q). \tag{4-32}$$

Summary of the functors: We summarize below the relationships of the functors discussed so far:

$$\begin{tikzcd}
C_{g,n} \arrow[dr] & \mathfrak{Q} & & Z^{\mathrm{SCI}}_{p,q,t}(Q) \arrow[r] & Z^{\mathrm{SCI}}_{p=0,q,t}(Q) \arrow[d] \arrow[dd, bend left] \\
 & S_{G,g,n} \quad \mathrm{Hyp}(V) & Q \arrow[ur] \arrow[r] \arrow[dr] & \mathbb{VOA}(Q) \arrow[r, "\mathrm{ch}"] \arrow[d, "\mathrm{avar}"] & Z^{\mathrm{SCI}}_{p=0,q=t}(Q) \\
V \arrow[ur] & & & \mathfrak{M}_{\mathrm{Higgs}}(Q) \arrow[r, dotted, "\mathrm{ch}"] & Z^{\mathrm{SCI}}_{p=q=0,t=\tau^2}(Q)
\end{tikzcd}$$

4.7 Consequences. Let us see a few consequences of the whole setup.

Class S theories of type $\mathrm{SU}(2)$ **:** We already noted that

$$S_{\mathrm{SU}(2),0,4} = (T_{\mathrm{SU}(2)} \times T_{\mathrm{SU}(2)}) /\!\!/\!\!/ \mathrm{SU}(2) \tag{4-33}$$

and also

$$T_{\mathrm{SU}(2)} = \mathrm{Hyp}(V \otimes_{\mathbb{C}} V \otimes_{\mathbb{C}} V) \tag{4-34}$$

where V is the defining 2-dimensional representation of $\mathrm{SU}(2)$.

By applying the functor $\mathfrak{M}_{\mathrm{Higgs}}$, we have

$$\mathfrak{M}_{\mathrm{Higgs}}(S_{\mathrm{SU}(2),0,4}) = [V_a \otimes_{\mathbb{C}} V_b \otimes_{\mathbb{C}} V_x \oplus V_y \otimes_{\mathbb{C}} V_c \otimes_{\mathbb{C}} V_d] /\!\!/\!\!/ \mathrm{SU}(2). \tag{4-35}$$

Here, we put subscripts to various copies of V to distinguish them, and $\mathrm{SU}(2)$ used in the quotient is the diagonal subgroup of $\mathrm{SU}(V_x)$ and $\mathrm{SU}(V_y)$. The right hand side clearly has an action of $\prod_{i=a,b,c,d} \mathrm{SU}(V_i)$. But from the left hand side, we see that there should also be an action S_4 permuting $\mathrm{SU}(V_{a,b,c,d})$, which is not obvious from the right hand side. The right hand side, when written as

$$V \otimes_{\mathbb{R}} \mathbb{R}^8 /\!\!/\!\!/ \mathrm{SU}(V), \tag{4-36}$$

is the ADHM construction of the minimal nilpotent orbit of

$$\mathrm{SO}(8) \supset [\prod_{i=a,b,c,d} \mathrm{SU}(V_i)]/\mathbb{Z}_2$$

and the S_4 permutations of V_i’s are given by elements of $\mathrm{Aut}(\mathrm{SO}(8))$.

By applying the functor $Z^{\mathrm{SCI}}_{p,q,t}$, we have the equality

(4-37)
$$Z^{\mathrm{SCI}}(S_{\mathrm{SU}(2),0,4})(a,b,c,d) = \frac{1}{\Gamma_{p,q}(t)\Gamma'_{p,q}(1)}\frac{1}{2}\oint \frac{dz}{2\pi\sqrt{-1}z}\prod_{\pm}\frac{1}{\Gamma_{p,q}(z^{\pm 2})\Gamma_{p,q}(tz^{\pm 2})} \times \prod_{\pm\pm\pm}\Gamma_{p,q}(t^{1/2}a^{\pm}b^{\pm}z^{\pm})\prod_{\pm\pm\pm}\Gamma_{p,q}(t^{1/2}c^{\pm}d^{\pm}z^{\pm})$$

where a, b, c, d are now thought of as $\in \mathrm{U}(1) \subset \mathrm{SU}(2)$. The left hand side should be symmetric under an arbitrary permutation of a, b, c, d. This symmetry is however nontrivial on the right hand side. This was pointed out from this physical argument in Gadde, Pomoni, Rastelli, and Razamat [2010], and completely independently studied and proved in a mathematical work van de Bult [2011].

By applying the functor $\mathbb{VOA}$, we have the equality

$$\mathbb{VOA}(S_{\mathrm{SU}(2),0,4}) = \mathbb{SB}[V_a \otimes_{\mathbb{C}} V_b \otimes_{\mathbb{C}} V_x \oplus V_y \otimes_{\mathbb{C}} V_c \otimes_{\mathbb{C}} V_d]/\mathrm{SU}(2). \tag{4-38}$$

There is a simple physics argument that the left hand side is just $\hat{\mathfrak{so}}(8)_{-2}$. Then the equality above is a new free-field construction of this particular vertex operator algebra based on the affine Lie algebra, which remains to be proven.

Class S theories of general type: In general, we have the relation

$$S_{G,0,4} = (T_G \times T_G)/\!/\!/G, \tag{4-39}$$

where $T_G = S_{G,0,3}$ as we defined above. The left hand side is symmetric under G^4 together with S_4 permuting four factors of G. The right hand side is symmetric under G^4, but only a subgroup $(S_2 \times S_2) \rtimes S_2 \subset S_4$ permuting four factors of G is manifest.

By applying the functors $\mathfrak{M}_{\mathrm{Higgs}}$ or $\mathbb{VOA}$, we are led to the following conjectures. Let $X_G = \mathfrak{M}_{\mathrm{Higgs}}(T_G)$, a hyperkähler cone with G^3 action, and $\mathbb{V}_G = \mathbb{VOA}(T_G)$, a vertex operator algebra with a $\hat{\mathfrak{g}}^3$ subalgebra. They satisfy the following:

- $(X_G \times X_G)/\!/\!/G$ is a hyperkähler cone with G^4 action together with S_4 permuting four factors of G.
- $(\mathbb{V}_G \otimes \mathbb{V}_G)/G$ is a vertex operator algebra with $\hat{\mathfrak{g}}^{\oplus 4}$ subalgebra with S_4 permuting four factors of $\hat{\mathfrak{g}}$.
- $X_G = \operatorname{avar} \mathbb{V}_G$.

In addition, there is a way to determine $Z^{\mathrm{SCI}}_{p=0,q,t}(T_G)$ explicitly using the theory of Macdonald polynomials. By taking a further limit $q = t$ or $q = 0$, $t = \tau^2$, we see the properties

$$\mathrm{ch}\, X_G(z_1, z_2, z_3) = \sum_\lambda \frac{\prod_{i=1,2,3} K_0(z_i)\underline{H}_\lambda(z_i)}{K_{e_{\mathrm{prin}}}\underline{H}_\lambda(q^\rho)} \tag{4-40}$$

$$\mathrm{ch}\, \mathbb{V}_G(z_1, z_2, z_3) = \sum_\lambda \frac{\prod_{i=1,2,3} K_0(z_i)\chi_\lambda(z_i)}{K_{e_{\mathrm{prin}}}\chi_\lambda(q^\rho)} \tag{4-41}$$

where λ runs over all irreducible representation of G, $\chi_\lambda(z)$ is the character in that representation, and $\underline{H}_\lambda(z) = N_\lambda H_\lambda(z)$ where $H_\lambda(z)$ is the standard Hall-Littlewood polynomial of type G and N_λ is a normalization constant so that $\underline{H}_\lambda$ is orthonormal under the following measure:

$$\delta_{\mu\nu} = \frac{1}{|W_G|}\int_{T^r} \underline{H}_\lambda(z)\underline{H}_\mu(1/z)\frac{1}{(1-\tau^2)^r}\prod_{\alpha:\text{roots of } G}\frac{1-z^\alpha}{1-\tau^2 z^\alpha}\prod_{i=1}^r \frac{dz_i}{2\pi\sqrt{-1}z_i}. \tag{4-42}$$

Finally, we should have

$$X_{\mathrm{SU}(3N)} \wr [N^3], [N^3], [N^3] = \mathfrak{M}^{\mathrm{inst}}_{E_6,N}, \tag{4-43}$$

$$X_{\mathrm{SU}(4N)} \wr [2N, 2N], [N^4], [N^4] = \mathfrak{M}^{\mathrm{inst}}_{E_7,N}, \tag{4-44}$$

$$X_{\mathrm{SU}(6N)} \wr [3N, 3N], [2N, 2N, 2N], [N^6] = \mathfrak{M}^{\mathrm{inst}}_{E_8,N}, \tag{4-45}$$

and

$$\mathbb{V}_{\mathrm{SU}(3)} = (\hat{e_6})_{-6} \tag{4-46}$$

$$\mathbb{V}_{\mathrm{SU}(4)} \wr [2, 2], [1^4], [1^4] = (\hat{e_7})_{-8} \tag{4-47}$$

$$\mathbb{V}_{\mathrm{SU}(6)} \wr [3, 3], [2, 2, 2], [1^6] = (\hat{e_8})_{-12} \tag{4-48}$$

where the last three equations follow from the property of the Minahan-Nemeschansky theory.

The properties satisfied by X_G were already given in Moore and Tachikawa [2011] in a slightly different language, as a 2d topological QFT taking values in the category of holomorphic symplectic varieties. Such X_G has now been constructed in Ginzburg and Kazhdan [n.d.] and Braverman, Finkelberg, and Nakajima [2017] as holomorphic symplectic varieties. The construction of the vertex operator algebras $\mathbb{V}_G$ satisfying these relations is also announced Arakawa [2018].

Instantons and W-algebras: Finally a brief remark is made about the conjecture that the direct sum $\mathcal{H}_G := \oplus_n H^*_G(\mathfrak{M}^{\text{inst}}_{G,n})$ of the equivariant cohomology of the instanton moduli space of simply-laced group G has an action of the W-algebra of the corresponding type, originally made in Alday, Gaiotto, and Tachikawa [2010], and now proved by Schiffmann and Vasserot [2012], Maulik and Okounkov [2012], and Braverman, Finkelberg, and Nakajima [2014], from the point of view of the present note.

The essential point is that there is an another functor Z_{Nek} defined on $\mathcal{Q}_c(G)$ taking values in $\mathcal{H}_G$, such that a family of objects in $\mathcal{Q}_c(G)$ parameterized by $\mathcal{M}$ is sent to a section of an $\mathcal{H}_G$ bundle over $\mathcal{M}$. When $\mathcal{M} = \mathcal{M}_{g,n}$ as in the class S theory, this has a natural relationship with the theory of the 2d conformal blocks. A consideration in this line of thought naturally leads to the conjecture that $\mathcal{H}_G$ has to have the action of the W-algebra of type G. More details can be found in the longer version of this article on the author's webpage.

References

Luis F. Alday, Davide Gaiotto, and Yuji Tachikawa (2010). "Liouville Correlation Functions from Four-Dimensional Gauge Theories". *Lett. Math. Phys.* 91, pp. 167–197. arXiv: 0906.3219 (cit. on p. 2750).

Tomoyuki Arakawa (2012). "A remark on the C_2-cofiniteness condition on vertex algebras". *Math. Z.* 270.1-2, pp. 559–575. arXiv: 1004.1492. MR: 2875849 (cit. on p. 2746).

– (2018). "Representation theory of W-algebras and Higgs branch conjecture". In: *Proc. of the International Congress of Mathematicians, Rio de Janeiro 2018*. Ed. by Boyan Sirakov, Paulo Ney de Souza, and Marcelo Viana. In this volume. World Scientific, pp. 1263–1282 (cit. on pp. 2745, 2746, 2749).

Michael Atiyah (1988). "Topological quantum field theories". *Inst. Hautes Études Sci. Publ. Math.* 68, 175–186 (1989). MR: 1001453 (cit. on p. 2728).

Christopher Beem, Madalena Lemos, Pedro Liendo, Wolfger Peelaers, Leonardo Rastelli, and Balt C. van Rees (2015). "Infinite Chiral Symmetry in Four Dimensions". *Commun. Math. Phys.* 336.3, pp. 1359–1433. arXiv: 1312.5344 (cit. on p. 2745).

Christopher Beem and Leonardo Rastelli (2017). "Vertex Operator Algebras, Higgs Branches, and Modular Differential Equations". arXiv: 1707.07679 (cit. on p. 2746).

Alexander Braverman, Michael Finkelberg, and Hiraku Nakajima (2014). "Instanton Moduli Spaces and $\mathcal{W}$-Algebras". arXiv: 1406.2381 (cit. on p. 2750).

– (2017). "Ring Objects in the Equivariant Derived Satake Category Arising from Coulomb Branches". arXiv: 1706.02112 (cit. on p. 2749).

Fokko J. van de Bult (2011). “An elliptic hypergeometric integral with $W(F_4)$ symmetry”. *Ramanujan J.* 25.1, pp. 1–20. MR: 2787288 (cit. on p. 2748).

Nils Carqueville, Catherine Meusburger, and Gregor Schaumann (2016). “3-Dimensional Defect Tqfts and Their Tricategories”. arXiv: 1603.01171 (cit. on p. 2732).

Xie Chen, Zheng-Cheng Gu, Zheng-Xin Liu, and Xiao-Gang Wen (2013). “Symmetry Protected Topological Orders and the Group Cohomology of Their Symmetry Group”. *Phys. Rev.* B87.15, p. 155114. arXiv: 1106.4772 (cit. on p. 2735).

Xian-zhe Dai and Daniel S. Freed (1994). “Eta Invariants and Determinant Lines”. *J. Math. Phys.* 35. [Erratum: J. Math. Phys.42,2343(2001)], pp. 5155–5194. arXiv: hep-th/9405012 (cit. on p. 2735).

Daniel S. Freed (2013). “The cobordism hypothesis”. *Bull. Amer. Math. Soc. (N.S.)* 50.1, pp. 57–92. arXiv: 1210.5100. MR: 2994995 (cit. on p. 2732).

Daniel S. Freed and Michael J. Hopkins (2016). “Reflection Positivity and Invertible Topological Phases”. arXiv: 1604.06527 (cit. on p. 2735).

Daniel S. Freed and Constantin Teleman (2012). “Relative Quantum Field Theory”. arXiv: 1212.1692 (cit. on p. 2733).

Edward Frenkel and David Ben-Zvi (2004). *Vertex algebras and algebraic curves*. Second. Vol. 88. Mathematical Surveys and Monographs. Providence, RI: American Mathematical Society, pp. xiv+400. MR: 2082709 (cit. on p. 2745).

Abhijit Gadde, Elli Pomoni, Leonardo Rastelli, and Shlomo S. Razamat (2010). “S-Duality and 2D Topological QFT”. *JHEP* 03, p. 032. arXiv: 0910.2225 (cit. on p. 2748).

Abhijit Gadde, Leonardo Rastelli, Shlomo S. Razamat, and Wenbin Yan (2013). “Gauge Theories and Macdonald Polynomials”. *Commun.Math.Phys.* 319, pp. 147–193. arXiv: 1110.3740 (cit. on p. 2743).

Abhijit Gadde, Shlomo S. Razamat, and Brian Willett (2015). “”Lagrangian” for a Non-Lagrangian Field Theory with $\mathfrak{N} = 2$ Supersymmetry”. *Phys. Rev. Lett.* 115.17, p. 171604. arXiv: 1505.05834 (cit. on p. 2741).

Davide Gaiotto (2012). “$\mathfrak{N} = 2$ Dualities”. *JHEP* 1208, p. 034. arXiv: 0904.2715 (cit. on p. 2742).

Davide Gaiotto, Gregory W. Moore, and Andrew Neitzke (2009). “Wall-Crossing, Hitchin Systems, and the WKB Approximation”. arXiv: 0907.3987 (cit. on p. 2742).

Victor Ginzburg and David Kazhdan (n.d.). *Construction of symplectic varieties arising in ‘Sicilian theories’*. unpublished (cit. on p. 2749).

Rudolf Haag (1996). *Local quantum physics*. Second. Texts and Monographs in Physics. Berlin: Springer-Verlag, pp. xvi+390. MR: 1405610 (cit. on p. 2729).

Arthur Jaffe and Edward Witten (2006). “Quantum Yang-Mills theory”. In: *The millennium prize problems*. Clay Math. Inst., Cambridge, MA, pp. 129–152. MR: 2238278 (cit. on p. 2731).

Anton Kapustin (2010). "Topological Field Theory, Higher Categories, and Their Applications". arXiv: 1004.2307 (cit. on p. 2732).

Alexei Kitaev (n.d.). "Homotopy-Theoretic Approach to SPT Phases in Action: $\mathbb{Z}_{16}$ Classification of Three-Dimensional Superconductors" (cit. on p. 2735).

Davesh Maulik and Andrei Okounkov (2012). "Quantum Groups and Quantum Cohomology". arXiv: 1211.1287 (cit. on p. 2750).

Joseph A. Minahan and Dennis Nemeschansky (1996). "An $\mathcal{N} = 2$ Superconformal Fixed Point with E_6 Global Symmetry". *Nucl. Phys.* B482, pp. 142–152. arXiv: hep-th/9608047 (cit. on p. 2741).

– (1997). "Superconformal Fixed Points with E_N Global Symmetry". *Nucl. Phys.* B489, pp. 24–46. arXiv: hep-th/9610076 (cit. on p. 2741).

Gregory W. Moore and Yuji Tachikawa (2011). "On 2D TQFTs Whose Values are Holomorphic Symplectic Varieties". arXiv: 1106.5698 (cit. on p. 2749).

W. Nahm (1978). "Supersymmetries and Their Representations". *Nucl. Phys.* B135, p. 149 (cit. on p. 2734).

Olivier Schiffmann and Eric Vasserot (2012). "Cherednik algebras, W algebras and the equivariant cohomology of the moduli space of instantons on $\mathbb{A}^2$". arXiv: 1202.2756 (cit. on p. 2750).

Christoph Schweigert, Jurgen Fuchs, and Ingo Runkel (2006). "Categorification and Correlation Functions in Conformal Field Theory". arXiv: math/0602079 (cit. on p. 2735).

Graeme Segal (2004). "The definition of conformal field theory". In: *Topology, geometry and quantum field theory*. Vol. 308. London Math. Soc. Lecture Note Ser. Cambridge: Cambridge Univ. Press, pp. 421–577. MR: 2079383 (cit. on p. 2728).

Vladimir Turaev (2010). *Homotopy quantum field theory*. Vol. 10. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, pp. xiv+276. MR: 2674592 (cit. on p. 2735).

Edward Witten (1988). "Topological Quantum Field Theory". *Commun. Math. Phys.* 117, p. 353 (cit. on p. 2730).

Received 2017-12-15.

Yuji Tachikawa
Kavli Institute for the Physics and Mathematics of the Universe
University of Tokyo
Kashiwa, Chiba 277-8583
Japan
yuji.tachikawa@ipmu.jp

PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 3 (2751–2776)

$(2+1)$-DIMENSIONAL INTERFACE DYNAMICS: MIXING TIME, HYDRODYNAMIC LIMIT AND ANISOTROPIC KPZ GROWTH

FABIO TONINELLI

Abstract

Stochastic interface dynamics serve as mathematical models for diverse time-dependent physical phenomena: the evolution of boundaries between thermodynamic phases, crystal growth, random deposition... Interesting limits arise at large space-time scales: after suitable rescaling, the randomly evolving interface converges to the solution of a deterministic PDE (hydrodynamic limit) and the fluctuation process to a (in general non-Gaussian) limit process. In contrast with the case of $(1+1)$-dimensional models, there are very few mathematical results in dimension $(d+1), d \geq 2$. As far as growth models are concerned, the $(2+1)$-dimensional case is particularly interesting: Dietrich Wolf in 1991 conjectured the existence of two different universality classes (called KPZ and Anisotropic KPZ), with different scaling exponents. Here, we review recent mathematical results on (both reversible and irreversible) dynamics of some $(2+1)$-dimensional discrete interfaces, mostly defined through a mapping to two-dimensional dimer models. In particular, in the irreversible case, we discuss mathematical support and remaining open problems concerning Wolf's conjecture on the relation between the Hessian of the growth velocity on one side, and the universality class of the model on the other.

1 Introduction

Many phenomena in nature involve the evolution of interfaces. A first example is related to phenomena of deposition on a substrate, in which case the interface is the boundary of the deposed material: think for instance of crystal growth by molecular beam epitaxy or,

This work was partially funded by the ANR-15-CE40-0020-03 Grant LSD, by the CNRS PICS grant "Interfaces aléatoires discrètes et dynamiques de Glauber" and by MIT-France Seed Fund "Two-dimensional Interface Growth and Anisotropic KPZ Equation".

MSC2010: primary 82C20; secondary 60J10, 60K35, 82C24.

Keywords: Interacting particle systems, Dimer model, Interface growth, Hydrodynamic limit, Anisotropic KPZ equation, Stochastic Heat Equation.

closer to everyday experience, of the growth of a layer of snow during snowfall (see e.g. Barabási and Stanley [1995] for a physicist's introduction to growth phenomena). Another example is the evolution of the boundary between two thermodynamic phases of matter. Think of a block of ice immersed in water: the shape of the ice block, hence the water/ice boundary, changes with time and of course the dynamics is very different according to whether temperature is above, below or exactly at 0 °C.

A common feature of these examples is that on macroscopic (i.e. large) scales the interface evolution appears to be deterministic, while a closer look reveals that the interface is actually rough and presents seemingly random fluctuations (this is particularly evident in the snow example, since snowflakes have a visible size).

To try to model mathematically such phenomena, a series of simplifications are adopted. First, the so-called *effective interface approximation*: the d-dimensional interface in $(d+1)$-dimensional space is modeled as a height function $h : x \in \mathbb{R}^d \mapsto h_x(t) \in \mathbb{R}$, where $h_x(t)$ gives the height of the interface above point x at time t (think of $d = 2$ in the case of snow falling on your garden, but $d = 1$ for instance for snow falling and sliding down on your car window). This approximation implies that one ignores the presence of overhangs in the interface. (More often than not, the model is discretized and $\mathbb{R}^d, \mathbb{R}$ are replaced by $\mathbb{Z}^d, \mathbb{Z}$.) Secondly, in the usual spirit of statistical mechanics, the complex phenomena leading to microscopic interface randomness (e.g. chaotic motion of water molecules in the case of the ice/water boundary, or the various atmospheric phenomena determining the motion of individual snowflakes) are simplified into a probabilistic description where the dynamics of the height function is modeled by a Markov chain with simple, "local", transition rules.

We already mentioned that on macroscopic scales the interface evolution looks deterministic: this means that rescaling space as $\epsilon^{-1}x$, height as ϵh and time as $\epsilon^{-\alpha}t$ (we will discuss the scaling exponent $\alpha > 0$ later) and letting $\epsilon \to 0$, the random function $\epsilon h_{\epsilon^{-1}x}(\epsilon^{-\alpha}t)$ converges to a deterministic function $\phi(x,t)$ that in general is the solution of a certain non-linear PDE. This is called the *hydrodynamic limit* and is the analog of a law of large numbers for the sum of independent random variables. When we say that convergence holds, it does not mean it is easy to prove it or even to write down the PDE explicitly. Indeed, as discussed in more detail in the following sections, above dimension $(d+1) = (1+1)$ the hydrodynamic limit has been proved only for a handful of models, and one of the goals of this review is to report on recent results for $d = 2$.

On a finer scale, the interface fluctuations around the hydrodynamic limit are expected to converge, after proper rescaling, to a limit stochastic process, not necessarily Gaussian. In some situations, but not always, this is described via a Stochastic PDE. Again, while much is now known about $(1+1)$-dimensional models (for one-dimensional growth models and their relation with the so-called KPZ universality class, we refer to the recent reviews Corwin [2012] and Quastel [2012]), results are very scarce in higher dimension and we will present some recent ones for $d = 2$.

Before entering into more details of the models we consider, it is important to distinguish between two very different physical situations. In the case of deposition phenomena, the interface grows irreversibly and asymmetrically in one direction (say, vertically upward). The same is the case for the ice/water example if temperature is not 0 °C: for instance if $T > 0$ °C then ice melts and the water phase eventually invades the whole space. In these situations, the Markov process modeling the phenomenon is irreversible and the correct scaling for the hydrodynamic limit is the so-called hyperbolic or Eulerian one: the scaling exponent α introduced above equals 1. The situation is very different for the ice/water example exactly at 0 °C: in this case, the two coexisting phases are at thermal equilibrium and none is a priori favored. If the ice block occupied a full half-space then the flat water/ice interface would macroscopically not move and indeed a finite ice cube evolves only thanks to curvature of its boundary. In terms of hydrodynamic limit, one needs to look at longer time-scales than the Eulerian one: more precisely, one needs to take time of order $\epsilon^{-\alpha}, \alpha = 2$ (diffusive scaling).

We will discuss in more detail the Eulerian and diffusive cases, together with the new results we obtained for some $(2+1)$-dimensional models, in Sections 2 and 3 respectively. A common feature of all our recent results is that the interface dynamics we analyze can be formulated as dynamics of dimer models on bipartite planar graphs, or equivalently of tilings of the plane. See Figure 1 for a randomly sampled lozenge tiling of a planar domain.

Such models have a family of translation-invariant Gibbs measures, with an integrable (actually determinantal) structure Kenyon [2009], that play the role of stationary states for the dynamics.

2 Stochastic interface growth

In a stochastic growth process, the height function $h(t) = \{h_x(t)\}_{x\in\mathbb{Z}^d}$ evolves asymmetrically, i.e. has an average non-zero drift, say positive. For instance, growth can be totally asymmetric: only moves increasing the height are allowed. It is then obvious that such Markov chain cannot have an invariant measure. One should look at interface gradients $\nabla h(t) = \{h_x(t) - h_{x_0}(t)\}_{x\in\mathbb{Z}^d}$ instead, where x_0 is some reference site (say the origin). Since the growth phenomenon we want to model satisfies vertical translation invariance, the transition rate at which h_x jumps, say, to $h_x + 1$ depends only on the interface gradients (say, the gradients around x) and not on the absolute height $h(t)$. Therefore, the projection of the Markov chain $h(t)$ obtained by looking at the evolution of $\nabla h(t)$ is still a Markov chain. For natural examples one expects that given a slope

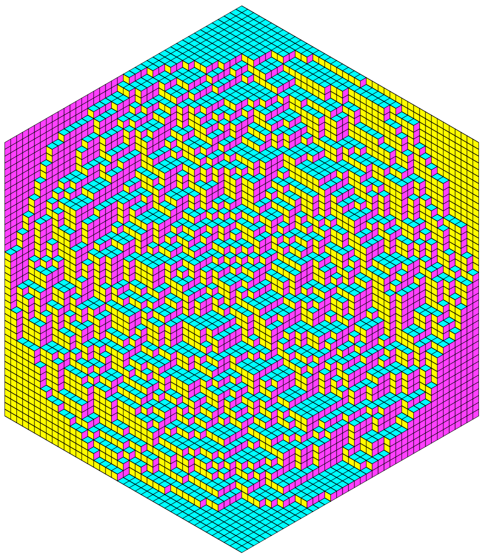


Figure 1: A random, uniformly chosen, lozenge tiling of a hexagon Λ_ϵ of diameter ϵ^{-1}. As $\epsilon \to 0$, the corresponding random interface presents six "frozen regions" near the corners and a "liquid region" inside the arctic circle Cohn, Larsen, and Propp [1998].

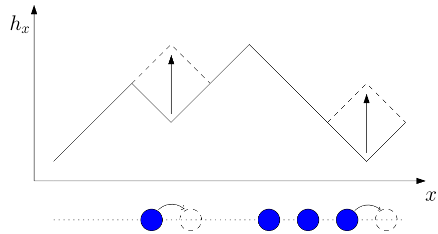


Figure 2: In the corner-growth process, heights increase by 1, with transition rate equal to 1, at local minima. Interpreting a negative gradient $h_{x+1} - h_x$ as a particle and a negative one as a hole, the dynamics of the gradients is the TASEP: particles try independently length-1 jumps to the right, with rate 1, subject to an exclusion constraint (at most one particle per site is allowed).

$\rho \in \mathbb{R}^d$, there exists a unique translation-invariant stationary state μ_ρ for the gradients, with the property that $\mu_\rho(h_{x+e_i} - h_x) = \rho_i, i = 1, \dots, d$. A very well known example is the $(1+1)$-dimensional corner growth model: the evolution of interface gradients is just the 1-dimensional Totally Asymmetric Simple Exclusion (TASEP), whose invariant measures are iid Bernoulli product measures labelled by the particle density ρ. See Figure 2.

If the initial height profile is sampled from μ_ρ (more precisely, the height gradients are sampled from μ_ρ, while the height h_{x_0} is assigned some arbitrary value, say zero), then on average the height increases exactly linearly with time:

$$\mathbb{E}_{\mu_\rho}(h_x(t) - h_x(0)) = v(\rho)t, \quad v(\rho) > 0. \tag{2-1}$$

Now suppose that the initial height profile is instead close to some non-affine profile ϕ_0, i.e.

$$\epsilon h_{\epsilon^{-1}x} \stackrel{\epsilon \to 0}{\to} \phi_0(x), \quad \forall x \in \mathbb{R}. \tag{2-2}$$

Then, one expects that, under so-called hyperbolic rescaling of space-time where $x \to \epsilon^{-1}x, t \to \epsilon^{-1}t$, one has

$$\epsilon h_{\epsilon^{-1}x}(\epsilon^{-1}t) \xrightarrow[\epsilon \to 0]{\mathbb{P}} \phi(x,t) \tag{2-3}$$

where $\phi(x,t)$ is non-random and solves the first order PDE of Hamilton-Jacobi type

$$\partial_t \phi(x,t) = v(\nabla \phi(x,t)) \tag{2-4}$$

with $v(\cdot)$ *the same function* as in (2-1). A couple of remarks are important here:

- Unless $v(\cdot)$ is a linear function (which is a very uninteresting case), the PDE (2-4) develops singularities in finite time. Then, one expects $\phi(x,t)$ to solve (2-4) in the sense of vanishing viscosity.

- Viscosity solutions of (2-4) are well understood when $v(\cdot)$ is convex, since they are given by the variational Hopf-Lax formula. However, there is no fundamental reason why $v(\cdot)$ should be convex (this will be an important point in next section). Then, much less is known on the analytic side, aside from basic properties of existence and uniqueness.

- The example of the TASEP is very special in that invariant measures μ_ρ are explicitly known. One should keep in mind that this is an exception rather than the rule and that most examples with known stationary measures are $(1+1)$-dimensional. As a consequence, the function $v(\cdot)$ in (2-4) is in general unknown.

Next, let us consider fluctuations in the stationary process started from μ_ρ. On heuristic grounds, one expects height fluctuations $\hat{h}_x(t)$ with respect to the average, linear, height profile $\mathbb{E}_{\mu_\rho}(h_x(t)) = \langle \rho, x\rangle + v(\rho)t$ to be somehow described, *on large space-time scales*, by a stochastic PDE (KPZ equation) of the type Kardar, Parisi, and Zhang [1986]

$$\partial_t \psi(x,t) = \nu \Delta \psi(x,t) + \langle \nabla \psi(x,t), H_\rho \nabla \psi(x,t)\rangle + \xi(x,t), \tag{2-5}$$

where:

- the Laplacian is a diffusion term that tends to locally smooth out fluctuations and $\nu > 0$ is a model-dependent constant;

- the $d \times d$ symmetric matrix H_ρ is the Hessian of the function $v(\cdot)$ computed at ρ and $\langle \cdot, \cdot \rangle$ denotes scalar product in $\mathbb{R}^d$. This non-linear term comes just from expanding to second order[1] the hydrodynamic PDE (2-4) around the flat solution of slope ρ.

- $\xi(x,t)$ is a space-time noise that models the randomness of the Markov evolution. It is well known that Eq. (2-5) is extremely singular if ξ is a space-time white noise (Hairer's theory of regularity structures Hairer [2014] gives a meaning to the equation for $d = 1$ but not for $d > 1$). Since however we are interested in properties on large space-time scales and since lattice models have a natural "ultraviolet" space cut-off of order 1 (the lattice spacing), we can as well imagine that the noise is not white in space and its correlation function has instead a decay length of order 1. As a side remark, in the physics literature (e.g. Kardar, Parisi, and Zhang [1986],

[1]The first-order term $\langle \nabla v(\rho), \nabla \psi(x,t)\rangle$ in the expansion is omitted because it can be absorbed into $\partial_t \psi$ via a linear (Galilean) transformation of space-time coordinates.

Wolf [1991], and Barabási and Stanley [1995]) the presence of a noise regularization in space is implicitly understood, and explicitly used in the renormalization group computations: this is the cut-off $\Lambda = 1$ that appears e.g. in Barabási and Stanley [1995, App. B].

One should not take the above conjecture in the literal sense that the law of the space-time fluctuation process $\hat{h}_x(t)$ converges to the law of the solution of (2-5). Only the large-scale correlation properties of the two should be asymptotically equivalent.

For $(1+1)$-dimensional models like the TASEP and several others, a large amount of mathematical results by now supports the following picture[2]: starting say with the deterministic condition $h_x(0) \equiv 0$, the standard deviation of $h_x(t)$ grows as t^β, $\beta = 1/3$, the space correlation length grows like $t^{1/z}$, where $z = 3/2$ is the so-called dynamic exponent and the fluctuation field $\hat{h}_x(t)$ rescaled accordingly tends as $t \to \infty$ to a (non-Gaussian) limit process. We do not enter into any more detail for $(1+1)$-dimensional models of the KPZ class here, see for instance the reviews Quastel [2012] and Corwin [2012]; let us however note that this behavior is very different from the (Gaussian) one of the stochastic heat equation with additive noise (called "Edwards-Wilkinson equation" in the physics literature), obtained by dropping the non-linear term in (2-5).

On the other hand, for $(d+1)$-dimensional models, $d \geq 3$, renormalization group computations Kardar, Parisi, and Zhang [1986] applied to the stochastic PDE (2-5) suggest that, if the non-linear term is sufficiently small (in terms of the microscopic growth model: if the speed function $v(\cdot)$ is sufficiently close to an affine function) then non-linearity is irrelevant, meaning that the large-scale fluctuation properties of the model (or of the solution of (2-5)) are asymptotically the same as those of the stochastic heat equation: these models belong to the so-called Edwards-Wilkinson universality class. There is very recent mathematical progress in this direction: indeed, Magnen and Unterberger [2017] states that for $d \geq 3$ the solution of (2-5) tends on large space-time scales to the solution of the Edwards-Wilkinson equation, if $H_\rho = \lambda \mathbb{I}$ with $\mathbb{I}$ the $d \times d$ identity matrix and λ small enough. See also Gu, Ryzhik, and Zeitouni [2017] where similar results are stated for the $d \geq 3$ dimensional stochastic heat equation with multiplicative noise, that is obtained from (2-5) via the Cole-Hopf transform.

The situation is richer in the borderline case of the critical dimension $d = 2$, to which the next two sections are devoted.

2.1 $(2+1)$**-dimensional growth: KPZ and Anisotropic KPZ (AKPZ) classes.** For $(1+1)$-dimensional models, the non-linear term in (2-5) equals $v''(\rho)(\partial_x \psi(x,t))^2$: multiplying ψ by a suitable constant, we can always replace $v''(\rho) \neq 0$ by a positive constant.

[2]As (2-5) suggests, for the following to hold one needs $v''(\rho) \neq 0$, otherwise the fluctuation process should be described simply the linear stochastic heat equation with additive noise.

The picture is richer for $d > 1$, and in particular in the case $d = 2$ we consider here. In fact, one should distinguish two cases:

1. (Isotropic) KPZ class: $\det(H_\rho) > 0$ (strictly);

2. Anisotropic KPZ (AKPZ) class: $\det(H_\rho) \leq 0$.

According to whether a growth model has a speed function $v(\cdot)$ whose Hessian satisfies the former or latter condition, the large-scale behavior of its fluctuations is conjectured to be very different.

The **isotropic KPZ class** is the one considered in the original KPZ work Kardar, Parisi, and Zhang [1986]. In this case, perturbative renormalization-group arguments[3] suggest that fluctuations of $h_x(t)$ (or of the solution $\psi(x,t)$ of (2-5)) grow in time like t^β and that, in the stationary states, fluctuations grow in space as $\mathrm{Var}_{\mu_\rho}(h_x - h_y) \sim |x-y|^{2\alpha}$, with two exponents $\beta > 0, \alpha = 2\beta/(\beta+1)$ that are *different* from those of the Edwards-Wilkinson equation: non-linearity is said to be *relevant*[4]. The Edwards-Wilkinson equation can be solved explicitly and in two dimensions one finds $\alpha_{EW} = \beta_{EW} = 0$ (growth in time and space is only logarithmic; the stationary state is the (log-correlated) massless Gaussian field). The values of α, β for the isotropic KPZ equation cannot be guessed by perturbative renormalization-group arguments and they are accessible only through numerical experiments (see discussion below). Note that $\alpha > 0$ means that stationary height profiles are much rougher than a lattice massless Gaussian field.

The **Anisotropic KPZ** case was analyzed later by Wolf [1991] with the same renormalization group approach and the result came out as a surprise: non-linearity turns out to be non-relevant in this case, i.e., the growth exponents α, β are predicted to be 0 as for the Edwards-Wilkinson equation.

Let us summarize this discussion into a conjecture:

Conjecture 2.1. *Let $v(\cdot)$ be the speed function of a (reasonable) $(2+1)$-dimensional growth model. If $\det(H_\rho) > 0$ with H_ρ the Hessian of $v(\cdot)$ computed at ρ, then height fluctuations grow in time as t^β for some model-independent $\beta > 0$ and height fluctuations in the stationary states grow as distance to the power $2\beta/(\beta+1)$. If instead $\det(H_\rho) \leq 0$, then $\beta = \alpha = 0$ and the stationary states have the same height correlations in space as a massless Gaussian field.*

[3] "perturbative" here means that, if we imagine that the non-linear term in (2-5) has a prefactor λ, then one expands the solution around the linear $\lambda = 0$ case, keeping only terms up to order $O(\lambda^2)$.

[4] The relation $\alpha = 2\beta/(\beta+1)$ is another way of writing a scaling relation between exponents that is usually written as $\alpha + z = 2$ where $z = \alpha/\beta$. Here z is the so-called dynamic exponent that equals $3/2$ for one-dimensional KPZ models.

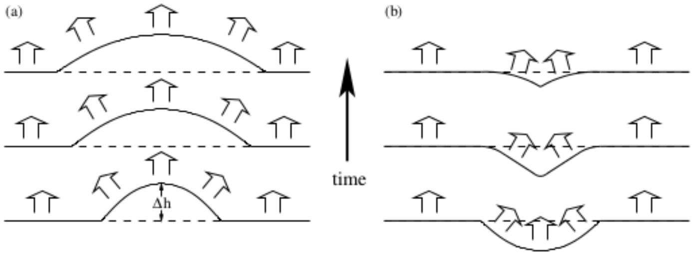


Figure 3: Left: The time evolution of a positive bump under equation (2-6), when $\det(H_\rho) > 0$. The height Δh of the bump is constant in time while its width grows as $t^{1/2}$. Right: a negative bump, on the other hand, develops a cusp and its height decreases as $1/t$. This figure is taken from Prähofer [2003].

Let us review the evidence in favor of this conjecture, apart from the renormalization-group argument of Kardar, Parisi, and Zhang [1986] and Wolf [1991] that does not provide much intuition and seems very hard to be turned into a mathematical proof:

1. A somewhat rough but suggestive argument that sheds some light on Conjecture 2.1 is given in Prähofer [2003, Sec. 2.2]. One imagines that in the evolution of the fluctuation field ψ there are two effects. Thermal noise adds random positive or negative "bumps", at random times, to the initially flat height profile; each bump then evolve following the hydrodynamic equation, expanded to second order:

$$\partial_t \psi = \langle \nabla \psi(x,t), H_\rho \nabla \psi(x,t) \rangle. \tag{2-6}$$

   It is not hard to convince oneself that, under (2-6), if both eigenvalues of H_ρ are, say, strictly positive, then a positive bump grows larger with time and a negative bump shrinks (the reverse happens if the eigenvalues of H_ρ are both negative). See Figure 3. On the other hand, if $\det(H_\rho) < 0$ (so that the eigenvalues of H_ρ have opposite signs) then a positive bump spreads in the direction where the curvature of $v(\cdot)$ is positive, but its height shrinks because of the concavity of $v(\cdot)$ in the other direction (the same argument applies to negative bumps). Then, it is intuitive that when $\det(H_\rho) < 0$ height fluctuations should grow slower with time than when $\det(H_\rho) > 0$, where the effects of spreading positive bumps accumulate.

2. There exist some growth models that satisfy a so-called "envelope property", saying essentially that given two initial height profiles $\{h_x^{(j)}\}_{x\in\mathbb{Z}^d}$, $j = 1, 2$, one can find a coupling between the corresponding profiles $\{h_x^{(j)}(t)\}_{x\in\mathbb{Z}^d}$ at time t

such that the evolution started from the profile $\check{h} := \{\max(h_x^{(1)}, h_x^{(2)})\}_{x\in\mathbb{Z}^d}$ equals $\check{h}(t) = \{\max(h_x^{(1)}(t), h_x^{(2)}(t))\}_{x\in\mathbb{Z}^d}$. One example is the $(2+1)$-dimensional corner-growth model analogous to that of Figure 2 except that the interface is two-dimensional and unit cubes instead of unit squares are deposed with rate one on it. For growth models satisfying the envelope property, a super-additivity argument implies that the hydrodynamic limit (2-3) holds and moreover that the function $v(\cdot)$ in (2-4) is convex Seppäläinen [2000] and Rezakhanlou [2002]. While for $(2+1)$-dimensional models in this class the stationary measures μ_ρ and the function $v(\cdot)$ cannot be identified explicitly, convexity implies (at least in the region of slopes where $v(\cdot)$ is smooth and strictly convex) that $\det(H_\rho) > 0$: these models must belong to the isotropic KPZ class. The $(2+1)$-dimensional corner-growth model was studied numerically in Tang, Forrest, and Wolf [1992] and it was found, in agreement with Conjecture 2.1, that $\beta \simeq 0.24$ (the numerics is sufficiently precise to rule out the value $1/4$ which was conjectured in earlier works). The same value for β is found numerically Halpin-Healy and Assdah [1992] from direct simulation of (a space discretization of) the stochastic PDE (2-5) with $\det(H_\rho) > 0$.

3. For models in the AKPZ class there is no chance to get the hydrodynamic limit by simple super-additivity arguments since, as we mentioned, $v(\cdot)$ would turn out to be convex. On the other hand, as we discuss in more detail in next section, there exist some $(2+1)$-dimensional growth models for which the stationary measures μ_ρ can be exhibited explicitly, and they turn out to be of massless Gaussian type, with logarithmic growth of fluctuations: $\alpha = 0$. For such models, one can prove also that $\beta = 0$ and one can compute the speed function $v(\cdot)$. In all the known examples, a direct computation shows that $\det(H_\rho) < 0$, as it should according to Conjecture 2.1.

Remark 2.2. *Let us emphasize that, in general, it is not possible to read a priori, from the generator of the process, the convexity properties of the speed function $v(\cdot)$, and therefore its universality class. This is somehow in contrast with the situation in equilibrium statistical mechanics, where usually the universality class of a model can be guessed from symmetries of its Hamiltonian. It is even conceivable, though we are not aware of any concrete example, that there exist growth models for which the sign of* $\det(H_\rho)$ *depends on ρ.*

2.2 Mathematical results for Anisotropic KPZ growth models. As we already mentioned, there are no results other than numerical simulations or non-rigorous arguments supporting the part of Conjecture 2.1 concerning the isotropic KPZ class. Fortunately, the

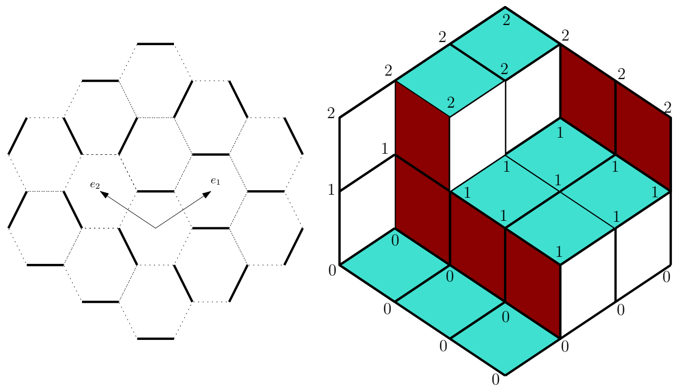


Figure 4: A perfect matching of the hexagonal lattice (left) and the corresponding lozenge tiling (right). Near each lozenge vertex is given the height of the interface w.r.t. the horizontal plane. For clarity let us emphasize that, while we draw only a finite portion of the matching/tiling, one should imagine that it extends to a matching/tiling of the infinite graph/plane.

situation is much better for the AKPZ class, which includes several models that are to some extent "exactly solvable".

Several of the AKPZ models for which mathematical results are available have a height function that can be associated to a two-dimensional dimer model (an exception is the Gates-Westcott model solved by Prähofer and Spohn [1997]). Let us briefly recall here a few well-known facts on dimer models (we refer to Kenyon [2009] for an introduction). For definiteness, we will restrict our discussion to the dimer model on the infinite hexagonal graph but most of what we say about the height function and translation-invariant Gibbs states extends to periodic, two-dimensional bipartite graphs (say, $\mathbb{Z}^2$). A (fully packed) dimer configuration is a perfect matching of the graph, i.e., a subset M of edges such that each vertex of the graph is contained in one and only one edge in M; as in Figure 4, in the case of the hexagonal graph the matching can be equivalently seen as a lozenge tiling of the plane and also as a monotone discrete two-dimensional interface in three dimensional space. "Monotone" here means that the interface projects bijectively on the plane $x + y + z = 0$. The height function is naturally associated to vertices of lozenges, i.e. to hexagonal faces. We will use the dimer, the tiling or the height function viewpoint interchangeably. If on the graph we choose coordinates $x = (x_1, x_2)$ according to the axes e_1, e_2 drawn in Figure 4, it is easy to see that the overall slope $\rho = (\rho_1, \rho_2)$ of the interface must belong to the triangle $\mathbb{T} \subset \mathbb{R}^2$ defined by the inequalities $0 \le \rho_1 \le 1, 0 \le \rho_2 \le 1, 0 \le \rho_1 + \rho_2 \le 1$. It is known that, given ρ in the interior

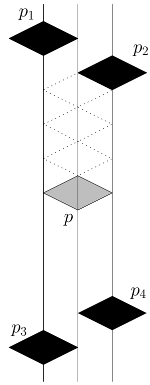


Figure 5: Each particle (or horizontal lozenge) p is constrained between its four neighboring particles $p_1, \dots p_4$. The three positions particle p can jump to (with rate 1) in the Borodin-Ferrari dynamics are dotted.

of $\mathbb{T}$, there exists a unique translation-invariant ergodic Gibbs state π_ρ of slope ρ. That is, π_ρ is a (translation invariant, ergodic) probability measure on dimer configurations of the infinite graph, such that the average height slope is ρ and such that, conditionally on the configuration outside any finite domain Λ, the law of the configuration inside Λ is uniform over all dimer configurations compatible with the outside (DLR condition). In fact, much more is known: as a consequence of Kasteleyn's theory Kasteleyn [1961], such measures have a determinantal representation. That is, the probability of a cylindrical event of the type "k given edges are occupied by dimers" is given by the determinant of a $k \times k$ matrix, whose elements are the Fourier coefficients of an explicit function on the two-dimensional torus $\{(z, w) \in \mathbb{C}^2 : |z| = |w| = 1\}$ Kenyon, Okounkov, and Sheffield [2006]. Thanks to this representation, much can be said about large-scale properties of the measures π_ρ. Notably, correlations decay like the inverse distance squared and the height function scales to a massless Gaussian field with logarithmic covariance structure.

Now that we have a nice candidate for a $(2+1)$-dimensional height function, we go back to the problem of defining a growth model that would hopefully be mathematically treatable and shed some light on Conjecture 2.1. To this purpose, let us remark first of all that, to a lozenge tiling as in Figure 4, one can bijectively associate a *two-dimensional system of interlaced particles*. For this purpose, we will call "particles" the horizontal (or blue) lozenges (the positions of the others are uniquely determined by these) and we note that particle positions along a vertical column are interlaced with those of the two neighboring columns. See Figure 5. A first natural candidate for a growth process would be the following immediate generalization of the TASEP: each particle jumps $+1$ vertically, with rate 1, provided the move does not violate the interlacement constraints. Actually, this is nothing but the three-dimensional corner-growth model. As we already mentioned, this should belong to the isotropic KPZ class and its stationary measures μ_ρ should be extremely different from the Gibbs measures π_ρ, with a power-like instead of logarithmic growth of fluctuations in space. Unfortunately, none of this could be mathematically proved so far.

In the work Borodin and P. L. Ferrari [2014], they considered instead another totally asymmetric growth model where each particle can jump an unbounded distance n upwards, with rate independent of n (say, rate 1), provided the interlacements are still satisfied after the move. See Figure 5. The situation is then entirely different with respect to the corner-growth process: the two processes belong to two different universality classes. If the initial condition of the process is a suitably chosen, deterministic, fully packed particle arrangement (see Fig. 1.1 in Borodin and P. L. Ferrari [ibid.]), it was shown that the height profile rescaled as in (2-3) does converge to a deterministic limit $\phi(\cdot, t)$, that solves the Hamilton-Jacobi equation (2-4) with

$$v(\rho) = \frac{1}{\pi} \frac{\sin(\pi\rho_1)\sin(\pi\rho_2)}{\sin(\pi(\rho_1+\rho_2))}. \tag{2-7}$$

A couple of remarks are important for the subsequent discussion:

- with the initial condition chosen in Borodin and P. L. Ferrari [ibid.], $\phi(x, t)$ turns out to be a *classical* solution of (2-4). That is, the characteristic lines of the PDE do not cross at positive times (we emphasize that this is due to the specific form of the chosen initial profile and not to the form of $v(\cdot)$).

- An explicit computation shows that $\det(H_\rho) < 0$ for the function (2-7): this growth model is then a candidate to belong to the AKPZ class.

As mentioned in Remark 2.2 above, let us emphasize that we see no obvious way to guess a priori that the corner growth process and the "long-jump" one should belong to different universality classes.

Various other results were proven in Borodin and P. L. Ferrari [ibid.], but let us mention only two of them, that support the conjecture that this model indeed belongs to the AKPZ class:

1. the fluctuations of $h_{\epsilon^{-1}x}(\epsilon^{-1}t)$ around its average value are of order $\sqrt{\log 1/\epsilon}$ (the growth exponent is $\beta = 0$) and, once rescaled by this factor, they tend to a Gaussian random variable;

2. the local law of the interface gradients at time $\epsilon^{-1}t$ around the point $\epsilon^{-1}x$ tends to the Gibbs measure π_ρ with $\rho = \nabla\phi(x, t)$.

Remark 2.3. *The basic fact behind the results of Borodin and P. L. Ferrari [2014] is that for the specific choice of initial condition, one can write Borodin and P. L. Ferrari [ibid., Th. 1.1] the probability of certain events of the type "there is a particle at position x_i at time $t_i, i \leq k$" as a $k \times k$ determinant, to which asymptotic analysis can be applied. The same determinantal properties hold for other "integrable" initial conditions, but they are not at all a generic fact.*

Point (2) above clearly suggests that the Gibbs measures π_ρ should be stationary states for the interface gradients. In fact, this is a result I later proved:

Theorem 2.4. *Toninelli [2017, Th. 2.4] For every slope ρ in the interior of $\mathbb{T}$, the measure π_ρ is stationary for the process of the interface gradients.*

Recall that, as discussed above, the Gibbs measures π_ρ of the dimer model have the large-scale correlation structure of a massless Gaussian field and indeed Conjecture 2.1 predicts that stationary states of AKPZ growth processes behave like massless fields. Most of the technical work in Toninelli [ibid.] is related to the fact that, since particles can perform arbitrarily long jumps with a rate that *does not decay* with the jump length, it is not clear a priori that the process exists at all: one can exhibit initial configurations such that particles jump to $+\infty$ in finite time (this issue does not arise in the work Borodin and P. L. Ferrari [2014] where, thanks to the chosen initial condition, there is no difficulty in defining the infinite-volume process). In Toninelli [2017] it is shown via a comparison with the one-dimensional Hammersley process Seppäläinen [1996] that, for a typical initial condition sampled from π_ρ, particles jump almost surely a finite distance in finite time and that, despite the unbounded jumps, perturbations do not spread instantaneously through the system. This means that if two initial configurations differ only on a subset S of the lattice, their evolutions can be coupled so that at finite time t they are with high probability equal sufficiently far away from S (how far, depending on t).

To follow the general program outlined above, once the stationary states are known, one would like to understand the growth exponent β for the stationary process. We proved the following, implying $\beta = 0$:

Theorem 2.5. *Toninelli [2017, Th. 3.1] For every lattice site x, we have*

$$\limsup_{t\to\infty} \mathbb{P}_{\pi_\rho}\left(|h_x(t) - \mathbb{E}_{\pi_\rho}(h_x(t))| \geq u\sqrt{\log t}\right) \stackrel{u\to\infty}{\longrightarrow} 0 \tag{2-8}$$

where

$$\mathbb{E}_{\pi_\rho}(h_x(t)) = v(\rho)t + \langle x, \rho\rangle. \tag{2-9}$$

To be precise:

- in the statement of Toninelli [ibid., Th. 3.1] there is a technical restriction on the slope ρ, that was later removed in joint work with Chhita, P. L. Ferrari, and Toninelli [2017];
- the proof that the speed of growth $v(\cdot)$ in (2-9) is the same as the function $v(\cdot)$ in (2-7), as it should, was obtained by Chhita and P. Ferrari [2017] and requires a nice combinatorial property of the Gibbs measures π_ρ.

With reference to Remark 2.3 above, it is important to emphasize that there is no known determinantal form for the space-time correlations of the stationary process; for the proof of (2-8) we used a more direct and probabilistic method.

Finally, it is natural to try to obtain a hydrodynamic limit for the height profile. Recall that in Borodin and P. L. Ferrari [2014] such a result was proven for an "integrable" initial condition that allowed to write certain space-time correlations, and as a consequence the average particle currents, in determinantal form. On the other hand, convergence to the hydrodynamic limit should be a very robust fact and not rely on such special structure. We have indeed:

Theorem 2.6. *Legras and Toninelli [2017, Th. 3.5 and 3.6] Let the initial height profile satisfy* (2-2)*, with ϕ_0 a Lipshitz function with gradient in the interior of $\mathbb{T}$. Let one of the following two conditions be satisfied:*

- *ϕ_0 is C^2 and the time t is smaller than T, the maximal time up to which a classical solution of* (2-4) *exists;*
- *ϕ_0 is either convex or concave (in which case we put no restriction on t).*

Then, the convergence (2-3) *holds, with $\phi(x,t)$ the viscosity solution of* (2-4).

The restriction to either small times or to convex/concave profile is due to the fact that we have in general little analytic control on the singularities of (2-4), due to the non-convexity of $v(\cdot)$. For convex initial profile, the viscosity solution of the PDE is given by a Hopf variational form and this allows to bypass these analytic difficulties. Let us emphasize, to avoid any confusion, that even in the case of convex initial profile the solution *does* in general develop singularities (shocks), i.e. discontinuities in space of the gradient $\nabla\phi(x,t)$.

Open problem 2.7. *Are height fluctuations still $O(\sqrt{\log t})$ at the location of shocks?*

Another important observation is the following. Given that we know explicitly the stationary states of the process and that the dynamics is monotone (i.e, if an initial profile is higher than another, under a suitable coupling it will stay higher as time goes on), it is

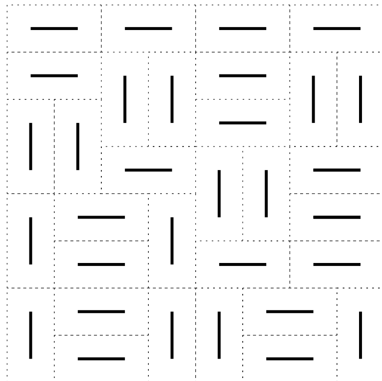


Figure 6: A perfect matching of the square lattice and the corresponding domino tiling (dotted). See Chhita, P. L. Ferrari, and Toninelli [2017, Fig. 2] for the definition of "particles" and their interlacing relations, and Kenyon [2009] for the definition of the height function.

tempting to try to apply the method developed by Rezakhanlou [2001], that gives under such circumstances convergence of the height profile of a growth model to the viscosity solution of the limit PDE. The delicate point is however that Rezakhanlou [ibid.] crucially requires that perturbations spread at finite speed through the system, so that one can analyze the evolution "locally", in small enough windows where the profile can be approximated by one sampled from μ_ρ, with suitably chosen slope ρ that depends on the window location. Due to unboundedness of particle jumps, however, the "finite-speed propagation property" might fail in our case and in any case it cannot hold uniformly for all initial conditions. Most of the technical work in Legras and Toninelli [2017] is indeed devoted to proving that one can localize the dynamics despite the long jumps. A crucial fact is that we show that the growth process under consideration can be reformulated through a so-called Harris-like graphical construction.

2.2.1 Extensions and open problems. There are various ways how the "lozenge tiling dynamics with long particle jumps" of previous section can be generalized to provide other $(2+1)$-dimensional growth processes in the AKPZ class. One such generalization was given in Toninelli [2017, Sec. 3.1]. There, one starts with the observation that: (i) as was the case for lozenge tilings, also domino tilings of $\mathbb{Z}^2$ (dominoes being 2×1 rectangles, horizontal or vertical, see Figure 6) have a natural height function interpretation, and (ii) a domino tiling can be bijectively mapped to a two-dimensional system of interlaced particles (interlacement constraints are different than in lozenge case).

This suggests a growth process where particles jump in an asymmetric fashion and the transition rate is independent of the jump length, jumps being limited only by the interlacement constraints. Then, the same results that were proven for the lozenge dynamics (notably, stationarity of the Gibbs measures, logarithmic correlations in space in the stationary states ($\alpha = 0$) and logarithmic growth of fluctuations in the stationary process implying $\beta = 0$) hold in this case too. The speed of growth $v(\cdot)$ for the domino dynamics was later computed in a joint work with Chhita, P. L. Ferrari, and Toninelli [2017, Th. 2.3]: it turns out to be rather more complicated than (2-7), but it is still an explicit function for which one can prove with some effort that the Hessian has negative determinant, in agreement with Conjecture 2.1.

Finally, there is yet another class of driven two-dimensional interlaced particle systems, that was introduced in Borodin, Bufetov, and Olshanski [2015]. While these have rather a group-theoretic motivation, these processes can also be viewed as $(2+1)$-dimensional growth models and actually the main result of Borodin, Bufetov, and Olshanski [ibid.] can be seen as a hydrodynamic limit for the height function Borodin, Bufetov, and Olshanski [ibid., Sec. 3.3]. Once again, direct inspection of the Hessian of the velocity function shows that these models belong to the AKPZ class[5], so these provide other natural candidates where Wolf's prediction of logarithmic growth of fluctuations can be tested (the logarithmic nature of fluctuation correlations is conjectured in Borodin, Bufetov, and Olshanski [ibid.]; we are not aware of an actual proof).

In conclusion, there are now quite a few $(2+1)$-dimensional growth models in the AKPZ class for which Wolf's predictions in Conjecture 2.1 can be verified. There is however one aspect one may find rather unsatisfactory. Both for the lozenge tiling dynamics, where the speed function turns out to be given by (2-7) and for its domino tiling generalization, where $v(\cdot)$ is a much more complicated-looking combination of ratios of trigonometric functions (see Chhita, P. L. Ferrari, and Toninelli [2017, Eq. (2.6)]) and also for the interlaced particle dynamics of Borodin, Bufetov, and Olshanski [2015], one verifies via brute-force computation that the Hessian of H_ρ of the corresponding velocity function $v(\cdot)$ has negative determinant. The frustrating fact is that via the explicit computation one does not see at all how the sign of the determinant the Hessian is related to the model being in the Edwards-Wilkinson universality class! We are still far from having a meta-theorem saying "if the exponents α and β are zero, then the determinant of the Hessian is negative". Up to now, we have essentially heuristic arguments and "empirical evidence" based on a few mathematically treatable models.

[5]For the growth models of Borodin, Bufetov, and Olshanski [2015], the determinant of the Hessian of the speed was computed by Weixin Chen, as mentioned in the unpublished work http://math.mit.edu/research/undergraduate/spur/documents/2012Chen.pdf

Open problem 2.8. *It would be very interesting to prove that the Hessian of the velocity function for the growth models just mentioned has negative determinant* without *going through the explicit computation of the second derivatives. For recent progress in this direction see Chhita and Toninelli [2018], where a complex-analytic argument is devised to prove* $\det(H_\rho) < 0$ *for an AKPZ growth model.*

2.2.2 Slow decorrelation along the characteristics. The results we discussed in Sections 2.2 and 2.2.1 (growth of fluctuation variance with time and spatial correlations in the stationary state) concern fluctuation properties *at a single time*. Another question of great interest is how fluctuations at different space-time points (x_i, t_i) are correlated. For $(1+1)$-dimensional growth models in the KPZ class, the following picture has emerged P. L. Ferrari [2008]: correlation decay slowly along the characteristic lines of the PDE (2-4), and faster along any other direction. For instance, take two space-time points (x_1, t_1) and (x_2, t_2), $t_1 < t_2$ and think of t_2 large. If the two points are on the same characteristic line, then the height fluctuations (divided by the rescaling factor $t_i^\beta = t_i^{1/3}$) will be almost perfectly correlated as long as $t_2 - t_1 \ll t_2$. If instead the two points are *not* along a characteristic line, then correlation will be essentially zero as soon as $t_2 - t_1 \gg t_2^{1/z}$, $z = 3/2$.

It has been conjectured P. L. Ferrari [2008] and Borodin and P. L. Ferrari [2014] that a similar phenomenon of slow decorrelation along the characteristic lines should occur for $(2+1)$-dimensional growth. For the AKPZ models described in the previous sections, it is still an open problem to prove anything in this direction. In the work Borodin, Corwin, and Toninelli [2017] in collaboration with A. Borodin and I. Corwin, we studied a growth model that depends on a parameter $q \in [0, 1)$: for $q = 0$ it reduces to the long-jump lozenge dynamics of Borodin and P. L. Ferrari [2014] and Toninelli [2017], while if $q \to 1$ and particle distances are suitably rescaled the dynamics simplifies in that fluctuations become Gaussian. In this limit, we were able to prove that, if height fluctuations are computed along characteristic lines, their correlations converge to those of the Edwards-Wilkinson equation and in particular they are large as long as $t_2 - t_1 \ll t_2$. If correlations are computed instead along a different direction, then they are essentially zero as soon as $t_2 - t_1 \gg t_2^{1/z}$, where $z = 2$ is the dynamic exponent of the Edwards-Wilkinson equation.

3 Interface dynamics at thermal equilibrium

Let us now move to reversible interface dynamics (we refer to Spohn [1993] and Funaki [2005] for an introduction). We can imagine that the interface is defined on a finite subset Λ_ϵ of $\mathbb{Z}^d$ of diameter $O(\epsilon^{-1})$, say the cubic box $[0, \dots, \epsilon^{-1}]^d$ so that after the rescaling $x = \epsilon^{-1}\xi$, the space coordinate ξ is in the unit cube. We impose Dirichlet boundary conditions, i.e., for $x \in \partial\Lambda_\epsilon$ the height $h_x(t)$ is fixed to some time-independent value $\bar{h}_x$.

The way to model the evolution of a phase boundary at thermal equilibrium is to take a Markov process with stationary and reversible measure of the Boltzmann-Gibbs form (we absorb the inverse temperature into the potential V)

$$\pi_{\Lambda_\epsilon}(h) \propto e^{-1/2 \sum_{x \sim y} V(h_x - h_y)} \tag{3-1}$$

where the sum runs, say, on nearest neighboring pairs of vertices. Note that the potential V depends only on interface gradients and not on the absolute height itself: this reflects the vertical translation invariance of the problem (apart from boundary condition effects). A minimal requirement on V is that it diverges to $+\infty$ when $|h_x - h_y| \to \infty$: the potential has the effect of "flattening" the interface and suppressing wild fluctuations, in agreement with the observed macroscopic flatness of phase boundaries. (Much more stringent conditions have to be imposed on V to actually prove any result.) Note also that the measure π_{Λ_ϵ} depends on the boundary height $\bar{h}_\cdot$: if $\bar{h}_\cdot$ is fixed so that the average slope is $\rho \in \mathbb{R}^d$, i.e. $\pi_{\Lambda_\epsilon}(h_x - h_{x+e_i}) = \rho_i, i \le d$, then we write $\pi_{\Lambda_\epsilon,\rho}$.

There are various choices of Markov dynamics that admit (3-1) as stationary reversible measure. A popular choice is the heat-bath or Glauber dynamics: with rate 1, independently, each height $h_x(t)$ is refreshed and the new value is chosen from the stationary measure π_{Λ_ϵ} *conditioned* on the values of $h_y(t)$ with y ranging over the nearest neighbors of x. Another natural choice, when the heights are in $\mathbb{R}$ rather than in $\mathbb{Z}$, is a Langevin-type dynamics where each $h_x(t)$ is subject to an independent Brownian noise, plus a drift that depends on the height differences between x and its neighboring sites, chosen so that (3-1) is reversible.

As we mentioned, under reasonable assumptions, a diffusive hydrodynamic limit is expected:

$$\epsilon h_{\epsilon^{-1}x}(\epsilon^{-2}t) \xrightarrow[\epsilon \to 0]{\mathbb{P}} \phi(x,t) \tag{3-2}$$

where ϕ is deterministic. Due to the diffusive scaling of time, the PDE solved by ϕ will be of second order and in general non-linear:

$$\partial_t \phi(x,t) = \mu(\nabla\phi(x,t)) \sum_{i,j=1}^{d} \sigma_{i,j}(\nabla\phi(x,t)) \frac{\partial^2}{\partial_{x_i}\partial_{x_j}} \phi(x,t). \tag{3-3}$$

The factors μ and $\sigma_{i,j}$ have a very different origin, which is why we have not written the equation in terms of the combination $\tilde{\sigma}_{i,j} := \mu\sigma_{i,j}$ instead. The slope-dependent prefactor $\mu > 0$ is called *mobility* and will be discussed in a moment. As for $\sigma_{i,j}$, let the convex function $\sigma : \rho \in \mathbb{R}^d \mapsto \sigma(\rho) \in \mathbb{R}$ denote the surface tension of the model at slope ρ Funaki [2005], i.e. minus the limit as $\epsilon \to 0$ of $1/|\Lambda_\epsilon|$ times the logarithm of the normalization constant of the probability measure (3-1) when $\pi_{\Lambda_\epsilon} = \pi_{\Lambda_\epsilon,\rho}$. Then, $\sigma_{i,j}$ denotes the second derivative of σ w.r.t. the i^{th} and j^{th} argument. Convexity of σ implies that the matrix $\{\sigma_{i,j}(\nabla\phi)\}_{i,j=1,\dots,d}$ is positive definite, so the PDE (3-3) is of parabolic type. We emphasize that the surface tension, hence $\sigma_{i,j}$, are defined purely in terms of the stationary measure (3-1). All the dependence on the Markov dynamics is in the mobility μ. Remark also that one can rewrite (3-3) in the following more evocative form:

$$\partial_t\phi(x,t) = -\mu(\nabla\phi(x,t))\frac{\delta F[\phi(\cdot,t)]}{\delta\phi(x,t)} \tag{3-4}$$

where $F[\phi(\cdot)] = \int dx\,\sigma(\nabla\phi)$ is the surface tension functional and $\delta F/\delta\phi$ denotes its first variation. In other words, the hydrodynamic equation is nothing but the gradient flow w.r.t. the surface tension functional, modulated by a slope-dependent mobility prefactor.

Via linear response theory one can guess a Green-Kubo-type expression for the mobility Spohn [1993]. This turns out to be given as[6] (say that the heights h_x are discrete, so that the dynamics is a Markov jump process; the formula is analogous for Langevin-type dynamics)

$$\mu(\rho) = \lim_{\epsilon\to 0}\frac{1}{2|\Lambda_\epsilon|}\pi_{\Lambda_\epsilon,\rho}\left[\sum_{x\in\Lambda_\epsilon}\sum_n c_x^n(h)n^2\right] \tag{3-5}$$

$$-\int_0^\infty dt\,\lim_{\epsilon\to 0}\frac{1}{|\Lambda_\epsilon|}\sum_{x,x'\in\Lambda_\epsilon}\sum_{n,n'}\mathbb{E}_{\pi_{\Lambda_\epsilon,\rho}}\left[c_x^n(h(0))n\; c_{x'}^{n'}(h(t))n'\right] \tag{3-6}$$

where $c_x^n(h)$ is the rate at which the height at x increases by $n\in\mathbb{Z}$ in configuration h, $\mathbb{E}_{\pi_{\Lambda_\epsilon,\rho}}$ denotes expectation w.r.t. the stationary process started from $\pi_{\Lambda_\epsilon,\rho}$ and $h(t)$ denotes the configuration at time t. Note that the first term involves only equilibrium correlation functions in the infinite volume stationary measure $\pi_\rho = \lim_{\Lambda\to\mathbb{Z}^d}\pi_{\Lambda_\epsilon,\rho}$[7]. The same is not true for the second one, which involves a time integral of correlations at different times for the stationary process. These are usually not explicitly computable even when π_ρ is known. It may however happen for certain models that, by a discrete summation by parts w.r.t. the x variable, $\sum_x\sum_n c_x^n(h)$ is deterministically zero, for any configuration h: one says then that a *gradient condition* is satisfied (a classical example is symmetric simple exclusion). In this case (3-6) identically vanishes and one is in a much better position to prove convergence to the hydrodynamic equation.

[6] One can express μ also via a variational principle, see Spohn [2012].

[7] For models in dimension $d\le 2$ the law of the interface does not have a limit as $\epsilon\to 0$, since the variance of h_x diverges as $\epsilon\to 0$. However, the law of the *gradients* of h does have a limit and the transition rates $c_x^n(h)$ are actually functions of the gradients of h only, by translation invariance in the vertical direction.

For the "Ginzburg-Landau (GL)" model Spohn [ibid.] where heights h_x are continuous variables and the dynamics is of Langevin type, if the potential $V(\cdot)$ is convex and symmetric then, in any dimension d, the gradient condition is satisfied and moreover the remaining average in the Green-Kubo formula is immediately computed, leading to a constant mobility: $\mu(\rho) = 1$. In this situation, Funaki and Spohn [1997] proved convergence of the height profile to (the weak solution of) (3-3) for the GL model, for any $d \geq 1$. (They look at weak solutions because for $d > 1$ it is not known whether the surface tension of the GL model is C^2 and the coefficients $\sigma_{i,j}$ are well defined and smooth). Until recently, to my knowledge, there was no other known interface model in dimension $d > 1$ where mathematical results of this type were available.

Before presenting our recent results for $(2+1)$-dimensional interface dynamics let us make two important observations:

- Not only for most natural interface dynamics in dimension $d > 1$ one is unable to prove a hydrodynamic convergence of the type (3-2): the situation is actually much worse. As (3-2) suggests, the correct time-scale for the system to reach stationarity (measured either by $T_{rel} := 1/\text{gap}(\mathfrak{L})$, with $\text{gap}(\mathfrak{L})$ denoting the spectral gap of the generator, or by the so-called total variation mixing time T_{mix}) should be of order ϵ^{-2} (logarithmic corrections are to be expected for the mixing time). On the other hand, for most natural models it not even proven that such characteristic times are upper bounded by a polynomial of ϵ^{-1}! For instance, for the well-known $(2+1)$-dimensional SOS model at low temperature, the best known upper bound for T_{rel} and T_{mix} is a rather poor $O(\exp(\epsilon^{-1/2+o(1)}))$ Caputo, Lubetzky, Martinelli, Sly, and Toninelli [2014, Th. 3].

- In dimension $d = 1$, natural Markov dynamics of discrete interfaces are provided by conservative lattice gases on $\mathbb{Z}$ (e.g. symmetric exclusion processes or zero-range processes), just by interpreting the number of particles at site x as the interface gradient $h_x - h_{x-1}$ at x. Similarly, conservative continuous spin models on $\mathbb{Z}$ translate into Markov dynamics for one-dimensional interface models with continuous heights. Then, a hydrodynamic limit for the height function follows from that for the particle density (see e.g. Kipnis and Landim [1999, Ch. 4 and 5] for the symmetric simple exclusion and for a class of zero-range processes, and for instance Fritz [1989] for the $d = 1$ Ginzburg-Landau model). For $d > 1$, instead, there is in general no obvious way of associating a height function to a particle system on $\mathbb{Z}^d$. Also, for $d = 1$ there are robust methods to prove that the inverse spectral gap is $T_{rel} = O(\epsilon^{-2})$, see e.g. Kipnis and Landim [1999] and Caputo [2004].

3.1 Reversible tiling dynamics, mixing time and hydrodynamic equation. In this section, I briefly review a series of results obtained in recent years in collaboration with

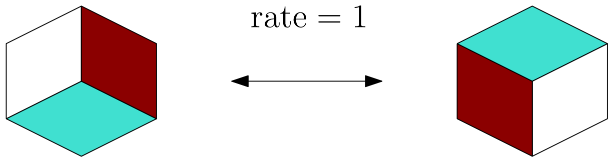


Figure 7: The elementary updates for the Glauber dynamics of the lozenge tiling model correspond to the rotation of three lozenges (equivalently, three dimers) around a hexagonal face. The transition rate is 1 both for the update and the reverse one. Note that in the three-dimensional corner growth model only the forward transition would be allowed.

Pietro Caputo, Benoît Laslier and Fabio Martinelli. In these works we study $(2+1)$-dimensional interface dynamics where the height function $\{h_x\}_{x\in\Lambda_\epsilon}$ is discrete and is given by the height function of a tiling model, either by lozenges or by dominoes, as explained in Section 2.2. In contrast with the $(2+1)$-dimensional Anisotropic KPZ growth models described in Section 2.2, that are also Markov dynamics of tiling models, here we want a *reversible* process because we wish to model interface evolution at thermal equilibrium. A natural candidate is the "Glauber" dynamics obtained by giving rate 1 to the elementary rotations of tiles around faces of the graph, see Figure 7 for the case of lozenge tilings.

In terms of the height function, elementary moves correspond to changing the height by ± 1 at single sites. Since all elementary rotations have the same rate, the uniform measure over the finitely many tiling configurations in Λ_ϵ is reversible. As a side remark, this measure can be written in the Boltzmann-Gibbs form (3-1) with a potential V taking values 0 or $+\infty$. Let us also remark that, as discussed in Caputo, Martinelli, and Toninelli [2012], this dynamics is equivalent to the zero-temperature Glauber dynamics of the *three-dimensional* Ising model with Dobrushin boundary conditions.

In agreement with the discussion of the previous section, if the tiled region is a reasonably-shaped domain Λ_ϵ of diameter $O(\epsilon^{-1})$, one expects T_{rel} and T_{mix} to be $\approx \epsilon^{-2}$ and the height profile to converge under diffusive rescaling to the solution of a parabolic PDE. Until recently, however, all what was known rigorously was that T_{rel} and T_{mix} are upper bounded as $O(\epsilon^{-n})$ for some finite $n > 2$!

Open problem 3.1. *This polynomial upper bound was proven in Luby, Randall, and Sinclair [2001] for the Glauber dynamics on either lozenge or domino tiling (the same proof works for tilings associated to the dimer model on certain graphs with both hexagonal and square faces, as shown in Laslier and Toninelli [2015a]). The method does not seem to work, however, for general planar bipartite graphs. For instance, a polynomial upper* *bound for T_{rel} or T_{mix} of the Glauber dynamics of the dimer model on the square-octagon graph (see Fig. 9 in Kenyon [2009]) is still unproven.*

Under suitable conditions, we improved this $O(\epsilon^{-n})$ upper bound into an almost optimal one:

Theorem 3.2 (Informal statement). *If the boundary height on $\partial\Lambda_\epsilon$ is such that the average height under the measure π_{Λ_ϵ} tends to an affine function as $\epsilon \to 0$, then T_{mix} and T_{rel} are $O(\epsilon^{-2+o(1)})$ Caputo, Martinelli, and Toninelli [2012] and Laslier and Toninelli [2015a].*

Later Laslier and Toninelli [2015b], we proved a result in the same spirit under the sole assumption that the limit average height profile is smooth and in particular has no "frozen regions" Kenyon [2009].

Let us emphasize that there are very natural domains Λ_ϵ such that the average equilibrium height profile in the $\epsilon \to 0$ limit does have "frozen regions":

Open problem 3.3. *Let Λ_ϵ be the hexagonal domain of side ϵ^{-1}, see Figure 1. Prove that, for the lozenge tiling Glauber dynamics, T_{rel} and/or T_{mix} are $O(\epsilon^{-2+o(1)})$. The best upper bound that can be extracted from Wilson [2004] plus the so-called Peres-Winkler censoring inequalities Peres and Winkler [2013] is $O(\epsilon^{-4+o(1)})$.*

The proofs of the previously known polynomial upper bounds on the mixing time were based on smart and rather simple path coupling arguments Luby, Randall, and Sinclair [2001]. To get our almost optimal bounds Caputo, Martinelli, and Toninelli [2012] and Laslier and Toninelli [2015a,b], there are at least two new inputs:

- our proof consists in a comparison between the actual interface dynamics and an auxiliary one that evolves on almost-diffusive time-scales $\approx \epsilon^{-2+o(1)}$ and that essentially follows the conjectural hydrodynamic motion where interface drift is proportional to its curvature;

- to control the auxiliary process, we crucially need very refined estimates on height fluctuation for the equilibrium measure π_{Λ_ℓ} on domains of mesoscopic size $\ell = \epsilon^{-a}, 1/2 \le a \le 1$, with various types of boundary conditions.

For the Glauber dynamics with elementary moves as in Figure 7, it seems hopeless to prove a hydrodynamic limit on the diffusive scale. In particular, no form of "gradient condition" is satisfied. Fortunately, there exists a more friendly variant of the Glauber dynamics, introduced in Luby, Randall, and Sinclair [2001], where a single update consists in "tower moves" changing the height by the same amount ± 1 at $n \ge 0$ aligned sites, as in Figure 8. The integer n is not fixed here, in fact transitions with any n are allowed but the transition rate decreases with n and actually it is taken to equal to $1/n$. It is immediate to

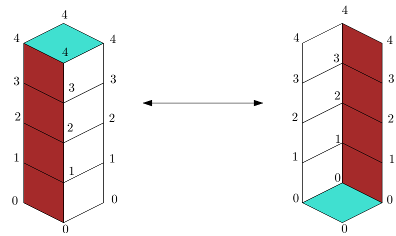


Figure 8: A "tower move" transition with $n = 4$ and the reverse transition. The transition rates equal $1/n = 1/4$. Note that the height decreases by 1 at four points.

verify that this dynamics is still reversible w.r.t. the uniform measure. For this modified dynamics, together with B. Laslier we realized in Laslier and Toninelli [2017] that a microscopic summation by parts implies that the term (3-6) in the definition of the mobility vanishes, and actually we could explicitly compute μ, that turns out to be *non-trivial and non-linear*:

$$\mu(\rho) = \frac{1}{\pi} \frac{\sin(\pi\rho_1)\sin(\pi\rho_2)}{\sin(\pi(\rho_1+\rho_2))}. \tag{3-7}$$

Recall that, in contrast, the mobility of the Ginzburg-Landau model is independent of the slope Spohn [1993]. Later, in Laslier and Toninelli [2015c, Th. 2.7], we could turn our arguments into a full proof of convergence of the height profile to the solution of the PDE:

Theorem 3.4 (Informal statement)**.** *If the initial profile ϕ_0 is sufficiently smooth, one has for every $t > 0$*

$$\lim_{\epsilon\to 0} \frac{1}{|\Lambda_\epsilon|} \sum_{x\in\Lambda_\epsilon} \mathbb{E}\left|\epsilon h_x(\epsilon^{-2}t) - \phi(\epsilon x, t)\right|^2 = 0, \tag{3-8}$$

with $\phi(x,t)$ the solution of (3-3)

(For technical reasons, we had to work with periodic instead of Dirichlet boundary conditions). A couple of comments are in order:

- As the reader may have noticed, the function (3-7) is exactly the same as the "speed function" $v(\rho)$ of the growth model discussed in Section 2.2, see formula (2-7). This is not a mere coincidence. Actually, one may see this equality as an instance of the so-called Einstein relation between diffusion and conductivity coefficients Spohn [2012].

- We mentioned earlier that convergence of the height profile of the Ginzburg-Landau model to the limit PDE has been proved Funaki and Spohn [1997] only in a weak sense. In our case, instead, we have strong convergence to classical solutions of (3-3) that exist globally because the coefficients $\mu(\cdot), \sigma_{i,j}(\cdot)$ turn out to be smooth functions of the slope.[8]

- A fact that plays a crucial role in the proof of the hydrodynamic limit is that the PDE (3-3) contracts the L^2 distance $D_2(t) = \int dx(\phi^{(1)}(x,t) - \phi^{(2)}(x,t))^2$ between solutions. I believe this is not a trivial or general fact: in fact, to prove contraction Laslier and Toninelli [2015c], we use the specific form (3-7) of μ and the explicit expression of $\sigma_{i,j}$ for the dimer model. (Note that if the mobility were constant, as it is for the Ginzburg-Landau model, L^2 contraction would just be a consequence of convexity of the surface tension σ). I think it is an intriguing question to understand whether the identities (see Laslier and Toninelli [ibid., Eqs. (6.19)-(6.22)]) leading to $dD_2(t)/dt \le 0$ have any thermodynamic interpretation.

To conclude this review, let us mention that new dynamical phenomena, taking place on time-scales much longer than diffusive, can occur at low temperature, for interface models undergoing a so-called "roughening transition". That is, up to now we considered situations where the equilibrium Gibbs measure for the interface in a $\epsilon^{-1} \times \epsilon^{-1}$ box Λ_ϵ scales to a massless Gaussian field as $\epsilon \to 0$ and in particular $\mathrm{Var}_{\pi_{\Lambda_\epsilon}}(h_x) \approx \log(1/\epsilon)$ if x is, say, the center of the box. The interface is said to be "rough" in this case, because fluctuations diverge as $\epsilon \to 0$. For some interface models, notably the well-known Solid-on-Solid (SOS) model where the potential V in (3-1) equals $T^{-1}|h_x - h_y|$ and heights are integer-valued and fixed to 0 around the boundary, it is known that at low enough temperature T the interface is instead rigid, with $\limsup_{\epsilon\to 0} \mathrm{Var}_{\pi_{\Lambda_\epsilon}}(h_x) < \infty$, while the variance grows logarithmically at high temperature Fröhlich and Spencer [1981]. The temperature T_r separating these two regimes is called "roughening temperature".

In a work with Caputo, Lubetzky, Martinelli, Sly, and Toninelli [2014] we discovered that rigidity of the interface can produce a dramatic slowdown of the dynamics, if the interface is constrained to stay above a fixed level, say level 0:

Theorem 3.5. *Caputo, Lubetzky, Martinelli, Sly, and Toninelli [ibid.] Consider the Glauber dynamics for the $(2+1)$-dimensional SOS model at low enough temperature, with 0 boundary conditions on $\partial\Lambda_\epsilon$ and with the positivity constraint $h_x \ge 0$ for every $x \in \Lambda_\epsilon$. Then, the relaxation and mixing times satisfy*

$$T_{mix} \ge T_{rel} \ge c \exp[c\, \epsilon^{-1}] \tag{3-9}$$

[8] The apparent singularity of the formula (3-7) for $\mu(\cdot)$ when $\rho_1 + \rho_2 = 0$ is not really dangerous: recall from Section 2.2 that the slope ρ is constrained in the triangle $\mathbb{T} = \{(\rho_1, \rho_2) : 0 \le \rho_1, \rho_2, \rho_1 + \rho_2 \le 1\}$ so that the mobility is C^∞ and strictly positive in the interior of $\mathbb{T}$.

for some positive, temperature-dependent constant c.

What we actually prove is that there is a cascade of metastable transitions, occurring on all time-scales $\exp(\epsilon^{-a})$, $a < 0 \le 1$. Strange as this may look, these results *do not* exclude that a hydrodynamic limit on the *diffusive scale*, as in (3-2)-(3-3), might occur. That is, the rescaled height profile $\epsilon h_{\epsilon^{-1}x}(\epsilon^{-2}t)$ could follow an equation like (3-3), so that at times $\gg \epsilon^{-2}$ the profile would be macroscopically zero (because $\phi \equiv 0$ is the equilibrium point of the PDE (3-3) with zero boundary conditions) but smaller-scale height fluctuations would need enormously more time, of the order $T_{mix} \approx \exp(\epsilon^{-1})$, to relax to equilibrium.

We are light years away from being able to actually prove a hydrodynamic limit for the $(2+1)$-dimensional SOS model. The following open problem is given just to show how little we know in this respect:

Open problem 3.6. *Take the Glauber dynamics for the $(2+1)$-d SOS model at low temperature, with initial condition $\epsilon\, h_x = 1$ for every $x \in \Lambda_\epsilon$. Is it true that, for some $N < \infty$, at time $t = 1/\epsilon^N$ all rescaled heights $\epsilon\, h_x(t)$ are with high probability lower than, say, $1/2$ (which is much larger than $\epsilon \log \epsilon^{-1}$, that is the typical value under the equilibrium measure of $\max_{x\in\Lambda_\epsilon}[\epsilon\, h_x]$)?*

References

Albert-László Barabási and H. Eugene Stanley (1995). *Fractal concepts in surface growth*. Cambridge University Press, Cambridge, pp. xx+366. MR: 1600794 (cit. on pp. 2752, 2756, 2757).

Alexei Borodin, Alexey Bufetov, and Grigori Olshanski (2015). "Limit shapes for growing extreme characters of $U(\infty)$". *Ann. Appl. Probab.* 25.4, pp. 2339–2381. MR: 3349009 (cit. on p. 2767).

Alexei Borodin, Ivan Corwin, and Fabio Lucio Toninelli (2017). "Stochastic heat equation limit of a $(2+1)$d growth model". *Comm. Math. Phys.* 350.3, pp. 957–984. MR: 3607467 (cit. on p. 2768).

Alexei Borodin and Patrik L. Ferrari (2014). "Anisotropic growth of random surfaces in $2+1$ dimensions". *Comm. Math. Phys.* 325.2, pp. 603–684. MR: 3148098 (cit. on pp. 2763–2765, 2768).

Pietro Caputo (2004). "Spectral gap inequalities in product spaces with conservation laws". In: *Stochastic analysis on large scale interacting systems*. Vol. 39. Adv. Stud. Pure Math. Math. Soc. Japan, Tokyo, pp. 53–88. MR: 2073330 (cit. on p. 2771).

Pietro Caputo, Eyal Lubetzky, Fabio Martinelli, Allan Sly, and Fabio Lucio Toninelli (2014). "Dynamics of $(2 + 1)$-dimensional SOS surfaces above a wall: slow mixing induced by entropic repulsion". *Ann. Probab.* 42.4, pp. 1516–1589. MR: 3262485 (cit. on pp. 2771, 2775).

Pietro Caputo, Fabio Martinelli, and Fabio Lucio Toninelli (2012). "Mixing times of monotone surfaces and SOS interfaces: a mean curvature approach". *Comm. Math. Phys.* 311.1, pp. 157–189. MR: 2892467 (cit. on pp. 2772, 2773).

Sunil Chhita and Patrik Ferrari (2017). "A combinatorial identity for the speed of growth in an anisotropic KPZ model". *Ann. Inst. Henri Poincaré D* 4.4, pp. 453–477. MR: 3734414 (cit. on p. 2765).

Sunil Chhita, Patrik L. Ferrari, and Fabio Lucio Toninelli (May 2017). "Speed and fluctuations for some driven dimer models". To appear in Ann. Inst. Henri Poincaré D. arXiv: 1705.07641 (cit. on pp. 2765–2767).

Sunil Chhita and Fabio Lucio Toninelli (2018). "A (2+ 1)-dimensional Anisotropic KPZ growth model with a rigid phase". *arXiv preprint arXiv:1802.05493* (cit. on p. 2768).

Henry Cohn, Michael Larsen, and James Propp (1998). "The shape of a typical boxed plane partition". *New York J. Math.* 4, pp. 137–165. MR: 1641839 (cit. on p. 2754).

Ivan Corwin (2012). "The Kardar-Parisi-Zhang equation and universality class". *Random Matrices Theory Appl.* 1.1, pp. 1130001, 76. MR: 2930377 (cit. on pp. 2753, 2757).

Patrik L Ferrari (2008). "Slow decorrelations in Kardar–Parisi–Zhang growth". *Journal of Statistical Mechanics: Theory and Experiment* 2008.07, P07022 (cit. on p. 2768).

J. Fritz (1989). "On the hydrodynamic limit of a Ginzburg-Landau lattice model. The law of large numbers in arbitrary dimensions". *Probab. Theory Related Fields* 81.2, pp. 291–318. MR: 982659 (cit. on p. 2771).

Jürg Fröhlich and Thomas Spencer (1981). "The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas". *Comm. Math. Phys.* 81.4, pp. 527–602. MR: 634447 (cit. on p. 2775).

T. Funaki and H. Spohn (1997). "Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model". *Comm. Math. Phys.* 185.1, pp. 1–36. MR: 1463032 (cit. on pp. 2771, 2775).

Tadahisa Funaki (2005). "Stochastic interface models". In: *Lectures on probability theory and statistics*. Vol. 1869. Lecture Notes in Math. Springer, Berlin, pp. 103–274. MR: 2228384 (cit. on pp. 2768, 2769).

Yu Gu, Lenya Ryzhik, and Ofer Zeitouni (Oct. 2017). "The Edwards-Wilkinson limit of the random heat equation in dimensions three and higher". arXiv: 1710.00344 (cit. on p. 2757).

M. Hairer (2014). "A theory of regularity structures". *Invent. Math.* 198.2, pp. 269–504. MR: 3274562 (cit. on p. 2756).

Timothy Halpin-Healy and Amine Assdah (1992). “On the kinetic roughening of vicinal surfaces”. *Physical Review A* 46.6, p. 3527 (cit. on p. 2760).

Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang (1986). “Dynamic scaling of growing interfaces”. *Physical Review Letters* 56.9, p. 889 (cit. on pp. 2756–2759).

Pieter W Kasteleyn (1961). “The statistics of dimers on a lattice: I. The number of dimer arrangements on a quadratic lattice”. *Physica* 27.12, pp. 1209–1225 (cit. on p. 2762).

Richard Kenyon (2009). “Lectures on dimers”. In: *Statistical mechanics*. Vol. 16. IAS Park City Math. Ser. Amer. Math. Soc., Providence, RI, pp. 191–230. MR: 2523460 (cit. on pp. 2753, 2761, 2766, 2773).

Richard Kenyon, Andrei Okounkov, and Scott Sheffield (2006). “Dimers and amoebae”. *Ann. of Math. (2)* 163.3, pp. 1019–1056. MR: 2215138 (cit. on p. 2762).

Claude Kipnis and Claudio Landim (1999). *Scaling limits of interacting particle systems*. Vol. 320. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, pp. xvi+442. MR: 1707314 (cit. on p. 2771).

Benoît Laslier and Fabio Lucio Toninelli (2015a). “How quickly can we sample a uniform domino tiling of the $2L \times 2L$ square via Glauber dynamics?” *Probab. Theory Related Fields* 161.3-4, pp. 509–559. MR: 3334275 (cit. on pp. 2772, 2773).

– (2015b). “Lozenge tilings, Glauber dynamics and macroscopic shape”. *Comm. Math. Phys.* 338.3, pp. 1287–1326. MR: 3355815 (cit. on p. 2773).

– (2015c). “Lozenge tilings, Glauber dynamics and macroscopic shape”. *Comm. Math. Phys.* 338.3, pp. 1287–1326. arXiv: 1701.05100. MR: 3355815 (cit. on pp. 2774, 2775).

– (2017). “Hydrodynamic limit equation for a lozenge tiling Glauber dynamics”. *Ann. Henri Poincaré* 18.6, pp. 2007–2043. MR: 3649450 (cit. on p. 2774).

Martin Legras and Fabio Lucio Toninelli (Apr. 2017). “Hydrodynamic limit and viscosity solutions for a 2D growth process in the anisotropic KPZ class”. To appear in Comm. Pure Appl. Math. arXiv: 1704.06581 (cit. on pp. 2765, 2766).

Michael Luby, Dana Randall, and Alistair Sinclair (2001). “Markov chain algorithms for planar lattice structures”. *SIAM J. Comput.* 31.1, pp. 167–192. MR: 1857394 (cit. on pp. 2772, 2773).

Jacques Magnen and Jérémie Unterberger (Feb. 2017). “The scaling limit of the KPZ equation in space dimension 3 and higher”. To appear in Journ. Stat. Phys. arXiv: 1702.03122 (cit. on p. 2757).

Yuval Peres and Peter Winkler (2013). “Can extra updates delay mixing?” *Comm. Math. Phys.* 323.3, pp. 1007–1016. MR: 3106501 (cit. on p. 2773).

M. Prähofer and H. Spohn (1997). “An exactly solved model of three-dimensional surface growth in the anisotropic KPZ regime”. *J. Statist. Phys.* 88.5-6, pp. 999–1012. MR: 1478060 (cit. on p. 2761).

Michael Prähofer (2003). “Stochastic surface growth”. PhD thesis. München (cit. on p. 2759).

Jeremy Quastel (2012). “Introduction to KPZ”. In: *Current developments in mathematics, 2011*. Int. Press, Somerville, MA, pp. 125–194. MR: 3098078 (cit. on pp. 2753, 2757).

Fraydoun Rezakhanlou (2001). “Continuum limit for some growth models. II”. *Ann. Probab.* 29.3, pp. 1329–1372. MR: 1872745 (cit. on p. 2766).

– (2002). “Continuum limit for some growth models”. *Stochastic Process. Appl.* 101.1, pp. 1–41. MR: 1921440 (cit. on p. 2760).

Timo Seppäläinen (1996). “A microscopic model for the Burgers equation and longest increasing subsequences”. *Electron. J. Probab.* 1 (5), p. 51. MR: 1386297 (cit. on p. 2764).

– (2000). “Strong law of large numbers for the interface in ballistic deposition”. *Ann. Inst. H. Poincaré Probab. Statist.* 36.6, pp. 691–736. MR: 1797390 (cit. on p. 2760).

Herbert Spohn (1993). “Interface motion in models with stochastic dynamics”. *J. Statist. Phys.* 71.5-6, pp. 1081–1132. MR: 1226387 (cit. on pp. 2768, 2770, 2771, 2774).

– (2012). *Large scale dynamics of interacting particles*. Springer Science & Business Media (cit. on pp. 2770, 2774).

Lei-Han Tang, Bruce M Forrest, and Dietrich E Wolf (1992). “Kinetic surface roughening. II. Hypercube-stacking models”. *Physical Review A* 45.10, pp. 7162–7169 (cit. on p. 2760).

Fabio Lucio Toninelli (2017). “A $(2+1)$-dimensional growth process with explicit stationary measures”. *Ann. Probab.* 45.5, pp. 2899–2940. MR: 3706735 (cit. on pp. 2764–2766, 2768).

David Bruce Wilson (2004). “Mixing times of Lozenge tiling and card shuffling Markov chains”. *Ann. Appl. Probab.* 14.1, pp. 274–325. MR: 2023023 (cit. on p. 2773).

Dietrich E Wolf (1991). “Kinetic roughening of vicinal surfaces”. *Physical review letters* 67.13, p. 1783 (cit. on pp. 2756, 2758, 2759).

Received 2017-11-15.

Fabio Toninelli
CNRS and Institut Camille Jordan
Université Lyon 1
43 bd du 11 novembre 1918
69622 Villeurbanne
France
toninelli@math.univ-lyon1.fr

Index of Authors

World Scientific
www.worldscientific.com
11060 hc

ISBN 978-981-3272-87-3 (set)
9 789813 272873

ISBN 978-981-3272-92-7 (v3)
9 789813 272927