PROCEEDINGS OF THE INTERNATIONAL CONGRESS OF MATHEMATICIANS

Volume 2

The Proceedings of the International Congress of Mathematicians held in Vancouver, August 21—29, 1974, is in two volumes. Volume 1 contains an account of the organization of the Congress, the list of members, the work of the Fields medalists, the expository addresses, and the addresses in Sections 1—7. Volume 2 contains the addresses in Sections 8—20 and the list of short communications.

Canadian Shared Cataloguing in Publication Data

International Congress of Mathematicians,
Vancouver, B.C., 1974.
Proceedings of the International Congress
of Mathematicians / [editor, Ralph D. James].—
1. Mathematics—Congresses. I. James, Ralph Duncan, ed.
QA1.I8 1974 510.631
ISBN 0-919558-04-6

Library of Congress Catalog Card No. 74-34533

Printed in the United States of America

Contents

Section 8—Differential Geometry and Analysis on Manifolds

Section 9—General Topology, Real and Functional Analysis

Section 10—Operator Algebras, Harmonic Analysis and Representation of Groups

Section 11—Probability and Mathematical Statistics, Potential, Measure and Integration

Section 12—Complex Analysis

Section 13—Partial Differential Equations

Section 14—Ordinary Differential Equations and Dynamic Systems

Section 15—Control Theory and Related Optimization Problems

Section 16—Mathematical Physics and Mechanics

Section 17—Numerical Mathematics

Section 18—Discrete Mathematics and Theory of Computation

Section 19—Applied Statistics, Mathematics in the Social and Biological Sciences

Section 20—History and Education

Section 8

Differential Geometry and Analysis on Manifolds

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Invariants of Flat Bundles

Jeff Cheeger

Let G be a Lie group with finitely many components. We are going to discuss some cohomology invariants of flat principal G-bundles. Their properties were developed jointly with Simons [2] in an outgrowth of earlier work of Chern and Simons [3]; see also [4] for related ideas.

We begin with some notation. Let $I^*(G)$ denote the ring of invariant polynomials of G and w: $I^k(G) \to H^{2k}(B_G, R)$ the universal Weil homomorphism. Fix $\Lambda \subset R$, a proper subring and define

$$K^k(G, \Lambda) = \{(P, u) \in I^k(G) \times H^{2k}(B_G, \Lambda) \mid w(P) = r(u)\}.$$

Here $r(u) \in H^{2k}(B_G, R)$ is the real image of u. Let G° represent G equipped with the discrete topology. $B_{G^\circ} = K(G^\circ, 1)$ is the classifying space for flat principal G-bundles. Let b: $H^{2k-1}(\ , R/\Lambda) \to H^{2k}(\ , \Lambda)$ denote the Bockstein map of the coefficient sequence $0 \to \Lambda \to R \to R/\Lambda \to 0$.

THEOREM 1. *There is a homomorphism S: $K^k(G, \Lambda) \to H^{2k-1}(B_{G^\circ}, R/\Lambda)$ such that the following diagram commutes*

$$\begin{array}{ccccc} & & K^k(G, \Lambda) & & \\ & \swarrow^{S} & & \searrow^{\pi_2} & \\ H^{2k-1}(B_{G^\circ}, R/\Lambda) & \xrightarrow{b} & H^{2k}(B_{G^\circ}, \Lambda) & \xleftarrow{i^*} & H^{2k}(B_G, \Lambda) \end{array}$$

Equivalently, for each flat bundle α over a space M there exists $S_{P,u}(\alpha) \in H^{2k-1}(M, R/\Lambda)$ which is natural under bundle maps such that $b(S_{P,u}(\alpha)) = -u(\alpha) \in H^{2k-1}(M)$.

The $S_{P,u}$ are secondary characteristic classes in the sense that unless $P \equiv 0$, they are not the pull-back of classes in B_G.

For the details of their construction and its extension to the more general context of nonflat bundles with connection we refer to [3].

$K^*(G, \Lambda) = \bigoplus K^k(G, \Lambda)$ forms a graded ring with multiplication $(P_1, U_1) \cdot (P_2, U_2) = (P_1 \cdot P_2, U_1 \cup U_2)$. $H^*(R/\Lambda)$ is also a graded ring with respect to the $*$-product $f_1 * f_2 = (-1)^k (b * f_1) \cup f_2$ where $f_i \in H^{r_i}(\ , R/\Lambda)$. Note that if $\Lambda = Q$, the rationals, then $b \equiv 0$ and the $*$-product is trivial. Also it may be shown that S is a ring homomorphism.

In case the holonomy group H of α is finite then, as J. Millson pointed out, the $S_{P,u}$ have an explicit interpretation. Let $\rho : B_H \to B_G$. Since H is finite, $H^i(B_H, R) = 0$ for $i \geqq 1$ and $b : H^{2i-1}(B_H, R/\Lambda) \to H^{2i}(B_H, \Lambda)$ is an isomorphism. Then necessarily we must have $H^{2i-1}(B_H, R/\Lambda) \ni S_{P,u} = -b(\rho^*(u))$. In case H is infinite, it is no longer true that $H^i(B_H, R) = 0$, but the formula may still be generalized by introducing so-called Borel cohomology.

It is also possible to construct the $S_{P,u}$ intrinsically:

EXAMPLE 2. Let $G = \mathrm{SO}(2n)$, χ be the Euler class and P_χ the Pfaffian. Let α be a flat G-bundle and $S(E)$ the associated sphere bundle. Then we have the homology sequence

$$0 \to H^{2n-1}(S^{2n-1}) \to H^{2n}(S(E)) \to H^{2n}(M) \to 0.$$
$$\wr \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$$
$$\boldsymbol{Z}$$

Let Vol denote the volume form on the fibre normalized so that $\mathrm{Vol}(S^{2n-1}) = 1$. Since $S(E)$ is flat, Vol extends uniquely to a closed form ω on $S(E)$. Given $x \in H_{2n-1}(M, \boldsymbol{Z})$ choose $z \in Z_{2n-1}(E(M))$ such that $[\pi(z)] = u$. Set $S_{P,\chi}(x) = \hat{\chi}(x) = (\omega(z))^{\sim}$ where $^{\sim}$ denotes reduction mod $\boldsymbol{Z}$. It follows from the exact sequence that $\hat{\chi}$ is well defined and one may show that $\hat{\chi} = S_{P,\chi}$.

In analogy with Example 2 it is possible to represent $\hat{\chi}$ by an explicit Borel cocycle in the bar resolution of G°. The formula for this cocycle involves the transcendental function which expresses the volume of a totally geodesic simplex on S^{2n-1} in terms of its dihedral angles. We mention in passing that by means of the power series for this function it is possible to construct a p-adic analogue of $\hat{\chi}$.

With appropriate modifications, the discussion of $\hat{\chi}$ applies equally well to classes $\hat{c}_i$, $\hat{p}_i$ corresponding to Chern and Pontryagin classes of $\mathrm{GL}(n, C)$ and $\mathrm{GL}(n, R)$ bundles. Set $\hat{c} = 1 + \sum_{i>0} \hat{c}_i$. If we regard $\hat{c}$ as being defined for flat vector bundles then the expected relation $\hat{c}(E \otimes F) = \hat{c}(E) * \hat{c}(F)$ holds. Moreover, taking $\Lambda = Q$ we have $(\hat{ch})$ corresponding to the Chern character. Because the $*$-product is trivial mod Q,

$$(\hat{ch}) = n + \sum_{i>0} (-1)^{i-1} \frac{\hat{c}_i}{(i-1)!} \tag{3}$$

and

$$(\hat{ch})(E^{n_1} \otimes E^{n_2}) = n_2(\hat{ch})(E^{n_1}) + n_1(\hat{ch})(E^{n_2}) - n_1 n_2. \tag{4}$$

An important property of the $S_{P,u}$ is their rigidity under deformations in dimension > 1.

THEOREM 5. *Let θ_t be a 1-parameter family of flat connections on the bundle $G \to E \to M$. Let $S^t_{P,u} \in H^{2k-1}(M, R/\Lambda)$ be an invariant corresponding to θ_t. If $k > 1$, $S^t_{P,u}$ is independent of t.*

This result is somewhat mysterious in that the proof makes use of nonflat bundles. Observe that for flat circle bundles over the circle, $\hat{\chi} \in H^1(S^1, R/Z)$ is just given by the angle of holonomy which can take on any value. It is therefore not rigid.

The deepest questions about the $S_{P,u}$ relate to the values which they take on cycles. Let V denote the vector space over Q generated by the values of $S_{P,u} \in H^{2k-1}(B_{G^\delta}, R/Z)$. Although for $k > 1$ the $S_{P,u}$ are not known to take on irrational values, it seems reasonable to conjecture that the dimension of V is countably infinite. In fact if G is algebraic, an argument based on the rigidity theorem shows that the set of values of $S_{P,u}$ is countable for $k > 1$. The dimension of V furnishes a lower bound for the rank of the homology group $H_{2k-1}(B_{G^\delta}, R)$. However, these homology groups can often be shown to be infinite by more indirect arguments. For example,

THEOREM 6. *The groups $H_{2n-1}(\mathrm{SO}(2n, C))$ and $H_{2n-1}(\mathrm{SO}(2n, R))$ and $H_{2n-1}(\mathrm{SO}(2n-1, 1, R))$ all have infinite rank.*

The proof uses what might be called "real noncontinuous" cohomology classes to show that certain cycles constructed from number theory are independent. In fact, the invariant polynomial P_χ on $\mathrm{SO}(2n, R)$ extends by complex linearity to an invariant polynomial on $\mathrm{SO}(2n, C)$. If IP_χ denotes the imaginary part of this polynomial, then $w(IP_\chi) = 0$ and we have $\hat{I}_X = S_{IP_\chi,0} \in H^{2n-1}(B_{\mathrm{SO}(2n,C)}, R)$. Let σ be an automorphism of the algebraic numbers. Then σ may be extended to a (noncontinuous) automorphism of C and hence induces an automorphism of σ^* of $H^{2n-1}(B_{\mathrm{SO}(2n,C)}, R)$. The relevant classes are of the form $\sigma^*(\hat{I}_X)$ for suitable σ.

We should not neglect to mention the important problem of finding a natural and complete set of invariants for deciding when a cycle in $H_*(B_{G^\delta}, \mathbf{Z})$ is homologous to zero.

We will close by describing how the $S_{P,u}$ arise in the context of the "geometric index theorem" of Atiyah, Patodi, Singer [1]. They have defined a spectral invariant $\eta_{E^n}(M^{2k-1}, g)$ of a flat unitary bundle E^n over a riemannian manifold M^{2k-1} with metric g. If E^1 is globally flat, we just write $\eta(M^{2k-1}, g)$. k odd implies $\eta(M^{2k-1}, g) = 0$.

Now assume that $M^{2k-1} = \partial N^{2k}$ and that E^n extends to a possibly nonflat bundle F^n over N^{2k}. Choose a metric $\bar{g}$ on N^{2k} such that on $U = \partial N^{2k} \times [0, \varepsilon] \subset N^{2k}$, $\bar{g}$ is isometrically the product of g and the metric on $[0, \varepsilon)$. Likewise, let $\bar{\nabla}$ be a connection on $F^n|U$ which is the product of the flat connection on E^n and the trivial connection along $[0, \varepsilon]$. Let Ω, Φ denote the curvature forms of $\bar{g}$, $\bar{\nabla}$ and L, ch denote the Hirzebruch and Chern character polynomials. Then, there is an integer $\#$ which is the index of a certain boundary value problem associated to the data such that

$$- \# = \int_{N^n} L(\Omega) \wedge ch(\Phi) + (-1)^k \eta_E(g).$$

We observe

LEMMA 7. $\int_{N^n} L(\Omega) \wedge ch(\Phi) = L(M) \cup (\hat{ch})(E^n)(M) \mod Q$, *so that*

$$(-1)^{k+1} \eta_{E(g)} = L(M) \cup (ch)(E^n)(M) \mod Q.$$

The proof while not standard is quite short. Combining Lemma 7 with (3) and (4) gives

$$(8) \qquad \begin{aligned}(-1)^{k+1} \eta_{E^n}(M, g) &= (-1)^{k+1} n \cdot \eta(M, g) \\ &\quad + \sum_{i \geqq 0}^{k-1} \frac{(-1)^{2(k-i)-1}}{(2k-2i-1)!} L_i(M) \cup \hat{c}_{2(k-i)}(E^n)(M) \mod Q\end{aligned}$$

and

$$(9) \qquad \eta_{E^{n_1} \otimes E^{n_2}}(M, g) = n_2 \eta_{E^{n_1}}(M, g) + n_1 \eta_{E^{n_2}}(M, g) - n_1 n_2 \eta(M, g) \mod Q.$$

Because these formulas only hold mod Q, it seems doubtful that they can be proved directly by purely analytic methods.

Bibliography

1. M. F. Atiyah, V. Patodi and I. M. Singer, *Spectral asymmetry and riemannian geometry* (preprint).

2. J. Cheeger and J. Simons, *Differential characters and geometric invariants* (preprint).

3. S.S. Chern and J. Simons, *Characteristic forms and geometric invariants*, Ann. of Math. (2) **99** (1974), 48–69.

4. F.W. Kamber and P. Tondeur, *Characteristic invariants of foliated bundles* (preprint).

SUNY AT STONY BROOK
STONY BROOK, NEW YORK 11790, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Geometric Aspects of the Generalized Plateau Problem

H. Blaine Lawson, Jr.

Let M be a Riemannian manifold and denote by $\mathscr{R}_p(M)$ the rectifiable currents of dimension p with compact support on M (cf. [3]). Then a generalized Plateau problem on M is the following. Given a current $B \in \mathscr{R}_{p-1}(M)$ such that $B = dV_0$ for some given $V_0 \in \mathscr{R}_p(M)$, find a current of least mass (i.e., "weighted area") in the class $\mathscr{A}_{B,V_0} = \{V \in \mathscr{R}_p(M)\colon dV = B \text{ and } V \text{ is homologous to } V_0\}$. When $B = 0$, this becomes the problem of finding a representation of least mass in a given integral homology class on M. Analogous formulations of the problem exist using homology with $(\mathbf{Z}/m\mathbf{Z})$-coefficients. For $m = 2$, this is called the "unoriented Plateau problem". From the work of Federer and Fleming the existence of solutions to a generalized Plateau problem is guaranteed whenever there is a sequence of currents $\{V_j\}_{j=1}^{\infty} \subset \mathscr{A}_{B,V_0}$ with uniformly bounded supports, whose masses tend to the minimum. In particular, solutions are always guaranteed if M is compact or homogeneous. The question then is to determine the structure of such solutions. This falls into the general problem of determining the structure of stationary (or stable) currents on M. These latter classes are defined by infinitesimal criteria, namely, that the first variation of mass be zero (or that the first variation be zero and the second variation be $\geqq 0$) for smooth deformations which preserve the boundary; and so they lend themselves more directly to the methods of differential geometry.

Certainly it is of fundamental importance to understand the local (i.e., the singularity) structure of these currents, and the first step in approaching this question is to understand the structure of their tangent cones. In particular, this means studying the geometry of compact minimal submanifolds of the Euclidean sphere S^n. In dimension one, such submanifolds are simply great circles. However, matters quickly become more complicated (cf. [9], [10]).

THEOREM 1. *There exist compact orientable surfaces of every genus minimally embedded in S^3. When the genus is not prime, the embeddings are not all congruent. However, two minimally embedded surfaces of the same genus are always ambiently isotopic.*

The cones on these surfaces give examples of stationary 3-currents in Euclidean 4-space with isolated singularities. However, such cones cannot be stable.

THEOREM 2 (ALMGREN, SIMONS [15]). *Let V^{n-2} be a compact minimal submanifold of dimension $n-2$ in S^{n-1} such that the cone $C(V^{n-2})$ over V^{n-2}, with the origin as vertex, is stable in $\boldsymbol{R}^n$. Then if $n < 8$, V^{n-2} is a totally geodesic $(n-2)$-sphere.*

This result is the key differential geometric fact used in proving the complete interior regularity of mass-minimizing hypersurfaces in n-manifolds for $n < 8$. Simons points out that the cone $C(S^3 \times S^3)$ on $S^3 \times S^3 = \{(x, y) \in \boldsymbol{R}^4 \times \boldsymbol{R}^4 : |x|^2 = |y|^2 = \frac{1}{2}\} \subset S^7$ is stable; and Bombieri, De Giorgi and Giusti [1] proved that, in fact, $C(S^3 \times S^3)$ is the unique solution to the Plateau problem for $S^3 \times S^3$ in $\boldsymbol{R}^8$. The analogous statement is true for a number of homogeneous cones including the cones on $M_{p,q} = \{(x, y) \in \boldsymbol{R}^{p+1} \times \boldsymbol{R}^{q+1} : |x|^2 = p/(p+q) \text{ and } |y|^2 = q/(p+q)\} \subset S^{p+q+1}$ for $p + q + 2 \geqq 8$ and $\{p, q\} \neq \{1, 5\}$ (cf. [11], [14]). The mass-minimizing 7-cones in $\boldsymbol{R}^8$ are known to be finite up to ambient isotopies (of S^7). However, the general question of classifying stable cones of codimension 1 in $\boldsymbol{R}^n$ is one of the major outstanding problems in the theory of minimal varieties.

The global structure of the stable currents without boundary on a Riemannian manifold M is intimately related to the global geometry of M. Of course, every nonzero class $\alpha \in H_p(M; G)$ for $G = \boldsymbol{Z}$ or $\boldsymbol{Z}/m\boldsymbol{Z}$ gives rise to a stable p-current by solving the Plateau problem for α. However, under appropriate curvature assumptions (in this case extrinsic curvature) such currents can be shown not to exist (cf. [13]).

THEOREM 3. *Let M be a compact n-manifold immersed in the unit sphere with second fundamental form A, and suppose that $\|A\|^2 < \min\{2(pq)^{1/2}, pq\}$ on M where $p + q = n$. Then there are no stable currents without boundary of dimension p or q on M, and, in particular, $H_p(M; G) = H_q(M; G) = 0$ for any finitely generated abelian group G.*

This result is sharp, since there exist embeddings of $S^p \times S^q$ into S^{p+q+1} with $\|A\|^2 \equiv 2(pq)^{1/2}$.

An immediate corollary of Theorem 3 is that there are no stable closed currents on S^n. This leads naturally to the question of the structure of the stable currents on the other positively curved (rank-one) symmetric spaces. In these cases there is homology, so stable currents must exist. In fact, in complex projective space $\boldsymbol{P}^n(\boldsymbol{C})$ a closed rectifiable current is a solution to the Plateau problem in its homology class if and only if it is a positive integral chain of algebraic subvarieties. This follows from a result of Federer [2] together with King's solution of Lelong's conjecture [8]. More generally, from [13] and [7] the following is true.

THEOREM 4. *A closed rectifiable current on $\boldsymbol{P}^n(\boldsymbol{C})$ is stable if and only if it is an integral chain of algebraic subvarieties.*

In particular, there are no stable odd-dimensional cycles on $\boldsymbol{P}^n(\boldsymbol{C})$. Note that in Theorem 4, arbitrary (not just positive) integral combinations of subvarieties are allowed. It is far from true that every stable current is homologically mass minimizing. Interestingly, the hypothesis of stability in Theorem 4 can be reduced to the assumption that the average second derivative of mass with respect to holomorphic deformations is zero.

No similar characterization of stable currents in quaternionic projective space has yet been found.

In the light of these results it seems natural to conjecture the nonexistence of stable closed currents on compact simply-connected manifolds whose sectional curvature is strictly $\frac{1}{4}$-pinched.

The remarks preceding Theorem 4 give an indication of the special role played by the Plateau problem in complex geometry. In fact, from the results of Federer and King referred to there, the following is true. Let M be a Kähler manifold and suppose that $V \in \mathscr{R}_{2p}(M)$ is a *positive holomorphic p-chain*, i.e., V is given by integration over a positive integral chain of complex p-dimensional subvarieties in $M - \operatorname{supp}(dV)$. Then V is a solution to the Plateau problem for $B = dV$ in M, and any other solution must also be a positive holomorphic p-chain. If $M = \boldsymbol{C}^n$ with the flat metric or if M is a Stein manifold with metric induced from a proper embedding into $\boldsymbol{C}^m$, then V is, in fact, the unique solution.

This raises the interesting question of which currents $B \in \mathscr{R}_{2p-1}(M)$ on a Stein manifold M are the boundaries of holomorphic p-chains. There are certain obvious necessary conditions for this to be true. If $p = 1$, then B must satisfy the *moment condition*: $B(\omega) = 0$ for all holomorphic 1-forms ω. If $p > 1$, then B must be *maximally complex*, i.e., $B(\omega) = 0$ for all (r, s)-forms ω with $|r - s| > 1$. In the case that B corresponds to integration over a C^1 submanifold $B^{2p-1} \subset M$, maximal complexity is equivalent to the local condition that

$$\dim_{\boldsymbol{R}} [T_x B^{2p-1} \cap J(T_x B^{2p-1})] = 2p - 2$$

for all $x \in B^{2p-1}$, where J is the almost complex structure on M.

It follows from generalizations of work of John Wermer [**16**] that if B is a system of C^1 oriented closed curves that satisfies the moment condition, then B is the boundary of a (unique, not necessarily positive) holomorphic 1-chain with compact support in M. For $\dim B > 1$, the following is true (cf. [**6**]).

THEOREM 5. *Let B^{2p-1} be a compact, oriented C^2 manifold of dimension $2p - 1 > 1$ embedded in a Stein manifold M. If B is maximally complex, then it is the boundary of a unique holomorphic p-chain with compact support in M. In particular, if B is also connected, then it is the boundary of a unique, irreducible complex p-dimensional subvariety V in $M - B$. Furthermore, there is a closed set $S \subset B$ of Hausdorff $(2p - 1)$-measure zero, such that every point in $B - S$ has a neighborhood in which $\bar{V}$ is a regular C^2 manifold with boundary.*

Interestingly, the "soft" version of this theorem is false. That is, on any Stein manifold M with $H_*(M; \boldsymbol{R}) \neq 0$, there exist (nonrectifiable) maximally complex $(2p - 1)$-cycles which are not the boundaries of real (p, p)-currents with compact support.

I re-emphasize that the condition of maximal complexity, as opposed to the moment condition (when $\dim(B) = 1$), is local in nature. This is related to the appearance in several complex variables of Hartogs' phenomenon. In fact, Theorem 5 gives a direct generalization of a result of Bochner on extending functions which satisfy the boundary Cauchy-Riemann equations.

COROLLARY 6. *Let B be a connected manifold as in Theorem 5, and let V be the irreducible complex subvariety which it bounds in M. Let $f \in C^2(B)$ be a complex-valued function which satisfies the tangential Cauchy-Riemann equations (i.e., the differential of f is complex linear when restricted to the complex subspaces of $T(B)$). Then there exists a weakly holomorphic function F on V which is continuous on $\bar{V}$ (and class C^2 at points of boundary regularity of $\bar{V}$) such that $F|B = f$.*

The corollary follows essentially from the observation that the graph of a function satisfying the tangential Cauchy-Riemann equations on a maximally complex submanifold is again maximally complex.

These last results can be generalized to manifolds which are somewhat larger holomorphically.

THEOREM 7. *Let $M = X - Y$ where X is a complex projective n-manifold and Y is a k-dimensional complete intersection on X. Then any real $(2p-1)$-manifold $B \subset M$ where $p + k > n$, which satisfies the hypotheses of Theorem 5, is the boundary of a unique holomorphic p-chain with compact support in M and with boundary regularity almost everywhere on B.*

All these results remain true when B is assumed to be a maximally complex, real analytic cycle (i.e., supp B is a finite union of $(2p-1)$-dimensional analytic blocks). In this form, Theorem 6 generalizes a result of Rossi on extending complex subvarieties defined in a neighborhood of Y.

The discussion above leads naturally to the question of when a class $\alpha \in H_{2p}(M; \boldsymbol{Z})$, on a compact Kähler manifold M, can be represented by a positive holomorphic p-cycle. One necessary condition is that $\alpha_{\boldsymbol{R}}$, the image of α under the coefficient homomorphism $\boldsymbol{Z} \subsetneq \boldsymbol{R}$, be dual to a class $\hat{\alpha}_{\boldsymbol{R}} \in H^{2(n-p)}(M; \boldsymbol{R})$ represented by a closed $(n-p, n-p)$-form. (Equivalently, $\langle \alpha, [\varphi] \rangle = 0$ for all closed (r, s)-forms φ with $|r - s| > 1$.)

Interestingly, this question leads to the study of a pair of related Plateau problems. Let $\|\alpha\|$ denote the infimum of the masses of the closed rectifiable currents in α, and, similarly, let $\|\alpha_{\boldsymbol{R}}\|$ denote the infimum of the masses (in the sense of Federer and Fleming [4]) of all closed de Rham currents in $\alpha_{\boldsymbol{R}}$. Then $\|m\alpha\| \geqq \|(m\alpha)_{\boldsymbol{R}}\| = m\|\alpha_{\boldsymbol{R}}\|$ for all positive integers m, and Federer has shown that $\lim_{m\to\infty} m^{-1}\|m\alpha\| = \|\alpha_{\boldsymbol{R}}\|$. Suppose now that $\hat{\alpha}_{\boldsymbol{R}}$ contains a *positive* closed $(n-p, n-p)$-form ω. Then Harvey and Knapp [5] have shown that the current $[\omega] \in \alpha_{\boldsymbol{R}}$, defined by setting $[\omega](\varphi) = \int_M \omega \wedge \varphi$ for smooth $2p$-forms φ, is a current of least mass in $\alpha_{\boldsymbol{R}}$ and, moreover, that α contains a holomorphic p-cycle if and only if $\|\alpha\| = \|\alpha_{\boldsymbol{R}}\|$.

A well-known result of Atiyah and Hirzebruch says that it is at least necessary, in general, to pass to positive integral multiples of α in order to find holomorphic

cycles. Thus, one is led to conjecture that if α is a homology class on a compact Hodge manifold M such that α_R is dual to a positive closed $(n-p, n-p)$-form, then:

(S) *There exists an integer* $m \geqq 1$ *such that* $m^{-1}\|m\alpha\| = \|\alpha_R\|$.

Unfortunately, this is not quite true (cf. [12]).

THEOREM 8. *There exists a class* $\alpha \in H_4(T; \boldsymbol{Z})$ *on a four-dimensional abelian variety* T, *which is dual to a positive, invariant* (2, 2)*-form, but no multiple of* α *carries a positive holomorphic cycle.*

One might then ask whether statement (S) holds for the classes $\alpha_n = \alpha + n[H^p]$ for all sufficiently large integers n, where H^p is a p-dimensional linear section obtained from some embedding $M \subset \boldsymbol{P}^N(\boldsymbol{C})$. This refined statement is equivalent to the condition that α contain a (not necessarily positive) holomorphic cycle.

References

1. E. Bombieri, E. De Giorgi and E. Giusti, *Minimal cones and the Bernstein theorem*, Invent. Math. **7** (1969), 243–269. MR **40** #3445.
2. H. Federer, *Some theorems on integral currents*, Trans. Amer. Math. Soc. **117**(1965), 43–67. MR **29** #5984.
3. ———, *Geometric measure theory*, Die Grundlehren der math. Wissenschaften, Band 153, Springer-Verlag, New York, 1969. MR **41** #1976.
4. H. Federer and W. H. Fleming, *Normal and integral currents*, Ann. of Math. (2) **72**(1960), 458–520. MR **23** #A588.
5. R. Harvey and A. Knapp, *Positive (p, p)-forms, Wirtinger's inequality and currents*, Proc. Tulane Conf. on Complex Analysis (to appear).
6. R. Harvey and H. B. Lawson, Jr., *On boundaries of complex analytic varieties*. I, II (to appear).
7. R. Harvey and B. Shiffman, *A characterization of holomorphic chains*, Ann. of Math. (2) **99** (1974), 553–587.
8. J. King, *The currents defined by analytic varieties*, Acta Math. **127**(1971), 185–200.
9. H. B. Lawson, Jr., *Complete minimal surfaces in* S^3, Ann. of Math. (2) **92**(1970), 335–374. MR **42** #5170.
10. ———, *The unknottedness of minimal embeddings*, Invent. Math. **11**(1970), 183–187. MR **44** #4651.
11. ———, *The equivariant Plateau problem and interior regularity*, Trans. Amer. Math. Soc. **173**(1972), 231–249. MR **46** #8017.
12. ———, *The stable homology of a flat torus*, Math. Scand. **36** (1975).
13. H. B. Lawson, Jr. and J. Simons, *On stable currents and their application to global problems in real and complex geometry*, Ann. of Math. (2) **98** (1973), 427–450. MR **48** #1881.
14. P. Simões, *On a class of minimal cones in* R^n, Bull. Amer. Math. Soc. **80**(1974), 488–489.
15. J. Simons, *Minimal varieties in riemannian manifolds*, Ann. of Math. (2) **88**(1968), 62–105. MR **38** #1617.
16. J. Wermer, *The hull of a curve in* C^n, Ann. of Math. (2) **68**(1958), 550–561. MR **20** #6536.

UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720, U.S.A.

Actes du Congrès International des Mathématiciens
Vancouver, 1974

Problèmes de Géométrie Conforme

J. Lelong-Ferrand

Soit (M,g) une variété riemannienne de classe C^∞ et de dimension $n \geqq 3$. Plaçons nous d'abord au point de vue usuel de la géométrie différentielle et notons $C(M, g)$ [resp. $I(M, g)$] le groupe des C^∞-difféomorphismes conformes de M [resp. le groupe des isométries de M]. D'après les théorèmes de Montgomery et Kobayashi on sait que $C(M, g)$ est un groupe de Lie pour la topologie compacte ouverte: c'est en effect, le groupe des automorphismes de la structure conforme de M, qui est une G-structure de type fini. Un sous-groupe G de $C(M, g)$ est dit essentiel s'il n'existe aucune fonction régulière ρ sur M telle que $G \subset I(M, e^{2\rho}g)$, inessentiel dans le cas contraire. Cela étant le problème suivant (conjecture de Lichnérowicz) a fait l'objet de nombreux travaux:

Existe-t-il des variétés non conformes à l'espace euclidien (E^n, g_0) ou à la sphère standard (S^n, g_0) pour lesquelles la composante connexe $C_0(M, g)$ de $C(M, g)$ soit essentielle?

Ce problème a d'abord été résolu sous diverses hypothèses supplémentaires, en particulier: M est compacte [5], [9]. [10] et: (M, g) est une variété d'Einstein complète [13] (pour une bibliographie plus détaillée, voir [5] ou [9]).

Une réponse définitive (et négative) nous est maintenant donnée par le théorème suivant de D.V. Alekseevski [1]:

THÉORÈME 1. *Si $C_0(M, g)$ est essentiel, (M, g) est conforme à (E^n, g_0) ou à (S^n, g_0).*

On en déduit facilement:

THÉORÈME 1 BIS. *Si (M, g) est compacte et si $C_0(M, g)$ est non compact, (M, g) est conforme à (S^n, g_0).*

Ce dernier résultat, qui apparaît maintenant comme un corollaire du Théorème 1,

avait été établi presque simultanément par M. Obata et moi-même, par des méthodes essentiellement différentes: c'est par une étude précise des modules de continuité des transformations quasi-conformes que j'établis l'équicontinuité de $C(M, g)$ lorsque (M, g) n'est pas conforme à (S^n, g_0); M. Obata utilise au contraire les techniques propres à la géométrie riemannienne et prouve que, si $C_0(M, g)$ est non compact, (M, g) est conformément plate; il y parvient en étudiant l'action de $C_0(M, g)$ sur l'ouvert formé des points de M où le tenseur de courbure conforme est non nul. Par un théorème de N. Kuiper [4] on est alors ramené à l'étude d'un ouvert de (S^n, g_0); on utilise le fait que $C_0(M, g)$ contient un sous-groupe essentiel à un paramètre et qu'un tel sous-groupe G admet un point fixe; le résultat est obtenu par une analyse du comportement de G en ses points fixes.

La démonstration du Théorème 1 donnée par D.V. Alekseevski n'est pas sans analogie avec celle du Théorème 1 bis par Obata. Elle est fondée sur la notion de groupe isotropiquement compact, i.e., tel que le sous-groupe d'isotropie de chaque $P \in M$ soit compact; et elle se décompose en trois parties:

(A) Un sous-groupe essentiel G de $C(M, g)$ n'est pas isotropiquement compact.

(B) Si $C_0(M, g)$ n'est pas isotropiquement compact, il contient un sous-groupe essentiel à un paramètre.

(C) Si (M, g) admet un groupe essentiel à un paramètre de transformations conformes (M, g) est conforme à (E^n, g_0) ou à (S^n, g_0).

Pour établir (A), D.V. Alekseevski considère un sous-groupe fermé et isotropiquement compact G de $C(M, g)$ et montre tout d'abord que G opère proprement dans M au sens de Bourbaki; puis il construit, sur un voisinage G-invariant U_p d'un point arbitraire $p \in M$, une métrique G-invariante et conforme à g. Enfin, il construit une partition G-invariante de l'unité, subordonnée au recouvrement $\{U_p\}$, et il en déduit l'existence sur M d'une métrique G—invariante et conforme à g, prouvant ainsi que G est inessentiel.

Pour établir (B) il suffit de prouver que si le groupe d'isotropie C_p d'un point P est non compact, sa composante connexe est non compacte. On se ramène au cas où $C_0(M, g)$ est transitif en considérant l'orbite M' de p, et en prouvant que le groupe facteur C' agissant sur M' est essentiel. La fin de la démonstration utilise une représentation de C_p dans $C(E_n)$ et une étude fine des transformations conformes au voisinage de leurs points fixes, fondée sur un développement limité d'ordre 2.

Pour établir (C) on montre d'abord que chaque point fixe p d'un sous-groupe essentiel à un paramètre de $C_0(M, g)$ admet un voisinage ouvert et G-invariant V sur lequel le tenseur de courbure conforme s'annule; il ne reste ensuite qu'à prouver que ∂V est vide ou réduit à un point.

Enfin D.V. Alekseevski signale en note que le Théorème 1 reste vrai si l'on remplace $C_0(M, g)$ par $C(M, g)$ lui-même.

Par contre, il n'existe pas de résultat analogue pour les variétés pseudo-riemanniennes.

Etude directe du cas compact. Etant donnée la complexité des techniques utilisées par D.V. Alekseevski il n'est peut-être pas sans intérêt de rappeler ici la méthode

relativement élémentaire que j'ai utilisée dans le cas compact; de plus cette méthode fait apparaître le Théorème 1 bis comme une conséquence d'un théorème d'analyse plus général, relatif aux transformations quasi-conformes.

DÉFINITION 1. Soit (M, g) et $(\bar{M}, \bar{g})$ deux variétés riemanniennes de classe C^1 et de dimension $n \geqq 2$. Nous dirons qu'un homeomorphisme $\phi : M \to \bar{M}$ est k-quasi-conforme si:

(i) ϕ est ACL^n (i.e., si ϕ est absolument continue sur presque toute ligne coordonnée relative à un atlas donné, et si ses dérivées partielles appartiennent à $L^n_{\mathrm{loc}}(M)$).

(ii) Pour presque tout $x \in M$ la différentielle de ϕ et son jacobien métrique J_ϕ vérifient:

$$(1) \qquad J_\phi(x) \neq 0 \quad \text{et} \quad |\phi'(x)|^n \leqq k^n J_\phi(x)$$

(la condition (i) entraînant l'existence de $\phi'(x)$ pour presque tout x).

En tout point x où $\phi'(x)$ existe et vérifie $\phi'(x) \neq 0$, on a nécessairement $|\phi'(x)|^n \geqq J_\phi(x)$. On a donc toujours $k \geqq 1$; et si $k = 1$, $\phi'(x)$ est une similitude: les homéomorphismes 1-quasi-conformes peuvent donc être dits conformes; et, pour prouver que ce sont des transformations conformes au sens classique, il suffit de prouver que ce sont des difféomorphismes (voir plus loin).

Désignons par $Q_k(M, \bar{M})$ l'ensemble des homéomorphismes k-quasi-conformes de M sur $\bar{M}$, muni de la topologie compacte ouverte. On a alors [**5b**, Théorème 9.3]:

THÉORÈME 2. *Si M, $\bar{M}$ sont compactes et si $Q_k(M, \bar{M})$ n'est pas compact, il existe un homéomorphisme K-quasi-conforme de (M, g) sur (S^n, g_0).*

En prenant $\bar{M} = M$ et $k = 1$, on en déduit:

COROLLAIRE. *Si le groupe des homéomorphismes conformes d'une variété riemannienne compacte (M, g) n'est pas compact il existe un homéomorphisme conforme de (M, g) sur (S^n, g_0).*

Si (M, g) et $(\bar{M}, \bar{g})$ sont de classe C^∞, on peut prouver que les homéomorphismes conformes de (M, g) sur $(\bar{M}, \bar{g})$ sont des difféomorphismes (voir plus loin). Du corollaire précédent on déduit alors que le Théorème 1 bis reste vrai en remplaçant $C_0(M, g)$ par $C(M, g)$.

La théorie générale des transformations quasi-conformes permet de montrer que $Q_k(M, \bar{M})$ est une partie fermée de l'ensemble des homéomorphismes de M sur $\bar{M}$; pour prouver que $Q_k(M, \bar{M})$ est compact il suffit donc (si M et $\bar{M}$ sont compactes) de prouver que la famille $Q_k(M, \bar{M})$ et la famille formée par les réciproques des éléments de $Q_k(M, \bar{M})$ sont équicontinues. Nous allons donner un aperçu de la démonstration.

Problèmes d'équicontinuité. La construction de modules de continuité effectuée dans [**5b**] pour les applications quasi-conformes a été reprise de manière plus systématique dans [**8**]; pour la rendre plus intuitive, nous introduirons ici une définition:

DÉFINITION 2. Une partie X de la variété (M, g) sera dite de diamètre apparent

$\leqq \delta$ s'il existe une partie compacte propre de M, contenant X et difféomorphe à une boule de E^n, dont la frontière ait un diamètre géodésique $\leqq \delta$.

La proposition suivante résulte alors du Lemme (5.2) de [8] et du fait qu'une variété riemannienne est localement difféomorphe à E^n:

PROPOSITION 1. *Les notations étant celles de la Définition* 1, *soit* $\phi:(M, g) \to (\bar{M}, \bar{g})$ *une application* ACL^n *verifiant*

$$\int_M |\phi'|^n \, d\tau \leqq m. \tag{2}$$

Pour chaque compact $K \subset M$, *il existe un nombre* $\alpha_k > 0$, *ne dépendant que de K, tel que l'image par* ϕ *d'une boule géodésique centrée sur K et de rayon* $\rho < \alpha_k$ *ait, dans* $\bar{M}$, *un diamètre apparent inférieur à* $(mB_n/|\operatorname{Log} r|)^{1/n}$, *où* B_n *désigne une constante dépendant seulement de n.*

La condition (2) est vérifiée par tout homéomorphisme k-quasi-conforme $\phi:M \to \bar{M}$, si $\bar{M}$ est de mesure finie $\leqq mk^{-n}$. D'autre part, on a la proposition suivante, qui résulte du Lemme 7.3 de [5b] et qui permet d'établir un lien entre le diamètre apparent et le vrai diamètre.

PROPOSITION 2. *Si* $(\bar{M}, \bar{g})$ *est compacte il existe un nombre* $\gamma > 0$ *possédant la propriété suivante: si X est une partie de* $\bar{M}$ *de diamètre apparent* $\delta < \gamma$, *l'un des ensembles* X, $\bar{M}\backslash X$ *a un diamètre géodésique* $\leqq 4\delta$.

Des Propositions 1 et 2 on déduit le critère d'équicontinuité qui suit:

PROPOSITION 3. *Soit* Φ *une famille d'homéomorphismes k-quasi-conformes de la variété compacte* (M, g) *sur la variété compacte* $(\bar{M}, \bar{g})$; *pour que* Φ *soit uniformément équicontinue, il suffit qu'il existe un nombre* $h > 0$ *tel qu'à chaque* $\phi \in \Phi$ *on puisse associer trois points* a_1, a_2, a_3 *de M vérifiant* $d_M(a_i, a_j) \geqq h$ *et* $d_{\bar{M}}(\phi(a_i), \phi(a_j)) \geqq h$ *pour* $i, j = 1, 2, 3$ $(i \neq j)$ (d_M *et* $d_{\bar{M}}$ *désignant les distances géodésiques sur M et* $\bar{M}$).

Si la famille Φ n'est pas équicontinue, ce lemme montre, en gros, qu'on peut en extraire une suite convergeant vers une transformation dégénérée. De façon précise on a:

PROPOSITION 4. *Soit* Φ *une famille non équicontinue d'homéomorphismes k-quasi-conformes de la variété compacte* (M, g) *sur la variété compacte* $(\bar{M}, \bar{g})$; *alors il existe une suite* (ϕ_p) *extraite de* Φ, *un point a de M et un point b de* $\bar{M}$ *tels que*:

(i) *la suite* $(\phi_p(x))$ *converge vers b pour tout* $x \in M\backslash\{a\}$, *la convergence étant uniforme sur tout compact de* $M\backslash\{a\}$;

(ii) *la suite* $(\phi_p^{-1}(y))$ *converge vers a pour tout* $y \in \bar{M}\backslash\{b\}$, *la convergence étant uniforme sur tout compact de* $\bar{M}\backslash\{b\}$.

Par une nouvelle application de la Proposition 3, on montre que l'existence d'une telle suite (ϕ_p) implique celle d'un homéomorphisme k-quasi-conforme de $M\backslash\{a\}$ sur (E^n, g_0); et, par prolongement, on en déduit l'existence d'un homéomor-

phisme k-quasi-conforme de M sur (S^n, g_0) (cf. demonstration du Théorème 8 dans [**5b**]); d'où le Théorème 2.

Notons que l'existence d'une suite de difféomorphismes conformes de (M, g) sur $(\bar{M}, \bar{g})$ vérifiant (i) et (ii) permettrait de montrer que (M, g) et $(\bar{M}, \bar{g})$ sont conformément plates. On en déduirait une démonstration "mixte" du Théorème 1 bis.

Constructions d'invariantes conformes. C'est pour préparer une démonstration du Théorème 1 que, dans [7] j'avais cherché à construire, sur les variétés riemaniennes non compactes, des métriques (au sens général) conformément invariantes. En fait il s'agit d'invariants globaux liés à la structure conforme de (M, g), et qui présentent par eux-mêmes un intérêt, même lorsque le groupe conforme de (M, g) se réduit à l'identité. J'en dirai donc ici quelques mots.

A chaque variété (M, g) de dimension n associons la classe $H^*(M)$ formée des fonctions numériques u, continues et ACL^n sur M, telles que:

(i) $I(u, M) = \int_M |\nabla u|^n \, d\tau < \infty$;

(ii) pour chaque partie ouverte, connexe et relativement compacte U de M, on ait:

$$\sup_{x \in \partial U} u(x) = \sup_{x \in U} u(x), \quad \inf_{x \in \partial U} u(x) = \sup_{x \in U} u(x).$$

Si la famille $H^*(M)$ sépare les points de M, on obtient une distance δ_M sur M en posant:

$$\delta_M(x, y) = \sup_{u \in H^*(M)} \frac{|u(x) - u(y)|}{(I(u, M))^{1/n}}.$$

Cette distance est (comme l'intégrale $I(u, M)$) invariante par déformation conforme de M; et la topologie qu'elle définit est moins fine que celle de (M, g): pour cette distance $C(M, g)$ est donc un groupe d'isométries.

Interprétation fonctionnelle des homéomorphismes conformes. Une autre méthode d'approche des transformations conformes consiste à leur associer des opérateurs linéaires.

A chaque variété (M, g) de dimension $n \geqq 2$ associons l'algèbre de Banach $A(M, g)$ formée des fonctions numériques continues et ACL^n sur M, vérifiant: $I(u, M) = \int_M |\nabla u|^n \, d\tau < \infty$ et tendant vers zéro à l'infini, pour la norme $\nu(u) = \|u\|_\infty + [I(u, M)]^{1/n}$.

Si $\phi : M \to \bar{M}$ est une application continue, notons ϕ^* l'application $\mathscr{C}(\bar{M}) \to \mathscr{C}(M)$, $v \mapsto v \circ \phi$. On montre alors [**6**] que l'application $\phi \mapsto \phi^*$ définit une bijection de l'ensemble des homéomorphismes k-quasi-conformes $\phi : M \to \bar{M}$, sur l'ensemble des isomorphismes de norme $\leqq k$ de $A(\bar{M}, \bar{g})$ sur $A(M, g)$. Le groupe conforme $C(M, g)$ s'identifie donc à une partie de la sphère unité de $\mathscr{L}(A(M, g))$; et le problème qui se pose est de comparer les diverses topologies que l'on peut définir naturellement sur cet ensemble.

Problèmes de régularité. Pour faire le lien entre les méthodes d'analyse utilisées dans [**5**] et celles de la géométrie différentielle, il reste à prouver que tout homéo-

morphisme conforme d'une variété riemannienne de classe C^∞ sur une variété de classe C^∞ est lui-même de classe C^∞. L'historique du cas euclidien, que nous allons esquisser, montre que ce problème n'est pas trivial.

En 1850 Liouville prouvait que tout difféomorphisme conforme ϕ d'ouverts de (E^n, g_0) $n \geqq 3$ était la restriction d'une transformation de Möbius. Mais sa démonstration supposait implicitement ϕ de classe C^3; et à ma connaissance, ce n'est qu'en 1947 qu'il fut prouvé, par P. Hartman [3], qu'il suffisait de supposer ϕ de classe C^1: la démonstration de P. Hartman utilise les propriétés des systèmes différentiels elliptiques sur-déterminés.

En 1960, J.G. Resetnyak [11] prouve que tout homéomorphisme conforme d'ouverts de R^n est de classe C^3; et, en 1962 F.W. Gehring [2] donne une autre démonstration du même résultat: tous deux utilisent les propriétés des solutions d'équations elliptiques, mais la démonstration de Gehring utilise les extrémales de l'intégrale $\int|\nabla u|^n\, d\tau$, tandis que celle de Resetnyak est fondée sur des inégalités isopérimétriques et utilise la théorie du potentiel.

Aucune de ces démonstrations ne se généralise simplement aux variétés riemanniennes; mais l'étude du cas riemannien, bien que plus général, est facilitée par l'introduction de la fonction courbure scalaire. En effet, tout homéomorphisme conforme ϕ: $(M, g) \to (\bar{M}, \bar{g})$ peut être considéré comme une isométrie de $(M, |\phi'|^2 g)$ sur $(\bar{M}, \bar{g})$. Or, si u est de classe C^2 sur M, la courbure scalaire R_u de $(M, u^{4/(n-2)} g)$ est liée à la courbure scalaire R de (M, g) par

$$u^{4/(n-2)} R_u = R - 4\frac{n-1}{n-2}\frac{\Delta u}{u}.$$

On peut donc s'attendre à ce que la fonction $u = |\phi'|^{(n-2)/2}$ admette au sens des distributions, un laplacien vérifiant:

$$\Delta u = \frac{n-2}{4(n-1)}(Ru - \bar{R} \circ \phi u^{(n+2)/(n-2)}) \tag{3}$$

ou $\bar{R}$ désigne la courbure scalaire de $(\bar{M}, \bar{g})$; et on en déduit que u (donc aussi ϕ) est de classe C^∞ : on peut en effet se ramener localement au cas où $R = \bar{R} = 0$, auquel cas $u = |\phi'|^{(n-2)/2}$ est harmonique. Cependant, une démonstration rigoureuse de (3) s'impose: ce sera l'objet d'un autre article.

On notera que la demonstration de Resetnyak consiste à prouver, par des considérations geométriques, que, dans le cas euclidien, $|\phi'|^{(n-2)/2}$ est sous-harmonique.

Un problème ouvert. Convenons de dire qu'une application continue (non nécessairement bijective) ϕ: $(M, g) \to (\bar{M}, \bar{g})$ est conforme si elle est ACL^n, si elle conserve l'orientation et si, pour presque tout $x \in M$, $\phi'(x)$ est une similitude (c'est une extension aux variétés riemanniennes de la notion d'application 1—quasi-régulière).

En 1967, Resetnyak [12] a pu montrer que toute application conforme d'un ouvert de (E^n, g_0) dans E^n était un homéomorphisme (donc la restriction d'une transformation de Möbius). Pour les variétés riemanniennes, le problème se pose donc de savoir s'il existe des applications conformes qui ne sont pas localement des difféomorphismes.

Le problème se pose également de savoir si, pour des variétés de classe C^1, tout homéomorphisme conforme est aussi de classe C^1.

Bibliographie

1. D. V. Alekseevski, *Groups of conformal transformations of Riemannian spaces*, Mat. Sb. **89** (131) (1972), 280–296 = Math. USSR Sb. **18** (1972), 285–301.

2. F.W. Gehring, *Rings and quasiconformal mappings in space*, Trans. Amer. Math. Soc. **103** (1962), 353–393. MR **25** #3166.

3. P. Hartman, *On isometries and on a theorem of Liouville*, Math. Z. **69** (1958), 202–210. MR **21** #7521.

4. N.H. Kuiper, *On conformally-flat spaces in the large*, Ann. of Math. (2) **50** (1949), 916–924. MR **11**, 133.

5a. J. Lelong-Ferrand, *Transformations conformes et quasiconformes des variétés riemannienne; application à la démonstration d'une conjecture de A. Lichnerowicz*, C.R. Acad. Sci. Paris Sér. A-B **269** (1969), A583–A586. MR **40** #7989.

5b. ———, *Transformations conformes et quasiconformes des variétés riemanniennes compactes (démonstration de la conjecture de A. Lichnerowicz)*, Acad. Roy. Belg. Cl. Sci. Mém. Coll. in-8° (2) **39**, no. 5, (1971). MR **49** #1100.

6. ———, *Étude d'une classe d'applications liées à des homomorphismes d'algèbres de fonctions et généralisant les quasi conformes*, Duke. Math. J. **40** (1973), 163–186. MR **47** #4177.

7. J. Lelong-Ferrand, *Invariants conformes globaux sur les variétés riemanniennes*, J. Differential Geometry **8** (1973), 487–510.

8. ———, *Construction de modules de continuité dans le cas limite de Soboleff*, Arch. Rational Mech. Anal. **52** (1973), 297–311.

9. M. Obata, *The conjectures on conformal transformations of Riemannian manifolds*, Bull. Amer. Math. Soc. **77** (1971), 269–270. MR **42** #5286.

10. A.J. Ledger and M. Obata, *Compact Riemannian manifolds with essential groups of conformorphisms*, Trans. Amer. Math. Soc. **150** (1970), 645–651. MR **41** #7576.

11. Ju. G. Rešetnjak, *On conformal mappings of a space*, Dokl. Akad. Nauk SSSR **130** (1960), 1196–1198 = Soviet Math. Dokl. **1** (1960), 153–155. MR **22** #9935.

12. ———, *Liouville's conformal mapping theorem under minimal regularity hypotheses*, Sibirsk. Mat. Ž. **8** (1967), 835–840. (Russian) MR **36** #1630.

13. K. Yano and T. Nagano, *Einstein spaces admitting a one-parameter group of conformal transformations*, Ann. of Math. (2) **69** (1959), 451–461. MR **21** #345.

14. K. Yano and M. Obata, *Conformal changes of Riemannian metrics*, J. Differential Geometry **4** (1970), 53–72. MR **41** #6113.

UNIVERSITÉ DE PARIS VI
PARIS, FRANCE

Труды Международного Конгресса Математиков
Ванкувер, 1974

Дискретные Группы Движений Многообразий Неположительной Кривизны*

Г. А. Маргулис

Доклад в основном посвящен результатам об арифметичности и конечномерных представлениях дискретных групп движений симметрических пространств. Приводятся также обобщения этих результатов на дискретные подгруппы в p-адических группах, из которых, в свою очередь, вытекает, в частности, "почти окончательная" классификация конечномерных представлений группы k-точек произвольной полупростой односвязной алгебраической k-группы, где k—конечное расширение поля рациональных чисел. Далее формулируются некоторые нерешенные задачи о дискретных группах движений в симметрических пространствах и в более общих многообразиях неположительной кривизны. Последняя часть доклада посвящена эквивариантным измеримым отображениям, играющим основную роль в доказательстве упомянутых результатов. При построении этих отображений существенно используются методы эргодической теории и теории представлений.

1. Арифметичность и конечномерные представления решеток в полупростых вещественных и p-адических группах. Как обычно, $\boldsymbol{C}$, $\boldsymbol{R}$, $\boldsymbol{Q}$, $\boldsymbol{Z}$, $\boldsymbol{N}^+$, $\boldsymbol{Q}_p$ обозначают соответственно множества комплексных, вещественных, рациональных, целых, целых < 0, p-адических чисел. Для любой линейной алгебраической k-группы N и любого кольца K, вложенного в одно с k кольцо, через N_K или $N(K)$ обозначается множество точек группы N, определенных над K, определитель которых является обратимым элементом кольца K. Для любой алгебраической $\boldsymbol{R}$-группы H через $H^0_{\boldsymbol{R}}$ будем обозначать связную компоненту единицы группы Ли $H_{\boldsymbol{R}}$.

*Not presented in person.

Дискретная подгруппа Λ локально компактной группы H называется *решеткой*, если объем факторпространства H/Λ (относительно меры Хаара) конечен. Решетка Λ называется *равномерной*, если H/Λ компактно, и *неравномерной* в противном случае. Решетка $\Lambda \subset H$ называется *неприводимой*, если для любого непрерывного эпиморфизма $f\colon H \to F$ подгруппа $f(\Lambda)$ недискретна в F, как только группы $\operatorname{Ker} f$ и F бесконечны.

Для любого подмножества M алгебраической группы F через $\bar{M}$ будет обозначаться замыкание по Зарисскому M в F. Через $\operatorname{char} k$ обозначается характеристика поля k.

1. Скажем, что алгебраическая $\boldsymbol{R}$-группа G *не имеет компактных факторов*, если у группы Ли $G^0_{\boldsymbol{R}}$ нет бесконечных компактных факторгрупп. Две подгруппы называются *соизмеримыми*, если их пересечение имеет конечный индекс в каждой из них.

Теорема 1. *Пусть G—связная присоединенная полупростая алгебраическая $\boldsymbol{R}$-группа без компактных факторов, $\boldsymbol{R}$—ранг (т.е. размерность максимального расщепимого над $\boldsymbol{R}$ тора) которой больше* 1, *а Γ—неприводимая решетка в $G^0_{\boldsymbol{R}}$. Тогда Γ является арифметической подгруппой в G, т.е. найдется такая алгебраическая $\boldsymbol{Q}$-группа G' и такой определенный над $\boldsymbol{R}$ эпиморфизм $f\colon G' \to G$, что группа Ли $(\operatorname{Ker} f)_{\boldsymbol{R}}$ компактна, а подгруппы $f(G'_{\mathbf{Z}})$ и Γ соизмеримы.*

Теоремы 1 вместе с теоремой Бореля-Харриш-Чандры о конечности объема факторпространства полупростой группы Ли по арифметической подгруппе [25] (см. также [24], [30]) дает классификацию с точностью до соизмеримости всех неприводимых решеток в полупростых вещественных группах $\boldsymbol{R}$-ранга большего 1. Теорему 1 можно переформулировать на языке теории симметрических пространств.

Теорема 1′. *Пусть S—риманово симметрическое пространство некомпактного типа ранга большего* 1, *а Γ—неприводимая дискретная группа движений пространства S, для которой факторпространство S/Γ имеет конечный объем. Тогда Γ является арифметической подгруппой в группе движений пространства S.*

Теоремы 1 и 1′ высказывались в качестве гипотезы Сельбергом в случае, когда Γ неравномерна [44], [45] и Пятецким-Шапиро в общем случае [17], [18]. Ими же были получены важные результаты в направлении доказательства этой гипотезы [16], [17], [18], [44], [46]. Первоначально [45] предполагалось, что Теорема 1 верна всякий раз, когда $G \neq SL(2, \boldsymbol{R})$. Но построенные Макаровым [8] и Винбергом [4], [5] примеры неарифметических решеток в 3, 4 и 5-мерных пространствах Лобачевского заставили предположить, что $\boldsymbol{R}$-ранг группы G больше 1. Для неравномерных Γ Теорема 1 была ранее доказана докладчиком методом, который основан на изучении унипотентных элементов подгруппы Γ и поэтому неприменим к равномерным решеткам (доказательство для "почти всех" G см. [10]). Существенное продвижение в направлении

доказательства Теоремы 1 в случае неравномерных Γ было также получено Рагунатаном [43].

Если заданы конечномерное представление T группы H над полем k и вложение σ поля k в другое поле k', то, применяя гомоморфизм σ к коэффициентам матриц $T(h)$, $h \in H$, мы получаем новое представление, которое естественно обозначить через $\sigma \circ T$. Представление T подгруппы Γ алгебраической группы F над полем k назовем *псевдорациональным*, если найдется такое рациональное представление $\tilde{T}$ группы F над некоторым полем k' и такое вложение $\sigma: k \to k'$, что $\sigma \circ T = \tilde{T}|\Gamma$ (т.е. $\sigma \circ T(\gamma) = \tilde{T}(\gamma)$ для любого $\gamma \in \Gamma$). Представление T группы Γ назовем *вырожденным*, если группа $T(\Gamma)$ конечна.

Теорема 2. *Пусть G и Γ те же, что и в Теореме* 1, *а T—конечномерное представление группы Γ над полем k. Тогда*

(1) *если представление T неприводимо, а алгебраическая группа $\overline{T(\Gamma)}$ связна и проста, то T псевдорационально,*

(2) *если* $\operatorname{char} k \neq 0$, *то T вырождено,*

(3) *если* $\operatorname{char} k = 0$, *то T вполне приводимо.*

Отметим, что утверждение п. 1 стало бы неверным, если бы мы вместо простоты группы $\overline{T(\Gamma)}$ потребовали только почти простоту. Мы приведем набросок доказательства Теоремы 1 и п. 1 Теоремы 2 (п. 2 вытекает из п. 1, а доказательство п. 3 основано на близких соображениях). Основную роль в доказательстве Теоремы 1 и п. 1 Теоремы 2 играет

Теорема 3. *Пусть G и Γ те же, что и в Теореме* 1, *а T—такое (абсолютно) неприводимое конечномерное представление группы Γ над некоторым локальным (т.е. недискретным локально компактным) полем K, что группа $\overline{T(\Gamma)}$ связна и проста. Тогда*

(1) *если группа $T(\Gamma)$ не является относительно компактной в K-топологии, то K есть $\boldsymbol{R}$ или $\boldsymbol{C}$; иначе говоря, если K не есть $\boldsymbol{R}$ или $\boldsymbol{C}$, т.е. K вполне несвязно, то группа $T(\Gamma)$ относительно компактна в K-топологии,*

(2) *если K есть $\boldsymbol{R}$ или $\boldsymbol{C}$, а группа $T(\Gamma)$ не является относительно компактной в K-топологии, то T продолжается до рационального представления группы G.*

Чтобы доказать Теорему 1 (п. 1 Теоремы 2) надо применить Теорему 3 к неприводимым компонентам представлений $\sigma \circ \mathrm{Ad}$ (представлений $\sigma \circ T$) для всех вложений σ наименьшего поля определения [5] группы $\mathrm{Ad}\, \Gamma(T(\Gamma))$ в локальные поля. При этом используется теорема о том, что группа Γ конечно порождена (см. [42, Замечание 13.21]).

Теорему 3 нетрудно вывести из приводимых ниже Теорем 4, 5 и 6 и теоремы плотности Бореля [22]. В Теореме 4 утверждается конечность фактора группы Γ по ее коммутанту. Эта теорема применяется в случае, когда представление T одномерно. В Теореме 5 утверждается существование "хороших" эквивариантных измеримых отображений, а в Теореме 6 утверждается, что для групп G $\boldsymbol{R}$-ранга большего 1 эти (а также несколько более общие) отображения рациональны.

ТЕОРЕМА 4. *Пусть G—полупростая алгебраическая $\boldsymbol{R}$-группа без компактных факторов, $\boldsymbol{R}$-ранг которой больше* 1, *а Γ—неприводимая решетка в $G^0_{\boldsymbol{R}}$. Тогда фактор группы Γ по ее коммутанту конечен.*

В случае, когда G не имеет факторов $\boldsymbol{R}$-ранга 1 Теорема 4 доказана Канеуки и Нагано [**31**] для равномерных Γ и Кожданом [**3**], [**6**], [**26**], [**45**] для всех Γ. В случае, когда G не $\boldsymbol{R}$-проста, а Γ равномерна, Теорема 4 доказана Бернштейном и Кожданом [**1**]. В остальных случаях, т.е. когда G не $\boldsymbol{R}$-проста и имеет факторы $\boldsymbol{R}$-ранга 1, а Γ неравномерна, метод Бернштейна-Каждана также применим, если использовать результаты Рагунатана о строении фундаментальных областей [**42**, Теоремы 13.12 и 13.19].

Отображение пространства X с мерой μ в топологическое пространство (пространство с мерой) называется *измеримым* по мере μ, если оно определено почти всюду и прообраз любого открытого (измеримого) множества измерим. Для локального поля K и $n \in \boldsymbol{N}^+$ через $\mathrm{Gr}(n, K)$ обозначим грассманово многообразие (всех линейных подпространств в K^n) с K-топологией.

ТЕОРЕМА 5. *Пусть G—полупростая алгебраическая $\boldsymbol{R}$-группа без компактных факторов, Γ—неприводимая решетка в $G^0_{\boldsymbol{R}}$, $n \in \boldsymbol{N}^+$, K—локальное поле, а T—такое n-мерное (абсолютно) неприводимое представление группы Γ над полем K, что $T(\Gamma) \subset SL(n, K)$, подгруппа $T(\Gamma)$ не является относительно компактной в K-топологии, а алгебраическая группа $\overline{T(\Gamma)}$ связна. Предположим, что выполнено хотя бы одно из следующих трех условий*: (1) *решетка Γ равномерна,* (2) *в Γ найдется такая подгруппа Γ'' конечного индекса, что T переводит унипотентные элементы из Γ'' в унипотентные матрицы,* (3) *$\boldsymbol{R}$-ранг группы G больше* 1. *Тогда найдется такая параболическая $\boldsymbol{R}$-подгруппа P группы G и такое измеримое (по мере Хаара) отображение ω группы Ли $G^0_{\boldsymbol{R}}$ в грассманово многообразие $\mathrm{Gr}(n, K)$, что для почти всех (по мере Хаара) $g \in G^0_{\boldsymbol{R}}$*

(1) $0 < \dim \omega(g) < n$,

(2) *отображение ω эквивариантно, т.е. $\omega(g\gamma) = T(\gamma^{-1})\omega(g)$ для любого $\gamma \in \Gamma$,*

(3) $\omega(pg) = \omega(g)$ *для любого $p \in P \cap G^0_{\boldsymbol{R}}$.*

Теорема 5 занимает центральное место в доказательстве Теоремы 3. Доказательство Теоремы 5 основано на применении мультипликативной эргодической теоремы и методов теории представлений. Существенную роль в доказательстве играет также одна модификация теоремы Титса о свободных подгруппах линейных групп. Мы более подробно остановимся на доказательстве Теоремы 5, а также Теоремы 6, в последней части доклада.

ТЕОРЕМА 6. *Пусть G—полупростая алгебраическая $\boldsymbol{R}$-группа без компактных факторов, $\boldsymbol{R}$-ранг которой больше* 1. *Γ—неприводимая решетка в $G^0_{\boldsymbol{R}}$, P—параболическая $\boldsymbol{R}$-подгруппа в G, K—локальное поле, F—алгебраическая K-группа, K-рационально действующая на алгебраическом K-много-*

образии M, T—гомоморфизм подгруппы Γ в группу F_K, ω—измеримое (по мере Хаара) отображение группы Ли G_R^0 в множество M_K K-точек многообразия M. Предположим, что для почти всех (по мере Хаара) $g \in G_R^0$,

(1) $\omega(g\gamma) = T(\gamma^{-1})\omega(g)$ *для любого* $\gamma \in \Gamma$,

(2) $\omega(pg) = \omega(g)$ *для любого* $p \in P \cap G_R^0$.

Тогда

(1) *если K не есть $\boldsymbol{R}$ или $\boldsymbol{C}$, т.е. K вполне несвязно, то ω есть отображение в точку (точнее, из G_R^0 можно выбросить такое множество A меры нуль, что $\omega(x) = \omega(y)$ для любых $x, y \in G_R^0 - A$),*

(2) *если K есть $\boldsymbol{R}$ или $\boldsymbol{C}$, то отображение ω рационально (точнее, совпадает почти всюду с ограничением на G_R^0 некоторого рационального отображения группы G).*

Замечание. С Теоремой 3 связаны теоремы Сельберга-Вейля о локальной жесткости [**44**], [**50**] и теорема Мостова о строгой жесткости [**35**], [**37**], [**38**]. О представлениях арифметических подгрупп групп ненулевого $\boldsymbol{Q}$-ранга см. [**21**], [**41**], [**47**]. Прасад [**39**] применил метод Мостова для доказательства строгой жесткости неравномерных решеток $\boldsymbol{Q}$-ранга 1, а также доказал [**40**] теорему о строгой жесткости для p-адических групп. Мостов доказал теорему о строгой жесткости, основываясь на своих замечательных идеях о продолжении изоморфизмов дискретных подгрупп на компактификации симметрических пространств. Докладчик находился под сильным влиянием этих идей, и общий план доказательства Теоремы 3 похож на общий план доказательства теоремы Мостова. С другой стороны, отметим, что Мостов рассматривал только *непрерывные* эквивариантные отображения, а его метод построения отображений существенно отличается от нашего и является чисто геометрическим.

2. Пусть S—некоторое множество простых чисел, к которому присоединена ∞. Если $p = \infty$, то положим $\boldsymbol{Q}_p = \boldsymbol{R}$. Положим $\boldsymbol{Q}(S) = \{q \in \boldsymbol{Q};\ v_p(q) \leqq 1$ для любого простого $p \notin S\}$, где v_p—(мультипликативное) p-адическое нормирование. Пусть G_p, $p \in S$ есть односвязная полупростая алгебраическая $\boldsymbol{Q}_p$-группа, причем G_p ($\boldsymbol{Q}_p$) не имеет бесконечных компактных факторгрупп. Положим $G = \prod_{p \in S} G_p$, $G_S = \prod_{p \in S} G_p(\boldsymbol{Q}_p)$. Рангом группы G_S назовем сумму $\sum_{p \in S} \operatorname{rank}_{\boldsymbol{Q}_p} G_p$ $\boldsymbol{Q}_p$-рангов групп G_p. В приводимых ниже Теоремах 7, 8 и 9 мы предполагаем, что множество S конечно, а Γ—неприводимая решетка в G_S.

Теорема 7. *Если ранг группы G_S больше 1, то Γ является S-арифметической подгруппой в G, т.е. найдутся алгебраическая $\boldsymbol{Q}$-группа H, замкнутые подгруппы H_1 и H_2 локально компактной группы $H_S = \prod_{p \in S} H(\boldsymbol{Q}_p)$, открытая компактная подгруппа F группы H_2 и непрерывный изоморфизм f: $G_S \to H_1$ такие, что $H_S = H_1 \times H_2$, а подгруппы $d((H_1 \times F) \cap H(\boldsymbol{Q}(S)))$ и $f(\Gamma)$ соизмеримы, где d:$H_S \to H_1$—естественный эпиморфизм.*

Для некоторых G Теорема 7 высказывалась (правда, в несколько неточной форме) в качестве гипотезы Пятецким-Шапиро [**17**].

Теорема 8. *Пусть T—конечномерное представление группы Γ над полем k. Предположим, что ранг группы G_S больше* 1. *Тогда*

(1) *если представление T неприводимо, а алгебраическая группа $\overline{T(\Gamma)}$ связна и проста, то T псевдорационально,*

(2) *если* $\operatorname{char} k \neq 0$, *то T вырождено.*

(3) *если* $\operatorname{char} k = 0$, *то T вполне приводимо.*

Теорема 9. *Обозначим через $\tilde{\Gamma}$ группу Γ-рациональных элементов в G_S, т.е. $\tilde{\Gamma} = \{g \in G_S;\ g\Gamma g^{-1}$ соизмерима с $\Gamma\}$. Тогда, если $\tilde{\Gamma}$ всюду плотна в G_S, то Γ является S-арифметической подгруппой в G.*

Теорема 9 в случае, когда G—вещественная группа, была высказана в качестве гипотезы в работе Пятецкого-Шапиро и Шафаревича [18] и ранее доказана Кажданом для неравномерных решеток в $G = SL(2, \boldsymbol{R})$, состоящих из матриц, элементы которых являются целыми алгебраическими числами [7]. Теоремы 7 и 8 являются обобщениями Теорем 1 и 2. Метод доказательства Теорем 7, 8 и 9 близок к методу доказательства Теорем 1 и 2.

3. Подгруппу Γ алгебраической k-группы H, $\operatorname{char} k = 0$ назовем *допустимой*, если любое ее представление над полем ненулевой характеристики вырождено, а любое представление над полем нулевой характеристики разлагается в прямую сумму абсолютно неприводимых представлений, каждое из которых является тензорным произведением вырожденного представления и (нескольких) псевдорациональных представлений. Из Теоремы 8 нетрудно вывести

Теорема 10. *Пусть G, G_S и Γ те же, что и в Теоремах* 7—9. *Тогда, если ранг группы G_S больше* 1, *то Γ является допустимой подгруппой группы G.*

Из Теоремы 10, используя теорему Бореля о приведении [23], сильную аппроксимационную теорему и теорему Платонова о (почти) простоте группы p-адических точек односвязной некомпактной (почти) простой $\boldsymbol{Q}_p$-группы [15], нетрудно вывести

Теорема 11. *Пусть k—конечное расширение поля $\boldsymbol{Q}$, l—подкольцо в k, G—односвязная полупростая алгебраическая k-группа, S—некоторое множество нормализованных нормирований поля k, содержащее все неархимедовы нормализованные нормирования. Через k_s обозначим пополнение поля k относительно нормирования s. Положим $k(S) = \{q \in k;\ s(k) \leqq 1$ для любого $s \notin S\}$. Тогда, если $\sum_{s \in S} \operatorname{rank}_{k_s} G \geqq 2$ и $l \supset k(S)$, то любая подгруппа конечного индекса группы G_l является допустимой подгруппой группы G.*

Так как G_{k_s} изотропны для всех нормализованных нормирований s, кроме конечного числа, то из Теоремы 11 вытекает

Теорема 12. *Пусть k и G те же, что и в Теореме* 11. *Тогда G_k является допустимой подгруппой группы G.*

Теоремы 11 и 12 связаны с результатами об абстрактных изоморфизмах и

гомоморфизмах подгрупп алгебраических групп [26], [31], [32], [48] (о более ранних работах см. Дьедонне [29]). Методы, которыми получаются эти результаты, являются, в отличие от изложенного здесь, чисто алгебраическими и основаны, грубо говоря, на изучении "геометрии" группы G_k.

2. Некоторые нерешенные задачи. 1. В [42, следствие 6.13] алгебраическими методами доказано, что любая конечно порожденная группа движений симметрического пространства X содержит подгруппу конечного индекса, не имеющую элементов конечного порядка. Верна ли аналогичная теорема, если X—произвольное полное риманово многообразие неположительной кривизны (не обязательно симметрическое)?

2. В связи с Теоремой 1 и ранее упоминавшимися примерами Макарова и Винберга возникает вопрос, ответ на который совершенно неясен. В каких вещественных полупростых группах $\boldsymbol{R}$-ранга 1 есть неарифметические решетки?

3. Пусть G и Γ те же, что и в Теореме 1. Верно ли, что всякий нетривиальный нормальный делитель группы Γ имеет конечный индекс в Γ? Васерштейн [2] доказал, что ответ утвердителен для широкого класса арифметических подгрупп классических групп положительного $\boldsymbol{Q}$-ранга. Укажем также на уже упоминавшиеся результаты о конечности фактора по коммутанту.

4. Всегда ли фундаментальная группа компактного многообразия неположительной кривизны имеет точное конечномерное представление? Более обще, всякая ли группа движений полного односвязного многообразия неположительной кривизны имеет точное конечномерное представление?

5. Может ли группа автоморфизмов свободной группы быть группой движений полного односвязного многообразия неположительной кривизны? Тот же вопрос для группы автоморфизмов фундаментальной группы компактной римановой поверхности и для группы кос Артина. Если ответ отрицателен, то возможно, что здесь будут полезны результаты работ Лоусона и Яу [34], [51].

6. Если $G = SL(2, \boldsymbol{R})$, а Γ_1 и Γ_2—решетки в G, то, как хорошо известно, отношение объемов $v(G/\Gamma_1)/v(G/\Gamma_2)$ рационально. Что можно утверждать для произвольной полупростой группы Ли G? Уже в случае, когда $G = SL(3, \boldsymbol{R})$, возникают интересные теоретикочисловые вопросы.

7. Скажем, что риманово многообразие M *не имеет плоских слагаемых*, если M нельзя разложить в прямую сумму двух римановых многообразий, одно из которых изометрично евклидову пространству положительной размерности. Пусть n—натуральное число, а $0 < K' < K < \infty$. Обозначим через $A(n, K)$ и $A(n, K, K')$ классы всех полных односвязных многообразий неположительной кривизны, не имеющих плоских слагаемых, у которых кривизна k по любому двумерному направлению удовлетворяет соответственно следующим неравенствам: $-K \leqq k \leqq 0$, $-K \leqq k \leqq -K'$. Через $B(n, K)$ и $B(n, K, K')$ обозначим классы всех факторпространств вида X/Λ, где X принадлежит соответственно $A(n, K)$ и $A(n, K, K')$, а Λ—дискретная группа движений

многообразия X, а через $C(n, K) \subset B(n, K)$ и $C(n, K, K') \subset B(n, K, K')$—классы всех многообразий неположительной кривизны, у которых универсальная накрывающая принадлежит соответственно $A(n, K)$ и $A(n, K, K')$. Выскажем следующие предположения:

(а) Найдется такое $p(n, K) > 0$, что, если $X/\Lambda \in B(n, K)$, то объем $v(X/\Lambda)$ больше $p(n, K)$, и, в частности, если $M \in C(n, K)$, то объем $v(M)$ больше $p(n, K)$;

(б) Для любого $V > 0$ найдется конечное число таких групп $\Lambda_1, \cdots, \Lambda_m$, что, если $X/\Lambda \in B(n, K)$ и $v(X/\Lambda) < V$, то Λ изоморфно одной из групп Λ_i, $1 \leqq i \leqq m$. В частности, для любого $V > 0$ найдется только конечное число гомотопически неэквивалентных многообразий класса $C(n, K)$, объем которых не превосходит V;

(в) Если $X/\Lambda \in B(n, K)$, то для дискретной группы Λ существует фундаментальная область с конечным числом сторон. В частности, если $X/\Lambda \in B(n, K)$, то группа Λ конечно определена, т.е. имеет конечное число образующих и конечное число соотношений. В частности, если $M \in C(n, K)$, то фундаментальная группа $\pi_1(M)$ конечно определена.

Докладчик умеет доказывать эти предположения для класса $B(n, K, K')$.

3. Эквивариантные измеримые отображения. Остановимся теперь более подробно на доказательстве Теорем 5 и 6. Нам понадобится одно следствие из мультипликативной эргодической теоремы и один результат о свободных дискретных подгруппах линейных групп, которые мы сейчас приведем (в п. 1 и 2). Но прежде зафиксируем $n \in \mathbf{N}^+$ и локальное поле K, снабженное абсолютной величиной ϕ и положим $\|x\| = \max_{1 \leq i \leq n} \phi(x_i)$ для любого $x = (x_1, \cdots, x_n) \in K^n$ и $\|A\| = \sup_{x \in K^n, x \neq 0} \|Ax\|/\|x\|$ для любого $A \in \operatorname{End} K^n$.

1. Пусть X—пространство Лебега (см. [20]) с непрерывной нормированной мерой μ ($\mu(X) = 1$), т.е. X изоморфно (в смысле теории меры) с точностью до множеств меры нуль отрезку числовой прямой с обычной мерой Лебега. Предположим, что на X задан автоморфизм L, т.е. измеримое вместе с обратным взаимно однозначное (после выбрасывания множества меры нуль) преобразование, сохраняющее меру μ. Пусть на X задан измеримый *коцикл* $u(m, x)$, $m \in \mathbf{Z}$, $x \in X$, $u(m + k, x) = u(m, L^k x)u(k, x)$ относительно динамической системы $\{L^m\}$ со значениями в группе $GL(n, K)$. Коцикл $u(m, x)$ назовем *ограниченным*, если $\ln\|u(1, x)\| \in L^\infty(X, \mu)$ и *существенным*, если

$$\overline{\lim_{m \to +\infty}} \frac{1}{m} \int_X \ln\|u(m, x)\| d\mu > 0.$$

Предложение 1. *Предположим, что* $\det u(m, x) \equiv 1$, *автоморфизм* L *эргодичен, а коцикл* $u(m, x)$ *является ограниченным и существенным. Тогда найдется такое измеримое отображение* $\varphi: X \to \operatorname{Gr}(n, K)$, *что для почти всех (по мере* μ*)* $x \in X$: (1) *для любого ненулевого* $y \in K^n$ *существует предел*

$$\chi(y, x) = \lim_{m \to +\infty} \frac{1}{m} \ln\|u(m, x)y\|$$

причем $\chi(y, x) < 0$ *тогда и только тогда, когда* $y \in \varphi(x)$; (2) $0 < \dim \varphi(x) < n$.

При $K = \boldsymbol{R}$ Предложение 1 вытекает из мультипликативной эргодической теоремы (см. Теоремы 1, 2 и 4 работы Оселедеца [**14**]); близкие результаты (см. также работу Миллионщикова [**13**]). Анализ доказательства Теорем 1, 2 и 4 в [**14**] показывает, что аналоги этих теорем верны для всех локальных полей K, причем доказательства переносятся без существенных изменений.

2. Элемент g группы $GL(n, K)$ назовем *существенно некомпактным*, если абсолютная величина хотя бы одного из его собственных значений отлична от 1 (напомним, что ϕ единственным образом продолжается на любое конечное расширение поля K).

Предложение 2. *Предположим, что подгруппа* Λ *группы* $GL(n, K)$ *абсолютна неприводима (т.е. тождественное представление* $U : U(\lambda) = \lambda$ *абсолютно неприводимо) и что* Λ *не является относительно компактной в* K*-топологии. Тогда*

(1) Λ *содержит существенно некомпактный элемент*;

(2) *если* Λ *конечно порождена,* $\Lambda \subset SL(n, K)$ *и группа* $\bar{\Lambda}$ *связна, то* Λ *содержит неабелеву свободную дискретную (в* K*-топологии) подгруппу.*

Доказательство п. 1 основано на невырожденности формы следа на алгебре, порожденной подгруппой Λ (которая вытекает из теоремы Бернсайда) и аналогично доказательству Леммы 3 работы Мерзлякова [**12**]. В работе Титса (*Free subgroups in linear groups*, J. Algebra **20** (1972), 250–270) фактически доказано, что из п. 1 и неприводимости подгруппы Λ вытекает п. 2.

3. *Существование "хороших" эквивариантных измеримых отображений.* Приведем набросок доказательства Теоремы 5 в предположении, что решетка Γ равномерна. Зафиксируем компактную фундаментальную область $A \subset G^0_{\boldsymbol{R}}$ решетки Γ и максимальный расщепимый над $\boldsymbol{R}$ тор S группы G. Для любого $g \in G^0_{\boldsymbol{R}}$ найдется ровно один такой элемент $\gamma(g) \in \Gamma$, что $g \in A\,\gamma(g)$. Для любых $g, g_1 \in G^0_{\boldsymbol{R}}$ положим

$$D(g) = T(\gamma(g)) \quad \text{и} \quad B(g, g_1) = D(gg_1)\,(D(g_1))^{-1}.$$

Ясно, что $B(g, g_1\gamma) = B(g, g_1)$ для любого $\gamma \in \Gamma$. Поэтому $B(g, g_1)$ индуцирует $B(g, x)$, $x \in G^0_{\boldsymbol{R}}/\Gamma$. Положим

$$\alpha(g) = \int_X \ln \|B(g, x)\|\, d\tilde{\mu}(x),$$

где $\tilde{\mu}$—мера, индуцированная на X мерой Хаара μ на $G^0_{\boldsymbol{R}}$. Ясно, что $B(g, x)$ —коцикл, т.е. $B(g_1 g_2, x) = B(g_1, g_2 x)\, B(g_2, x)$. Поэтому функция $\alpha(g)$ *полуаддитивна*, т.е. $\alpha(g_1 g_2) \leqq \alpha(g_1) + \alpha(g_2)$. Предположим, что $s \in S^0_{\boldsymbol{R}}$, $s \neq e$ таков, что

$$\overline{\lim_{m \to +\infty}} \frac{1}{m} \alpha(s^m) > 0. \tag{1}$$

Тогда для $x \in X$ положим $L(x) = sx$, $u(m, x) = B(s^m, x)$. Так как $B(g, x)$—коцикл, то $u(m, x)$—коцикл относительно динамической системы $\{L^m\}$, причем ограни-

ченный (так как A компактна) и существенный (ввиду (1)). С другой стороны автоморфизм L эргодичен (ввиду Леммы 4 в [36]), а $\det u(m,x) \equiv 1$ (так как $T(\Gamma) \subset SL(n, K)$). Применим теперь Предложение 1 и положим $\omega(g) = (D(g))^{-1}\varphi(\pi(g))$, где $\pi\colon G_R^0 \to G_R^0/\Gamma$—естественная проекция. Ясно, что $0 < \dim \omega(g) < n$, $\omega(g\gamma) = T(\gamma^{-1})\omega(g)$. С другой стороны, из свойства 1 отображения φ вытекает, что

$$\omega(g) = \left\{ y \in K^n;\ \lim_{m \to +\infty} \frac{1}{m} \ln \left\| D(s^m g) y \right\| < 0 \right\}. \tag{2}$$

Из (2) и компактности области A легко вытекает, что $\omega(pg) = \omega(g)$ для любого $p \in P_s$, где $P_s = \{g \in G_R^0;$ множество $\{s^m g s^{-m}\}$, $m \in N^+$, относительно компактно в $G_R^0\}$. Но, как нетрудно показать, найдется такая параболическая R-подгруппа P, что $P_s \supset P \cap G_R^0$. Таким образом, остается доказать существование $s \in S_R^0$, для которого выполняется (1). Так как функция α ограничена на любом компакте (потому что A компактна) и полуаддитивна, то, ввиду разложения Картана, достаточно показать, что

$$\overline{\lim_{g \in G_R^0,\ \|g\| \to \infty}} \frac{\alpha(g)}{\ln \|g\|} > 0. \tag{3}$$

Доказательство неравенства (3) проводится методами теории бесконечномерных представлений. Положим $H = SL(n, K)$. Пусть E—левое регулярное представление группы H в пространстве $L^2(H, \mu')$, где μ'—мера Хаара на H, т.е. $E_h f(h') = f(h^{-1}h')$. Положим $F(\gamma) = E(T(\gamma))$ и пусть V—унитарное представление группы G_R^0, индуцированное в смысле Макки представлением F дискретной подгруппы Γ (пространство J представления V состоит из таких функций j на G_R^0 со значениями в $L^2(H, \mu')$, что $j(g\gamma) = F(\gamma^{-1})j(g)$ и $\|j\| \in L^2(A, \mu)$), а G_R^0 действует на J левыми сдвигами $(V_g j)(g_1) = j(g^{-1}g_1)$. Далее, так как Γ равномерна, то Γ конечно порождена. Поэтому, используя п. 2 Предложения 2, получаем, что $T(\Gamma)$ содержит неабелеву свободную дискретную подгруппу Δ. Так как Δ дискретна, то ограничение на Δ представления E разлагается в прямой интеграл представлений, эквивалентных регулярному представлению группы Δ. Но [28], [49] регулярное представление неабелевой свободной группы изолировано от единичного. Поэтому представление F изолировано от единичного. Но V индуцировано представлением F, а Γ равномерна. Поэтому представление V изолировано от единичного (этот факт в случае, когда G не имеет факторов R-ранга 1, непосредственно вытекает из результатов о свойстве T Каждана [3], [6], [28], [49]). Поэтому на G_R^0 найдется такая неотрицательная финитная непрерывная функция f, что

$$\int_{G_R^0} f\, d\mu = 1, \quad \text{но } \|V(f)\| < 1, \text{ где } V(f) = \int_{G_R^0} V_g f(g)\, d\mu(g).$$

Пусть $d > 0$ таково, что функция $b(h) = \|h\|^{-d}$, $h \in H$ принадлежит $L^2(H, \mu')$. Положим $\theta(g) = F((\gamma(g))^{-1})b$. Ясно, что $\theta \in J$. Так как (1) f финитна, (2) $(V(f))^m = V(f_m)$, где $f_1 = f$, $f_m = f_{m-1} * f$ и $*$ обозначает свертку, (3) скалярное произведение $(V(g)\theta, \theta) > 0$ для любого $g \in G_R^0$, (4) $\|V(f)\| < 1$, (5) $\int_{G_R^0} f\, d\mu = 1$, то

$$\overline{\lim_{g\in G_R^0,\ \|g\|\to\infty}} - \frac{\ln(V(g)\theta,\theta)}{\ln\|g\|} > 0. \tag{4}$$

Но простые вычисления показывают, что

$$(V(g)\theta,\theta) \geqq (b,b)e^{-d\alpha(g^{-1})}, \qquad g\in G_R^0. \tag{5}$$

Из (4) и (5) вытекает (3).

Замечание. Для неравномерных Γ предварительно доказываются следующие два утверждения:

(1) Γ порождается конечным подмножеством M, и, кроме того, найдется такая фундаментальная область A подгруппы Γ и такой компакт $N\subset G_R^0$, что $a\gamma\in Na$ для любых $a\in A$, $\gamma\in M$,

(2) для Γ найдется такая фундаментальная область A, что коцикл $B(g,x)$ удовлетворяет следующему условию

$$f_L(x) = \sup_{g\in L}\ln\|B(g,x)\| \in L^\infty(X,\mu)$$

для любого компакта $L\subset G_R^0$.

Доказательство утверждений (1) и (2) опирается на конструкции фундаментальных областей для арифметических подгрупп [**24**], [**25**], [**30**] и решеток $\boldsymbol{Q}$-ранга 1 [**42**, Глава XIII], а также на результаты работы [**10**] (или [**43**]) (при этом вначале доказывается, что из условия (3) в формулировке Теоремы 5 вытекает условие (2)).

4. *Рациональность эквивариантных измеримых отображений.* Дадим набросок доказательства п. 2 Теоремы 6 (п. 1 доказывается аналогично и даже несколько проще). Пусть P'—противоположная (см. [**26**]) к P параболическая $\boldsymbol{R}$-подгруппа, U'—унипотентный радикал группы P', S—максимальный расщепимый над $\boldsymbol{R}$ тор группы G, лежащий в $P\cap P'$. Пусть $s\in S_R^0$, $s\neq e$, $C=C(s)$—централизатор s в G, а $C_R^u = C(s)_R^u\subset C_R^0$—множество унипотентных элементов в C_R. Обозначим через Σ пространство измеримых отображений группы Ли C_R^0 в M_K с топологией сходимости по мере на каждом компактном подмножестве (мы не различаем отображения, отличающиеся на множестве меры нуль). Группа F_K действует на Σ: $(f\sigma)(c) = f(\sigma(c))$, $f\in F_K$, $\sigma\in\Sigma$, $c\in C_R^0$ и пусть $\tilde\Sigma$—пространство орбит этого действия с индуцированной с Σ топологией. Для любого элемента $\sigma\in\Sigma$ его стационарная подгруппа F_σ является алгебраической (здесь надо использовать то, что любая строго возрастающая последовательность алгебраических подгрупп группы F счетна). Размерность орбиты $F(\sigma)$, $\sigma\in\Sigma$ определим как разность $\dim F - \dim F_\sigma$, где dim понимается в смысле теории алгебраических многообразий. Используя лемму о замкнутых орбитах [**52**, 1.8], нетрудно доказать, что для любого $\sigma\in\Sigma$ граница орбиты $F(\sigma)$ является объединением орбит строго меньшей размерности. Поэтому пространство $\tilde\Sigma$ полуотделимо. Для любого $g\in G_R^0$ положим

$$\omega_g(c) = \omega(cg), \qquad c\in C_R^0,$$

и обозначим через $\tilde\omega(g)\in\tilde\Sigma$ ту орбиту, на которой лежит $\omega_g\in\Sigma$. Так как C_R^0 коммутирует с s, а $s\in P\cap G_R^0$, то из свойств 1 и 2 отображения ω вытекает,

что для почти всех $g \in G_R^0$ и любых $\gamma \in \Gamma$, $n \in \mathbf{Z}$

$$\tilde{\omega}(s^n g\gamma) = \tilde{\omega}(g). \tag{6}$$

Отображение ω, а, следовательно, и $\tilde{\omega}$ измеримы. С другой стороны [**36**, Лемма 4], s эргодически действует на G_R^0/Γ. Поэтому из (6) вытекает, что для любого открытого подмножества Y в $\tilde{\Sigma}$ либо $\mu(\tilde{\omega}^{-1}(Y)) = 0$, либо $\mu(G_R^0 - \tilde{\omega}^{-1}(Y)) = 0$. Но Σ, а, следовательно, и $\tilde{\Sigma}$ имеют счетную базу открытых множеств. Поэтому в G_R^0 найдется такое подмножество L меры нуль, что для любых g_1, $g_2 \in G_R^0 - L$ любая окрестность точки $\tilde{\omega}(g_1)$ содержит $\tilde{\omega}(g_2)$. Отсюда и из полуотделимости пространства $\tilde{\Sigma}$ вытекает, что $\tilde{\omega}(G_R^0 - L)$ есть точка. Поэтому для любого $c \in C_R^0$ для почти всех $g \in G_R^0$ найдется такое $h(g, c) \in F_K$, что

$$\omega_g(c_1 c) = \omega_{cg}(c_1) = h(g, c)\omega_g(c_1) \tag{7}$$

для почти всех $c_1 \in C_R^0$. Пусть $F_g \subset F_K$—стационарная подгруппа элемента $\omega_g \in \Sigma$, N_g—нормализатор подгруппы F_g в F_K, а $r_g : N_g \to N_g/F_g$—естественный эпиморфизм. Ясно, что $h(g, c) \in N_g$. Положим $\alpha_g(c) = r_g(h(g,c))$. Ясно, что α_g: $C_R^0 \to N_g/F_g$ является гомоморфизмом, причем, как можно показать, непрерывным. Но ограничение на C_R^u любого непрерывного гомоморфизма группы C_R^0 рационально. Поэтому, для почти всех $g \in G_R^0$, $\omega_g(c_1 c)$ рационально по $c \in C_R^u$ для почти всех $c_1 \in C_R^0$. Но $c_1 C_R^u = C_R^u c_1$ для любого $c_1 \in C_R^0$. Поэтому, для почти всех $g \in G_R^0$, $\omega_g(cc_1) = \omega_{c,g}(c)$ рационально по $c \in C_R^u$ для почти всех $c_1 \in C_R^0$, и, следовательно, верна

Лемма 4. *Для любого $s \in S_R^0$, $s \neq e$ верно, что $\omega(cg)$ рационально по $c \in C(s)_R^u$ для почти всех $g \in G_R^0$.*

Так как $\boldsymbol{R}$-ранг группы G больше 1, то в U' найдется такой нормальный ряд алгебраических $\boldsymbol{R}$-подгрупп $U' = N_k \supset N_{k-1} \supset \cdots \supset N_1 = \{e\}$, что N_i есть полупрямое произведение U_i на N_{i-1}, где $U_i \subset C(s_i)$, $s_i \in S_R^0$, $s_i \neq e$. Так как U' унипотентна, то для любого i, $1 \leqq i \leqq k$ найдется такое бирегулярное отображение $\lambda_i : (N_i)_R \to \boldsymbol{R}^{q_i + t_i}$, $q_i = \dim N_{i-1}$, $t_i = \dim U_i$ при котором левые смежные классы по $(U_i)_R$ переходят в подмножестве вида (x, R^{t_i}), $x \in \boldsymbol{R}^{q_i}$, а смежные классы по $(N_{i-1})_R$ — в подмножестве вида $(\boldsymbol{R}^{q_i}, y)$, $y \in \boldsymbol{R}^{t_i}$. Но, как нетрудно показать, любая измеримая функция от (x, y), $x \in \boldsymbol{R}^{q_i}$, $y \in \boldsymbol{R}^{t_i}$, которая рациональна по x для почти всех y и рациональна по y для почти всех x, рациональна. Поэтому, используя Лемму 4, индукцией по i получаем, что $\omega(ug)$ рационально по $u \in N_i \cap G_R^0$ для почти всех $g \in G_R^0$ и любого i, $1 \leqq i \leqq k$. Отсюда и из свойства 2 отображения ω вытекает, что $\omega(upg)$ рационально зависят от $(u,p) \in (U' \cap G_R^0) \times (P \cap G_R^0)$ для почти всех $g \in G_R^0$. Но [**26**, 4.10] отображение

$$\alpha : U' \times P \to G, \qquad \alpha(u, p) = up, \qquad u \in U', p \in P,$$

является бирегулярным отображением на открытое по Зарисскому подмножество в G. Поэтому отображение ω рационально.

Литература

1. И.Н. Бернштейн и Д.А. Каждан, *Об одномерных когомологиях дискретных подгрупп*, Функц. анализ **4** (1970), 1–5.

2. Л.Н. Васерштейн, *О группах, обладающих свойством T*, Функц. анализ **2** (1968), 86.

3. ———, *Строение классических арифметических групп ранга, большего* 1, Мат. сборник **91** (1973), 445–470.

4. Э.Б. Винберг, *Дискретные группы, порожденные отражениями, в пространстве Лобачевского*, Мат. сборник **72** (1967), 471–488.

5. ———, *Кольца определения плотных подгрупп полупростых линейных групп*, Известия АН СССР, сер. матем. **35** (1971), 46–55.

6. Д.А. Каждан, *О связи дуального пространства группы со строением ее замкнутых подгрупп*, Функц. анализ **1** (1967), 71–74.

7. ———, *Построение Γ-рациональных групп для некоторых дискретных подгрупп Γ группы* $SL(2,R)$, Функц. анализ **2** (1968), 36–39.

8. В.С. Макаров, *Об одном классе дискретных групп пространства Лобачевского, имеющих бесконечную фундаментальную область конечной меры*, Доклады АН СССР **167** (1966), 30–33.

9. Г.А. Маргулис, *Арифметичность неравномерных решеток*, Функц. анализ **7** (1973), 87–88.

10. ———, *Арифметические свойства дискретных подгрупп*, Успехи матем. наук **29** (1974), 49–98.

11. ———, *Арифметичность и конечномерные представления равномерных решеток*, Функц. анализ **8** (1974), 78–79.

12. Ю.И. Мерзляков, *О линейных группах с ограниченными циклическими подгруппами*, Сиб. матем. журн. **7** (1966), 318–322.

13. В.М. Миллионщиков, *Критерий устойчивости вероятностного спектра линейных систем дифференциальных уравнений с рекуррентными коэффициентами и критерий почти приводимости систем с почти периодическими коэффициентами*, Мат. сборник **78** (1969), 179–201.

14. В.И. Оселедец, *Мультипликативная эргодическая теорема. Характеристические показатели Ляпунова динамических систем*, Труды Москов. Мат. Общ. **19** (1969), 179–210.

15. В.П. Платонов, *Проблема сильной аппроксимации и гипотеза Кнезера-Титса для алгебраических групп*, Известия АН СССР сер. матем. **33** (1969), 1211–1220; дополнение Известия АН СССР **34** (1970), 775–777.

16. И.И. Пятецкий-Шапиро, *Дискретные группы аналитических автоморфизмов полицилиндра и автоморфные формы*, Доклады АН СССР **124** (1959), 760–763.

17. ———, *Автоморфные функции и арифметические группы*, Труды Межд. конгр. математиков, "Мир", Москва, 1968, стр. 232–247.

18. ———, *Дискретные подгруппы групп Ли*, Труды Москов. Мат. Общ. **18** (1968), 3–18.

19. И.И. Пятецкий-Шапиро и И.Р. Шафаревич, *Теория Галуа трансцендентных расширений и униформизация*, Известия АН СССР, сер. матем. **30** (1966), 671–704.

20. В.А. Рохлин, *Об основных понятиях теории меры*, Мат. сборник **25** (1949), 107–150.

21. H. Bass, J. Milnor and J.-P. Serre, *Solution of the congruence subgroup problem for* $SL_n (n \geq 3)$ *and* $Sp_{2n} (n \geq 2)$, I.H.E.S. Publ. Math. **33** (1967), 59–137.

22. A. Borel, *Density properties of certain subgroups of semisimple groups*, Ann. of Math. (2) **72** (1960), 179–188.

23. ———, *Some properties of adele groups attached to algebraic groups*, Bull. Amer. Math. Soc. **67** (1961), 583–585.

24. ———, *Ensembles fondamentaux pour les groupes arithmétiques et formes automorphes*, Faculté de science de Paris, cours mimeographie, 1967.

25. A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), 485–535.

26. A. Borel and J. Tits, *Groupes reductifs*, I.H.E.S. Publ. Math. **27** (1965), 55–150.

27. ———, *Homomorphismes "abstraites" des groupes algébriques semi-simples*, Ann. of Math. (2) **97** (1973), 499–571.

28. C. Delaroche and A. Kirillov, *Sur les relations entre l'éspace d'un groupe et la structure de ses sous-groupes fermes*, Séminaire Bourbaki, Exposé 343, 1968.

29. J. Dieudonné, *La géométrie des groupes classiques*, Springer-Verlag, Berlin and New York, 1971.

30. R. Godément, *Domaines fondamentaux des groupes arithmétiques*, Séminaire Bourbaki, Exposé 257, 1962–1963.

31. A. J. Hahn, *The isomorphisms of certain subgroups of the isometric groups of reflexive spaces*, J. Algebra **27** (1973), 205–242.

32. A. A. Johnson, *The automorphisms of the finite unitary groups over infinite fields*, Amer. J. Math. **95** (1973), 87–107.

33. S. Kaneuki and T. Nagano, *Quadratic forms related to symmetric riemannian space*, Osaka Math. J. **14** (1962), 241–252.

34. H. B. Lawson and S. T. Yau, *Compact manifolds of nonpositive curvature*, J. Differential Geometry **7** (1972), 211–228.

35. G. D. Mostow, *Quasi-conformal mappings in n-space and the rigidity of hyperbolic space forms*, I.H.E.S. Publ. Math. **34** (1968), 53–104.

36. ———, *Intersections of discrete subgroups with Cartan subgroups*, J. Indian Math. Soc. **34** (1970), 203–214.

37. ———, *The rigidity of locally symmetric spaces*, Proc. Internat. Congr. Math. (Nice), vol. 2, 1970, pp. 187–197.

38. ———, *Strong rigidity of locally symmetric spaces*, Ann. of Math. Studies, No. 78.

39. G. Prasad, *Strong rigidity of Q-rank 1 lattices*, Invent. Math. **21** (1973), 255–286.

40. ———, *Lattices in p-adic semi-simple groups* (preprint).

41. M. S. Raghunathan, *Cohomology of arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **86** (1967), 409–424; **87** (1968), 279–304.

42. ———, *Discrete subgroups of Lie groups*, Springer-Verlag, Berlin and New York, 1972.

43. ———, *Discrete groups and Q-structures on semisimple Lie groups* (preprint).

44. A. Selberg, *On discontinuous groups in higher dimensional symmetric spaces*, Contributions to Function Theory, Tata Institute of Fundamental Research, Bombay, 1960.

45. ———, *Discontinuous groups and harmonic analysis*, Proc. Internat. Congr. Math. (Uppsala), 1963, pp. 177–189.

46. ———, *Recent developments in the theory of discontinuous groups of motions of symmetric spaces*, Proc. 15th Scand. Congr., Oslo 1968; Lecture Notes in Math., vol. 118, 1970, pp. 99–120.

47. J.-P. Serre, *Le probleme des groupes de congruence pour* SL_2, Ann. of Math. (2) **92** (1970), 489–527.

48. R. Solazzi, *On the isomorphism between certain congruence subgroups*, Canad. J. Math. **25** (1973), 1006–1014.

49. S. P. Wang, *The dual space of semisimple Lie groups*, Amer. J. Math. **91** (1962), 921–937.

50. A. Weil, *On discrete subgroups of Lie groups.* I, II, Ann. of Math. (2) **72** (1960), 369–384; **75** (1962), 578–602.

51. S. T. Yau, *On the fundamental group of compact manifolds of non-positive curvature*, Ann. of Math. (2) **93** (1971), 573–585.

52. A. Borel, *Linear algebraic groups*, Benjamin, New York, 1969.

Институт Проблем Передачи Информации АН СССР
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Foliations and Local Homology of Groups of Diffeomorphisms

John N. Mather

Haefliger has defined a classifying space $B\Gamma_q^r$ for codimension q foliations of class C^r in [4], [5]. As a result of the classifying theorems of Haefliger [4], [5] and Thurston [13], [14] it is known that the homotopy and homology groups of $B\Gamma_q^r$ are of considerable geometric interest.

There is a "normal mapping" ν: $B\Gamma_q^r \to BO_q$ for $r \geqq 1$ and $B\Gamma_q^0 \to B\mathrm{Top}_q$ for $r = 0$. We will denote the homotopy theoretic fiber of this mapping by $F\Gamma_q^r$. Obviously the homotopy and homology groups of $B\Gamma_q^r$ are related to those of $F\Gamma_q^r$, and the latter groups are probably the more fundamental mathematical object.

The homotopy and homology of $F\Gamma_q^r$ are mostly unknown. However, some striking results have been obtained. It is an easy consequence of Bott's vanishing theorem [1] that the homology (or, equivalently, the homotopy) of $F\Gamma_q^r$ is infinitely generated [2] if $r, q \geqq 2$. Thurston has shown that there is a homomorphism of $H_{2q+1}(F\Gamma_q^r, \mathbf{Z})$ onto $\boldsymbol{R}$, if $r \geqq 2$. Several mathematicians have noticed a homomorphism of Gel'fand-Fuks cohomology into $H^*(F\Gamma_q^r, \boldsymbol{R})$ (cf. [3]). All of these results may be viewed as giving lower bounds for the homology of $F\Gamma_q^r$.

This article reports on some recent work which was motivated by the desire to find "upper bounds" for the homology and homotopy of $F\Gamma_q^r$. Before stating the main result, we mention the two principal consequences which have been obtained.

THEOREM 1. *$F\Gamma_q^0$ is contractible.*

THEOREM 2. *$F\Gamma_q^r$ is $(q+1)$-connected if $r \neq q+1$.*

The analogue of Theorem 1 is true for Lipschitz foliations. As a consequence, Thurston was able to show that every subbundle of the tangent bundle of a manifold is homotopic to an integrable subbundle. However, the foliation which the integrable subbundle defines is only a Lipschitz foliation. This contrasts with Bott's

vanishing theorem, which shows that there are subbundles which are not homotopic to integrable subbundles defining C^2 foliations.

Haefliger already showed [4] that $F\Gamma^r_q$ is q-connected if $r \geqq 1$. However, his proof showed that this is of no geometric interest. The first homotopy group which carries geometrically interesting information is $\pi_{q+1}(F\Gamma^r_q)$.

To state the main result, we need to define the *local homology* of a topological group. This is analogous to the Eilenberg-MacLane homology of a (discrete) group, the difference being that the homology is computed by means of chains supported in a small neighborhood of the identity.

DEFINITION. Let G be a topological group such that the identity has a neighborhood which is contractible in G to 1. Let G_δ be the underlying discrete group. Let $\bar{G}$ denote the homotopy theoretic fiber of the identity mapping $G_\delta \to G$. With its usual realization (paths in G, emanating from the identity, with the topology generated by the compact-open topology on the paths and the discrete topology on endpoints), $\bar{G}$ is a topological group. The *local homology* of G is the homology of $B\bar{G}$.

The local homology of G depends only on the algebraic and topological structure of G in a neighborhood of the identity. If G is a Lie group, $H_1(B\bar{G}) \approx \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$, where $\mathfrak{g}$ denotes the Lie algebra of G. However the other local homology groups of G are unknown, if G is not abelian.

Another way to describe the local homology of G is as $H_i(S(G)/G)$, where $S(G)$ denotes the singular complex of G and G acts on $S(G)$ on the left by pointwise multiplication.

Let $\mathscr{D}^r_q$ denote the group of C^r compactly supported diffeomorphisms of $\boldsymbol{R}^q$. If K is a compact subset of $\boldsymbol{R}^q$, let $\mathscr{D}^r_q(K)$ denote the subgroup of those diffeomorphisms having support in K. Topologize $\mathscr{D}^r_q(K)$ with the C^r topology, and let $\mathscr{D}^r_q = \operatorname{inj\,lim}_K \mathscr{D}^r_q(K)$ have the direct limit topology.

MAIN THEOREM. *There exists a mapping $B\bar{\mathscr{D}}^r_q \to \Omega^q F\Gamma^r_q$ which induces ismorphism in integer homology.*

This theorem is due to Thurston [12]. It generalizes to arbitrary q a theorem the author proved for $q = 1$ in [9]. This was announced in [8].

Since $\mathscr{D}^0_q$ is contractible (by the Alexander trick), and $B(\mathscr{D}^0_q)_\delta$ is acyclic [7], it follows that $B\bar{\mathscr{D}}^0_q$ is acyclic. Therefore, by the main theorem, $\Omega^q F\Gamma^r_q$ is acyclic. Since Haefliger already showed [4] that $F\Gamma^r_q$ is $(q-1)$-connected, we obtain Theorem 1. Similarly, Theorem 2 follows from the main theorem, and the study of commutators of diffeomorphisms made in [9], [10]. The case $r = \infty$ of the main theorem can also be derived from the main theorem, the Kolmogorov-Arnold-Moser implicit function theorems, and Thurston's "Hurewicz theorem", for the KAM theorems show that any diffeomorphism of T^n near the identity is a product of commutators near the identity [6]. This implies that the local H_1 of the group of C^∞ diffeomorphisms of T^n vanishes. Then Thurston's Hurewicz theorem [12] permits one to conclude the same result for the group of compactly supported C^∞ diffeomorphisms of $\boldsymbol{R}^n$.

References

1. R. Bott, *On a topological obstruction to integrability*, Proc. Internat. Congress Math. (Nice, 1970), vol. 1, Gauthier-Villars, Paris, 1971, pp. 27–36.

2. R. Bott and S. Heitsch, *A remark on the integral cohomology of $B\Gamma_q$*, Topology **11** (1972), 141–146. MR **45** #2738.

3. R. Bott and A. Haefliger, *On characteristic classes of Γ-foliations*, Bull. Amer. Math. Soc. **78** (1972), 1039–1044. MR **46** #6370.

4. A. Haefliger, *Feuilletages sur les variétés ouvertes*, Topology **9** (1970), 183–194. MR **41** #7709.

5. ———, *Homotopy and integrability*, Manifolds—Amsterdam 1970 (Proc. Nuffic Summer School), Lecture Notes in Math., vol. 197, Springer-Verlag, Berlin, 1971, pp. 133–163. MR **44** #2251.

6. M.R. Herman, *Simplicité du groupe des diffeomorphismes de classe C^∞, isotopes à l'identité, du tore de dimension n*, C.R. Acad. Sci. Paris Sér. A-B **273** (1971), A232–A234. MR **44** #4788.

7. J.N. Mather, *The vanishing of the homology of certain groups of homeomorphisms*, Topology **10** (1971), 297–298. MR **44** #5973.

8. ———, *On Haefliger's classifying space.* I, Bull. Amer. Math. Soc. **77** (1971), 1111–1115. MR **44** #1047.

9. ———, *Integrability in codimension* 1, Comment. Math. Helv. **48** (1973), 195–233.

10. ———, *Commutators of diffeomorphisms*, Comment. Math. Helv. **49** (1974), 512–528.

11. ———, *Commutators of diffeomorphisms.* II (1975) (to appear).

12. W. Thurston, *Foliations and groups of diffeomorphisms*, Bull. Amer. Math. Soc. **80** (1974), 304–307.

13. ———, *The theory of foliations in codimension* > 1, Comment. Math. Helv. **49** (1974), 214–231.

14. ———, *The theory of foliations in codimension* 1 (to appear).

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Riemannian Structures and Triangulations of Manifolds

V. K. Patodi

Let X be a closed C^∞-manifold of dimension N. Then two additional structures on X have been quite extensively studied. One is the Riemannian structure giving rise to Riemannian geometry and the other one is the triangulation of X giving rise to polyhedral or combinatorial topology.

Sometimes we come across a problem in which we have an invariant for the manifold which has a nice expression in terms of one of these structures and we want to express this invariant in terms of the other structure. Let us consider two examples. Our first example is related to Pontrjagin classes p_i, $1 \leqq i \leqq N/4$, of the manifold X. If we choose a Riemannian structure on X, then by the well-known Chern-Weil theory (see [2]) we can explicitly write down a closed $4i$-form representing the ith Pontrjagin class regarded as an element of $H^{4i}(X, \boldsymbol{R})$. It is an open problem to obtain a combinatorial formula for Pontrjagin classes.

We now come to our second example. Our second example refers to Reidemeister-Franz torsion. Let χ be a representation of $\pi_1(X)$ by orthogonal $n \times n$ matrices, n a fixed positive integer. Let ξ_χ be the flat vector bundle on X defined by this representation. Suppose that all the cohomology groups of X with coefficients in the vector bundle ξ_χ are zero. Then one can define an invariant known as Reidemeister-Franz torsion (see [3] and [4]) depending on X and χ. To define this torsion one first chooses a smooth triangulation of X and in terms of this triangulation one defines an invariant which one proves is not changed if we consider subdivisions of the triangulation. We do not have so far a definition for this torsion which does not make use of the choice of a triangulation. It is an interesting open problem to define this invariant in terms of a Riemannian structure.

There is a conjecture of Ray and Singer (see [4]) in this direction. They define an invariant which they call analytic torsion. The Ray-Singer analytic torsion is

defined in terms of the eigenvalues of the Laplace operator acting on differential forms with coefficients in the vector bundle ξ_χ. Ray and Singer have proved that this invariant is independent of the choice of the Riemannian metric and they conjecture that this analytic torsion is equal to the Reidemeister-Franz torsion.

By these examples we see that many times we come across an interesting problem which involves passing from one of these structures to the other one of these two structures. We will now describe a method which in a sense builds a bridge between these two structures and allows us a considerable interplay between them. The basic ideas are from Whitney's book [5] and Dodziuk's thesis [1].

Let g be a Riemannian metric on X and K a smooth triangulation of X. Let $T^*(X)$ be cotangent bundle of X and $\Lambda^q = \Lambda^q T^*(X)$, $0 \leqq q \leqq N$. Let $C_q(K)$, $0 \leqq q \leqq N$, be simplicial (oriented) real chain groups and $C^q(K)$ be the dual simplicial cochain groups.

Now in Riemannian geometry we consider de Rham complex:

$$0 \to C^\infty(\Lambda^0) \xrightarrow{d} C^\infty(\Lambda^1) \xrightarrow{d} \cdots \xrightarrow{d} C^\infty(\Lambda^N) \to 0.$$

And in simplicial topology we consider the complex:

$$0 \to C^0(K) \xrightarrow{\partial} C^1(K) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C^N(K) \to 0.$$

We have de Rham map R: $C^\infty(\Lambda^q) \to C^q(K)$, $0 \leqq q \leqq N$, from the de Rham complex to the cochain complex. This is given by integration.

Whitney in his book [5] has defined a map W which assigns to a cochain $A \in C^q(K)$, a differential form WA on X of type q. The map W is defined in the following way. Let $\sigma = (p_0, \cdots, p_q)$ be an oriented q-simplex, σ^* be dual simplex and Σ be an N-simplex. Then $W(\sigma^*)$ on $|\Sigma|$ is zero unless σ is a face of Σ. Also if σ is a face of Σ and $\Sigma = (p_0, \cdots, p_N)$, then any point $x \in |\Sigma|$ has barycentric coordinates μ_{p_0}, $\cdots$, μ_{p_N} and we define $W(\sigma)$ on $|\Sigma|$ as:

$$q! \sum_{i=0}^{q} (-1)^i \mu_{p_i} d\mu_{p_0} \cap d\mu_{p_1} \cap \cdots \cap d\mu_{p_{i-1}} \cap d\mu_{p_{i+1}} \cap \cdots \cap d\mu_{p_q}.$$

This defines W on dual simplices. We then extend W to the whole of $C^q(K)$ by linearity.

The form WA for $A \in C^q(K)$ is not a C^∞-differential form, but on any N-simplex Σ (N = dim of X) WA is C^∞, and if an r-simplex σ, $0 \leqq r \leqq N$, is a face of two N-simplices Σ, Σ', then

$$i^*(WA|_\Sigma) = (i')^*(WA|_{\Sigma'}),$$

where i, i' are inclusions of σ in Σ, Σ'.

Hence it makes sense to apply the de Rham map R to the form WA and we have

$$RWA = A, \tag{1}$$

$$dWA = W\partial A. \tag{2}$$

We can now use the Whitney map to pull the Riemannian inner product from the de Rham complex to the simplicial cochain complex. We thus obtain an inner product in cochain groups defined by

$$\langle C, C' \rangle =_{\text{def}} \langle WC, WC' \rangle, \qquad C, C' \in C^q(K), 0 \leqq q \leqq N.$$

Let $\partial^*: C^q(K) \to C^{q-1}(K)$ be the adjoint of ∂, with respect to this inner product and $\Delta_q = (\partial\partial^* + \partial^*\partial)$ be combinatorial Laplacian from $C^q(K)$ into itself, $0 \leqq q \leqq N$. Now let us consider subdivisions $S^n(K)$, $n \geqq 1$, of K. For any triangulation K of X we define mesh $\eta(K)$ of K by $\eta(K) = \text{Sup } d(x, y)$, where the supremum is taken over all points $x, y \in X$ for which there exists a simplex Σ such that $x, y \in |\Sigma|$; and d is the Riemannian distance function on $X \times X$ defined by the Riemannian metric.

Let us suppose that $\eta(S^nK) \to 0$ as $n \to \infty$ and subdivisions are reasonably nicely chosen so that there exists a constant $C > 0$ such that, for $n \geqq 1$, $d(\nu, \omega) > C\eta(S^n(K))$ for all vertices ν, ω, $\nu \neq \omega$, belonging to a simplex of $S^n(K)$.

Now for each n we have the de Rham map $R_n: C^\infty(\Lambda^q) \to C^q(S^nK)$, Whitney map W_n and inner product in $C^q(S^nK)$, defined by Whitney map W_n. With respect to this inner product we have the adjoint operator ∂_n^* and combinatorial Laplacians $\Delta_{q,n}$ from $C^q(S^nK)$ into itself.

The inner products induced by the Whitney map depend on the Riemannian metric g in $T^*(X)$ and a natural question now would be whether we can get back the Riemannian structure from these inner products in $C^q(S^nK)$, $n \geqq 1$. This is in fact true in a suitable sense and we now state these results.

The first approximation theorem, proved by Dodziuk (see [1]) is that for any C^∞ form f on X, W_nR_nf converges to f in L^2 as $n \to \infty$. Then by (2) it follows that dW_nR_nf $(= W_n\partial R_nf = W_nR_ndf)$ converges to df in L^2 as $n \to \infty$. It is however not true in general that $W_n\partial_n^*R_nf$ converges to δf in L^2 as $n \to \infty$. We nevertheless have approximation theorems for Hodge decompositions:

THEOREM 1. *Let* $f \in C^\infty(\Lambda^q)$ *and*

$$f = f_1 + df_2 + \delta f_3, \qquad R_nf = A_1 + \partial A_2 + \partial_n^* A_3$$

be Hodge decompositions (f_1 *is a harmonic form and* A_1 *is a harmonic cochain, i.e.,* $\partial A_1 = \partial_n^* A_1 = 0$). *Then,*

$$\| f_1 - W_nA_1 \| \leqq C\eta_n \sum_{i=0}^{k} \| \Delta^i f \|, \quad \| df_2 - W_n\partial A_2 \| \leqq C\eta_n \sum_{i=0}^{k} \| \Delta^i f \|,$$

and

$$\| \delta f_3 - W_n\partial_n^* A_3 \| \leqq C\eta_n \sum_{i=0}^{k} \| \Delta^i f \|$$

where C *and* k *are positive integers independent of* f, n *and* $\eta_n = \eta(S^nK)$ *is the mesh of* $S^n(K)$; $\| \ \|$ *is* L^2*-norm.*

THEOREM 2. *Let* $A \in C^q(S^nK)$ *and*

$$A = A_1 + \partial A_2 + \partial_n^* A_3, \quad W_nA = f_1 + df_2 + \delta f_3$$

be Hodge decompositions (W_nA, *though not smooth, is an* L^2*-form and therefore we still have its Hodge decomposition*). *Then*

$$\| W_nA_1 - f_1 \| \leqq C(\eta_n \log \eta_n)\, (\| W_nA \| + \| dW_nA \|),$$

$$\| W_n \partial A_2 - df_2 \| \leqq C(\eta_n \log \eta_n)(\| W_n A \| + \| dW_n A \|),$$

and

$$\| W_n \partial_n^* A_3 - \delta f_3 \| \leqq C(\eta_n \log \eta_n)(\| W_n A \| + \| dW_n A \|),$$

where C is a constant independent of n and A.

Theorem 1 was proved by Dodziuk [1].

Furthermore the eigenvalues of $\Delta_{n,q}$ converge to eigenvalues of smooth Laplacian Δ_q and we now proceed to state this result precisely.

Let $\Delta_q : C^\infty(\Lambda_q) \to C^\infty(\Lambda_q)$ be smooth Laplacian of Hodge theory, $0 \leqq q \leqq N$. Then it is known that Δ_q has an infinite sequence of eigenvalues $0 \leqq \lambda_{1,q} \leqq \lambda_{2,q} \leqq \cdots \leqq \lambda_{n,q} \leqq \cdots \uparrow \infty$, each eigenvalue being repeated as many times as its multiplicity indicates. Similarly combinatorial Laplacians $\Delta_{q,n}$ have eigenvalues (finitely many) $0 \leqq \lambda_{1,q}^n \leqq \lambda_{2,q}^n \leqq \cdots \leqq \lambda_{d(q,n),q}^n$, where $d(q, n)$ is the dimension of $C^q(S^n K)$. Our main theorem is

THEOREM 3. *Let j be any fixed positive integer. Then $\lambda_{j,q}^n$ converges to $\lambda_{j,q}$ as $n \to \infty$ (note that $\lambda_{j,q}^n$ is defined only when n is sufficiently large so that $d(q, n) \geqq j$). More precisely there exist a constant $C > 0$ and an integer $k > 0$ (independent of j and n) such that*

(a) $$\lambda_{j,q}\{1 - C\lambda_{j,q}^n(\eta_n \log \eta_n)^2\} \leqq \lambda_{j,q}^n \text{ and}$$

(b) $$\lambda_{j,q}^n\{1 - C(\lambda_{j,q})^k \eta_n\} \leqq \lambda_{j,q}$$

(provided that $\lambda_{j,q}^n$ is defined, i.e., $d(q, n) \geqq j$).

Since eigenvalues $\{\lambda_{n,q}\}$ are known to contain considerable geometric information of the Riemannian manifold (X, g), the above theorem means in a sense that by taking subdivisions $S^n(K)$ we get back a considerable amount of Riemannian structure.

As an immediate consequence of Theorem 3, one gets the following result, which may have more significance for the Ray-Singer conjecture about Reidemeister-Franz torsion. Let

$$\zeta_q(s) = \sum_{\lambda_{i,q} > 0} (\lambda_{i,q})^{-s}, \quad s \text{ a complex number;}$$

then the zeta function $\zeta_q(s)$ converges uniformly and absolutely for $\operatorname{Re} s > N/2$ ($N = \dim$ of X) and has an analytic continuation to the whole of the complex plane with some possible simple poles on the real line. Let

$$\zeta_{n,q}(s) = \sum_{\lambda_{i,q}^n > 0} (\lambda_{i,q}^n)^{-s}$$

be finite combinatorial zeta functions. Then we have

THEOREM 4. $\zeta_{n,q}(s) \to \zeta_q(s)$ *as* $n \to \infty$, *uniformly on compact subsets of the region defined by* $\operatorname{Re} s > N/2$.

This theorem was conjectured (for all q) by Dodziuk in [1] and was proved by him for the case $q = 0$. Theorem 3 follows quite easily by using Theorems 1 and

2 and defining eigenvalues of Δ_q, $\Delta_{q,n}$ in terms of maxima-minima principle applied to appropriate quadratic forms; the details will be published elsewhere.

Finally I would like to thank Professor I.M. Singer for bringing Dodziuk's thesis [1] to my attention.

References

1. Jozef Dodziuk, *Finite-difference approach to the Hodge theory of harmonic forms*, Thesis; Amer. J. Math. (to appear).

2. S. Kobayashi and K. Nomizu, *Foundations of differential geometry*. Vol. II, Interscience Tracts n Pure and Appl. Math., no. 15, Vol. II, Interscience, New York and London, 1969, Chapter XII. MR **38** #6501.

3. J.W. Milnor, *Whitehead torsion*, Bull. Amer. Math. Soc. **72** (1966), 358–426. MR **33** #4922.

4. D. B. Ray and I. M. Singer, *R-torsion and the Laplacian on Riemannian manifolds*, Advances n Math. **4** (1971), 145–210.

5. H. Whitney, *Geometric integration theory*, Princeton Univ. Press, Princeton, N. J., 1957. MR 19, 309.

TATA INSTITUTE OF FUNDAMENTAL RESEARCH
BOMBAY 400005, INDIA

Section 9

General Topology, Real and Functional Analysis

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Bases and Approximation by Splines

Z. Ciesielski

1. Introduction. Our main interest is in constructing a basis in the Banach space of functions of several complex variables analytic in the interior of a polydisc and continuously differentiable on its boundary a given number of times. The solution depends on results which are tied up with the history of spline bases of which we are giving a short account.

In the space $C(I)$, $I = \langle 0, 1\rangle$, there are essentially different bases: the interpolating Schauder basis and the orthonormal Franklin system. Most interesting are those two types of bases of splines of higher orders which are simultaneous in $C^m(I)$ [13]. The first simultaneous basis of splines of order 1 was exhibited in [3] (see also [11]), and it was used in [12] and [4] to construct a basis in $C^1(I^d)$. Extension to higher orders of orthogonal splines was carried out in [8] and [7]. Approximation properties of those bases are discussed in [5]. In the periodic case similar questions were considered in [10] and [6].

The periodic and nonperiodic simultaneous interpolating bases of splines are discussed in detail in [14], [13] and [13], [9], respectively.

Using the results for the periodic simultaneous spline orthonormal systems, which are stated in § 2, we were able to carry out the construction of § 3. It extends the result of [1] corresponding to the case of $m = 0$. The construction in [1] depends on the Franklin system, and the ideas contained there suggested the solution presented in this note.

2. Periodic spline orthonormal bases. Let us start with dyadic and periodic partition of the real line. For $n = 2^p + q$, $1 \leqq q \leqq 2^p$, $p \geqq 0$, let

$$\begin{aligned} s_{n,j} &= \pi(2j/2^{p+1} - 1) && \text{for } j = 0, \cdots, 2q, \\ &= \pi(2(j-q)/2^p - 1) && \text{for } j = 2q+1, \cdots, n; \end{aligned}$$

and $s_{n,i+kn} = s_{n,i} + 2\pi k$, $s_{n,k} = \pi(2k-1)$ for $k = 0, \pm 1, \cdots$.

The well-known B-splines of order $m \geqq -1$ corresponding to the partition $\{s_{n,i}\}$ are defined in terms of divided differences as follows:

$$N_{n,i}^{(m)}(t) = (s_{n,i+m+1} - s_{n,i-1})\,[s_{n,i-1}, \cdots, s_{n,i+m+1}; (s-t)_+^{m+1}].$$

In every interval the nontrivial B-splines are linearly independent and form a partition of unity. The corresponding 2π-periodic B-splines are given by the formula

$$T_{n,i}^{(m)}(t) = \sum_{k=0,\pm 1,\cdots} N_{n,i}^{(m)}(t + 2\pi k).$$

The set of 2π-periodic splines of order m corresponding to the nth partition $\{s_{n,i}\}$ is denoted by $S_n^m(T)$, $T = \langle -\pi, \pi\rangle$. Again, $\{T_{n,i},\ i = 0, \pm 1, \cdots\}$ forms a partition of unity and $\{T_{n,i},\ i = 1, \cdots, n\}$ is a basis in $S_n^m(T)$. In the space $C(T)$ the following scalar product is given

$$(f, g) = \int_T f(t)\overline{g(t)}dt. \tag{2.1}$$

Let $\pi_n^{(m)}$ denote the orthogonal w.r.t. (2.1) projection of $C(T)$ onto $S_n^m(T)$. Then

$$(\pi_n^{(m)} f)(t) = \int_T K_n^{(m)}(t, s) f(s)\, ds, \qquad f \in C(T),$$

where

$$K_n^{(m)} = \sum_{i,j=1}^{n} A_{n;i,j}^{(m)}\, T_{n,i}^{(m)} \otimes T_{n,j}^{(m)},$$

and the matrix $(A_{n;i,j}^{(m)})$ is the inverse to the Gram matrix corresponding to the functions $T_{n,i}^{(m)}$.

THEOREM 2.1 (J. DOMSTA [10]). *To each $m \geqq -1$ there are constants C_m and q_m, $0 < q_m < 1$, such that*

$$\left|A_{n;i,j}^{(m)}\right| < C_m n q_m^{d_n(i,j)}$$

where $d_n(i,j) = \min(|i-j|, n - |i-j|)$.

In what follows D denotes the usual differentiation operator and

$$(Hf)(t) = \int_t^{\pi} f(s)\, ds - \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\int_s^{\pi} f(u)\, du\right] ds.$$

Clearly, $DHf = -f$. Moreover, let

$$K_n^{(m;\,k,l)} = H^{(0,\,l)} D^{(k,\,0)} K_n^{(m)}.$$

THEOREM 2.2 (CF. [6]). *To each $m \geqq -1$ there are constants C_m and q_m, $0 < q_m < 1$, such that*

$$\left|\frac{1}{2\pi} + K_n^{(m;\,k,\,k)}(t, s)\right| < C_m n q_m^{n d(t,s)}$$

holds for $t, s \in T$, $0 \leqq k \leqq m+1$ with $d(t,s) = \min(|t-s|, 2\pi - |t-s|)$.

Since $S_j^m(T) \subset S_{j+1}^m(T)$ there is an ONC set of splines $F_j^{(m)} \in S_j^m(T)$ such that

$$K_n^{(m)} = \sum_{j=1}^{n} F_j^{(m)} \otimes F_j^{(m)}.$$

Clearly, $F_1^{(m)} = (2\pi)^{-1/2}$. Moreover, for each integer k, $|k| \leqq 1 + m$, we define a new set of splines

$$\begin{aligned} F_j^{(m,k)} &= D^k F_j^{(m)} \quad \text{for } k = 1, \cdots, m+1, \\ &= H^{-k} F_j^{(m)} \quad \text{for } k = -1, \cdots, -(m+1); \end{aligned}$$

and $F_j^{(m,0)} = F_j^{(m)}$ for $j = 1, \cdots$, $F_1^{(m,k)} = (2\pi)^{-1/2}$. The system $\{F_j^{(m,k)}, F_i^{(m,-k)}, i, j = 1, 2, \cdots\}$ for given k, $|k| \leqq m + 1$, is biorthogonal w.r.t. (2.1). Correspondingly, let

$$\pi_n^{(m,k)} f = \sum_{j=1}^{n} (f, F_j^{(m,-k)}) F_j^{(m,k)},$$

and let $\omega_1^{(p)}(f; 1/n)$ denote the $L_p(T)$ modulus of smoothness of order 1 corresponding to the increment $1/n$.

THEOREM 2.3 (CF. [6]). *For each k, $|k| \leqq m + 1$, $\{F_j^{(m,k)}, j = 1, \cdots\}$ is a basis in $L_p(T)$, $1 \leqq p < \infty$, and for $|k| \leqq m$ it is a basis in $C(T)$. Moreover, for $f \in L_p(T)$, $1 \leqq p \leqq \infty$ and $|k| \leqq m + 1$ we have*

$$\| f - \pi^{(m,k)} f \|_p \leqq C_m \omega_{m+2+k}^{(p)}(f; 1/n).$$

COROLLARY 2.1. *For each $m \geqq 0$, the system $\{F_j^{(m)}, j = 1, \cdots\}$ is a simultaneous basis in $C^m(T)$.*

COROLLARY 2.2. *A suitable d-fold tensor product of the basis $\{F_j^{(m)}, j = 1, \cdots\}$ is a simultaneous basis in $C^m(T^d)$.*

There are two more facts which play an essential role in the construction of §3. As in [7] we have

LEMMA 2.1. *Let $m \geqq -1$ and let $f = \sum_{j=1}^{n} a_j T_{n,j}^{(m)}$. Then, for some $C_m > 0$,*

$$C_m^{-1} \| f \|_p \leqq \left(\frac{1}{n} \sum_{j=1}^{n} |a_j|^p \right)^{1/p} \leqq C_m \| f \|_p.$$

LEMMA 2.2. *Let $m \geqq 0$. Then there is a constant C_m such that, for $f \in S_n^m(T)$,*

$$\left| \frac{f(t+s) - f(t-s)}{s} \right| \leqq C_m n^2 \int_{|t-u| < C_m/n} |f(u)| \, du$$

holds for $0 < s < 1/n$.

The last inequality was suggested by an inequality for polygonals derived by S. V. Botschkariev in [1].

3. Construction of a basis in $A^m(\Delta^d)$. We use the following notation: $\Delta = \{z \in \mathbf{C} : |z| \leqq 1\}$, $\Delta^d = \Delta_1 \times \cdots \times \Delta_d$, $\Delta_i = \Delta$. The set of all complex polynomials of d variables is denoted by $\mathbf{W}_d$. For given $m \geqq 0$ and $w \in \mathbf{W}_d$ we define the norm

$$\|w\|_d^{(m)} = \sum_{|\delta| \leq m} \|D^\delta w\|_d \tag{3.1}$$

where $\delta = (\delta_1, \cdots, \delta_d)$, $|\delta| = \delta_1 + \cdots + \delta_d$, $\| \ \|_d$ is the usual max norm over Δ^d, and the differentiation is taken in the complex variables. The completion in the norm (3.1) of the polynomials restricted to Δ^d is a Banach space and is denoted here by $A^m(\Delta^d)$. Our aim is to construct a basis in $A^m(\Delta^d)$. The problem is reduced to one variable by the following

THEOREM 3.1. *If there are simultaneous bases in $A^m(\Delta^{d_1})$ and $A^m(\Delta^{d_2})$, then their suitable tensor product is a simultaneous basis in $A^m(\Delta^{d_1+d_2})$.*

Now, let $A^m(T) = \{f: f(t) = F(e^{it}), F \in A^m(\Delta)\}$ and

$$\|f\|_T^{(m)} = \sum_{j=0}^{m} \|D_j f\|_T, \qquad \|g\|_T = \max\{|g(t)| : t \in T\}.$$

It can be seen that the two Banach spaces $\langle A^m(\Delta), \| \ \|_1^{(m)} \rangle$ and $\langle A^m(T), \| \ \|_T^{(m)} \rangle$ are linearly isomorphic. Thus, it is sufficient to construct a simultaneous basis in $\langle A^m(T), \| \ \|_T^{(m)} \rangle$.

Now, $f \in A(T) = A^0(T)$ iff $f = U + i(\tilde{U} + \mathrm{Im}(f, 1))$ and $U, \tilde{U} \in C(T)$, where

$$\tilde{U}(t) = -\frac{1}{\pi} \lim_{\varepsilon \to 0_+} \int_\varepsilon^\pi \frac{U(t+s) - U(t-s)}{2tgs/2} \, ds.$$

Using the basis $\{F_j^{(m)}\}$ in $C(T)$ for each fixed $m \geq 0$ the following system of complex valued functions is defined as

$$G_1 = 1, \qquad G_j = F_{j,+}^{(m)} + i\tilde{F}_{j,+}^{(m)} \quad \text{for } j = 2, 3, \cdots,$$

where F_+ is the even part of F. Moreover, for $0 \leq k \leq m$, let $G_1^{(\pm k)} = (2\pi)^{-1/2}$, $G_j^{(k)} = D^k G_j$ and $G_j^{(-k)} = H^k G_j$ for $j = 2, 3, \cdots$.

THEOREM 3.2. *For fixed $m \geq 0$ and for given k, $0 \leq k \leq m$, $\{G_j^{(k)}, j = 1, 2, \cdots\}$ is a basis in $\langle A(T), \|A\|_T \rangle$ and for f in $A(T)$ we have*

$$f = \sum_{j=1}^{\infty} (f, \mathrm{Re}\, G_j^{(-k)}) G_j^{(k)}.$$

THEOREM 3.3. *The system $\{G_j, j = 1, 2, \cdots\}$ is a simultaneous basis in $\langle A^m(T), \| \ \|_T^{(m)} \rangle$.*

An important role in the proof of Theorem 3.2 is played by Theorems 2.2 and 2.3, and Lemmas 2.1 and 2.2.

The complete proofs will appear in [6].

References

1. S. V. Botschkariev, *Existence of bases in the space of analytic functions, and some properties of the Franklin system*, Mat. Sb. **95 (137)** (1974), 3–18. (Russian)

2. Z. Ciesielski, *Properties of the orthonormal Franklin system*, Studia Math. **23** (1963), 141–157. MR **28** #419.

3. ———, *Properties of the orthonormal Franklin system*. II, Studia Math. **27** (1966), 289–323. MR **34** #3202.

4. Z. Ciesielski, *A construction of a basis in* $C^1(I^2)$, Studia Math. **33** (1969), 243–247. MR **40** #1759.

5. ———, *Constructive function theory and spline systems*, Studia Math. **53** (1974), 177–202.

6. ———, *Construction of a basis in the space of functions analytic in a polydisc and smooth on its boundary*, Studia Math. (to appear).

7. Z. Ciesielski and J. Domsta, *Construction of an orthonormal basis in* $C^m(I^d)$ *and* $W_p^m(I^d)$, Studia Math. **41** (1972), 211–224. MR **46** #9705.

8. J. Domsta, *A theorem on B-splines*, Studia Math. **41** (1972), 291–314. MR **46** #5897.

9. ———, *Approximation by interpolating spline bases*, Studia Math. (to appear).

10. ———, *A theorem on B-splines.* II: *The periodic case*, Bull. Acad. Polon. Sci. (to appear).

11. J. Radecki, *Orthonormal basis in the space* C_1 [0, 1], Studia Math. **35** (1970), 123–163. MR **42** #3484.

12. S. Schoenfeld, *Schauder bases in spaces of differentiable functions*, Bull. Amer. Math. Soc. **75** (1969), 586–590. MR **39** #6067.

13. ———, *A study of products and sums of Schauder bases in Banach spaces*, Dissertation, Purdue University, 1969; See also: *Schauder bases in the Banach space* $C^k(T^q)$, Trans. Amer. Math. Soc. **165** (1971), 309–318. MR **45** #2452.

14. Ju. N. Subbotin, *Spline approximation and smooth basis in* $C(0, 2\pi)$, Mat. Zametki **12** (1972), 43–51. (Russian) MR **47** #5493.

INSTITUTE OF MATHEMATICS OF THE POLISH ACADEMY OF SCIENCES
SOPOT, POLAND

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Recent Results on General Banach Spaces

Per Enflo*

Here we will consider the following general question: What can be said about the general Banach space? We will look into three aspects of this question, where substantial progress has been made over the last few years. The first we will call

1. The approximation problem and the basis problem. The approximation problem is the question whether every compact operator $T: B \to C$ between two Banach spaces can be approximated in the norm topology for operators by finite rank operators. In many special cases this problem has an affirmative answer. This is, for instance, the case when C has a Schauder basis. A Schauder basis of C is a sequence $\{e_j\}$ such that every $x \in C$ can be uniquely expressed as a convergent sum $x = \sum a_j e_j$. If we put $P_n: x \to \sum_{j=1}^{n} a_j e_j$ then the finite rank operators $T_n = P_n T$ will approximate T in the norm topology for operators.

The approximation problem and the basis problem—the question whether every separable Banach space has a Schauder basis—have now been solved negatively by myself. However, another problem in this direction as old as the ones mentioned —it appeared already in Banach's book from 1932—has recently been solved affirmatively by Pełczynski and Ovsepian. The question is the following: Is there in every separable Banach space a fundamental and total biorthogonal sequence (e_n, e_n^*) such that $\sup_n \|e_n\| \, \|e_n^*\| < \infty$?

A sequence (e_n, e_n^*) is called fundamental if the e_n's span the Banach space; it is called total if $e_n^*(x) = 0$ for all n implies $x = 0$; it is called biorthogonal if $e_n^*(e_m) = 0$ if $n \neq m$ and 1 if $n = m$. If a Banach space has a Schauder basis $\{e_n\}$ and if $\{e_n^*\}$ denotes the corresponding sequence of coefficient functionals, then (e_n, e_n^*) is a fundamental, total, biorthogonal sequence with $\sup_n \|e_n\| \, \|e_n^*\| < \infty$.

*This research was done under the sponsorship of National Science Foundation GP-43083.

There are several other properties weaker than having a Schauder basis for which the answer is not known in general, but for which an answer would be desirable in order to understand several long open problems about operators between Banach spaces. We will not go into these in this survey, but instead turn to

2. The role of the l_p-spaces in a general theory. The first result to be mentioned here is a now classical theorem by Dvoretsky: Every separable Banach space contains a $(1+\varepsilon)$-isomorph of n-dimensional Euclidean space for every $\varepsilon \to 0$ and n. Two Banach spaces B and C are said to be K-isomorphic if there is a continous, one-to-one operator T from B onto C with $\|T\| \, \|T^{-1}\| \leqq K$. They are said to be isomorphic if they are K-isomorphic for some K. If some part of the proof would generalize, one would have gotten the stronger result: Every separable Banach space contains an isomorph of l_p or c_0. l_p is the space of sequences $a = (a_1, a_2, \cdots)$ with the norm $\|a\| = (\sum |a_j|p)^{1/p}$, and c_0 is the space of sequences tending to 0 with the norm $\max_j |a_j|$.

This stronger result has been proved for Orlicz spaces by Lindenstrauss and Tzafriri. In general, however, it is false as was recently shown by Tzirelson. It seems that a weaker conjecture should be true: Every separable Banach space contains an isomorph of l_1 or c_0 or a reflexive Banach space. There is a recent characterization due to Rosenthal of Banach spaces containing l_1 which supports this conjecture:

A Banach space contains an isomorph of l_1 iff it contains a sequence $\{e_n\}$ with no weak Cauchy subsequence.

A sequence $\{e_n\}$ is a weak Cauchy sequence if $\{\varphi(e_n)\}$ is a Cauchy sequence for every functional φ.

The l_p -spaces also come up in several other contexts in connection with questions about general Banach spaces; for instance, some of them are characterized by special properties of their bases. Here we will mention the affirmative solution of the "complemented subspaces problem" due to Lindenstrauss and Tzafriri: A Banach space is isomorphic to l_2 if and only if there is a continuous linear projection onto every closed subspace.

3. The l_1^n-problem and uniform convexity. First in this section we should mention a result by R. C. James: Every nonreflexive Banach space contains a $(1+\varepsilon)$-isomorph of l_1^2 for every $\varepsilon > 0$. l_1^n is the space of n-tuples $(a_1, \cdots, a_n)$ with norm $\sum |a_j|$. The spaces which do not satisfy the conclusion of the theorem are called uniformly nonsquare. James later proved results for uniformly nonsquare Banach spaces and these were used by myself to prove the following result: Every uniformly nonsquare Banach space can be given an equivalent uniformly convex norm.

A norm $\| \ \|$ is called uniformly convex if there is a function $\delta(\varepsilon) > 0$ for $\varepsilon > 0$ such that $\|x\| \leqq 1$, $\|y\| \leqq 1$ and $\|x - y\| \geqq \varepsilon$ implies $\|x + y\| \leqq 2 - \delta(\varepsilon)$.

From the proof methods, the two results above were connected in the following way: If the first could be extended to l_1^n instead of l_1^2, then the second could be extended from "uniformly nonsquare" to "uniformly non-l_1^n". The question whether this is possible became known as the "l_1^n-problem". It was recently solved

negatively by James who proved the following result: There exists a nonreflexive Banach space B and an $\varepsilon > 0$ such that B does not contain a $(1 + \varepsilon)$-isomorph of l_1^3. The l_p-spaces are uniformly convex for $p > 1$ and for them $\delta(\varepsilon)$ can be chosen to be $K \cdot \varepsilon^q$ for some $K > 0$ and $q > 0$. The question, whether it is always possible to renorm a uniformly convex space so that $\delta(\varepsilon)$ can be chosen as $K \cdot \varepsilon^q$, was recently answered affirmatively by Pisier. For wide classes of uniformly convex spaces even more precise information has been obtained by Figiel and Johnson.

The development in the theory of general Banach spaces has gone fast during the last few years. Even so, there is certainly still a long way to go to get a satisfactory understanding to general Banach spaces. The richness in examples has for a long time made it hard to ask the right questions, a difficulty which is now overcome in many respects.

Of the results mentioned in this survey none is easy. The methods used to prove them have connections with many areas of mathematics such as harmonic analyisis, probability theory, integral geometry, combinatorics, etc.

STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305, U.S.A.

Труды Международного Конгресса Математиков
Ванкувер, 1974

Обзор Теории Размерности

В.В. Филиппов

Теория размерности общих топологических пространств имеет за последние несколько лет ряд крупных достижений, связанных с поиском естественных границ выполнения основных результатов классической теории. При этом с одной стороны получены сильные положительные результаты, с другой— построены интересные примеры.

Хорошо известно, какую роль играет совпадение различных размерностных инвариантов в теории размерности метрических пространств [**1**], [**4**], [**7**]. Этот вопрос естественно возникает и при построении теории размерности в более общей постановке. В частности, в последнее время привлекло внимание вопрос о соотношении размерностных функций в классе бикомпактных хаусдороровых пространств, поставленный П.С. Александровым в докладе на Московской международной топологической конференции в 1935 г. Как известно, для любого финально-компактного пространства X выполнено неравенство Александрова $\dim X \leqq \operatorname{ind} X \leqq \operatorname{Ind} X$, причем равенство 0 любой и 1 одной из индуктивных размерностей влечет равенство всех размерностей. В [**16**], [**17**] Филиппов построил пример бикомпактного хаусдоророва пространства X с $\operatorname{ind} X = 2$, $\operatorname{Ind} X = 3$. Эта конструкция видоизменялась Пасынковым и Лифановым [**8**], [**13**] и Филипповым. Наиболее сильный результат, полученный с ее помощью [**18**], есть

Предложение. *Для любого натурального числа $i \geqq 1$ существует бикомпакт X_i с* $\dim X_i = 1$, $\operatorname{ind} X_i = i$, $\operatorname{Ind} X_i = 2i - 1$.

В связи с вопросом о несовпадении размерностей dim и ind возник сильный метод построения примеров в классе бикомпактов [**14**], [**19**]. В [**19**] построена серия сравнительно простых примеров: для любых натуральных чисел m и n,

$1 \leqq m \leqq n$, существует бикомпакт X с первой аксиомой счетности с $\dim X = m$, $\operatorname{ind} X = n$. В предположении континуум—гипотезы этот же метод позволяет построить совершенно нормальный бикомпакт с теми же размерностными свойствами. Федорчук [15] для любого натурального $n \geqq 1$ построил бикомпакт, любое замкнутое подмножество которого либо n-мерно, либо нульмерно. В предположении $2^{\aleph_0} = \aleph_1$ Федорчук усилил этот результат, построив для любого натурального $n \geqq 1$ бикомпакт любое бесконечное замкнутое подмножество которого n-мерно.

В области изучения поведения размерности при замкнутых отображениях Филиппов получил две редукционные теоремы [20]. Первая есть:

Теорема. *Пусть $f: X \to Y$—замкнутое отображение нормального пространства X на нормальное пространство Y. Тогда существуют (метризуемые) компакты X' и Y', $\dim X' = \dim X$, $\dim Y' = \dim Y$ и непрерывное отображение $f': X' \to Y'$ при котором*

$$\dim\{y' : y' \in Y', |f'^{-1}(y')| \geqq k\} \leqq \operatorname{rd}_Y\{y : y \in Y, |f^{-1}(y)| \geqq k\}.$$

В другой теореме дополнительно предполагается, что Y есть F-пространство (см. [9], слабо паракомпактные пространства и σ-пространства являются F-пространствами) и тогда для конструируемого отображения мы можем считать выполненным следующее неравенство

$$\dim\{y' : y' \in Y', \dim f'^{-1}(y') \geqq k\} \leqq \operatorname{rd}_Y\{y : y \in Y, \dim f^{-1}(y) \geqq k\}.$$

Эти редукционные теоремы в ряде случаев делают тривиальным перенесение результатов со случая (метризуемых) компактов на общий случай. Так для замкнутого отображения $f: X \to Y$ выполнено неравенство Архангельского (см. [3])

$$(*) \qquad \dim Y \leqq \operatorname{rd}_Y\{y : y \in Y, |f^{-1}(y)| \geqq k\} + \dim X + k - 1,$$

если пространства X и Y нормальны, и оценки

$$(**) \quad \dim X \leqq \max\{\operatorname{rd}_Y\{y : y \in Y, \dim f^{-1}(y) \geqq k\} + k : k = 0, \cdots, \dim f\};$$
$$\dim Y \leqq \max\{\dim X, \operatorname{rd}_Y\{y : y \in Y, |f^{-1}(y)| \geqq 2\} + 1,$$
$$\operatorname{rd}_Y\{y : y \in Y, \dim f^{-1}(y) \geqq k\} + k : k = 1, \cdots, \dim f\};$$
$$\dim Y \leqq \operatorname{rd}_Y\{y : y \in Y, |f^{-1}(y)| \geqq k\} + \dim f + k - 1$$

(для $k = 1, \cdots, \dim Y - \dim X + 1$) если пространства X и Y нормальны и Y есть F-пространство.

Отметим, что формула Гуревича $\dim X \leqq \dim Y + \dim f$ (для замкнутого отображения $f: X \to Y$) не выполнена в классе нормальных пространств [20], но выполнена, когда образ есть F-пространство в силу неравенства Вайнштейна-Скляренко (**). В дополнительных теоретико-множественных предположениях может быть построен пример невыполнения формулы Гуревича в классе совершенно-нормальных пространств. Формула Гуревича $\dim Y \leqq \dim X +$ кратность $f - 1$ выполнена в силу (*). К ней примыкает следующая проблема, поставленная А.В. Архангельским: верно ли, что $\dim(X, \tau) \leqq n$ тогда и только

тогда, когда существуют топологии $\tau_1, \cdots, \tau_{n+1}$ на X, для которых $\dim(X, \tau_i) \leqq 0$ и $\tau = \bigcap_{i=1}^{n+1} \tau_i$? В классе метризуемых пространств ответ был получен В.П. Золотаревым [5] в следующем виде:

Теорема. *Пусть* (X, τ)*—метризуемое пространство. Тогда* $\dim(X, \tau) \leqq n$ *в том и только в том случае, если существуют такие топологии* $\tau_1, \cdots, \tau_{n+1}$, *что*:

(а) (X, τ_i)*—метризуемое пространство,* $i = 1, \cdots, n+1$,

(б) $(X, \bigcap_{i=1}^{k} \tau_i)$*—метризуемое пространство,* $k = 1, \cdots, n+1$,

(в) $\dim(X, \tau_i) \leqq 0$,

(г) $\tau = \bigcap_{i=1}^{n+1} \tau_i$.

Одной из центральных из оставшихся нерешенными задач общей теории размерности является вопрос о выполнении формулы

$$(***) \qquad \dim(X \times Y) \leqq \dim X + \dim Y$$

в классе нормальных пространств.

Наиболее сильными результатами, связанными с этим вопросом, являются

Теорема (Морита [6], Филиппов [21]). *Если в произведении* $X \times Y$ *вполне регулярных пространств проектирование на один из сомножителей есть замкнутое отображение, то*

$$\dim \beta(X \times Y) \leqq \dim \beta X + \dim \beta Y.$$

Теорема (Филиппов [21]). *Пусть пространство* $X \times Y$ *нормально и счетно паракомпактно,* Y*—паракомпактное p-пространство, тогда*

$$\dim(X \times Y) \leqq \dim X + \dim Y.$$

Аналогичный вопрос связан с неравенством

$$\operatorname{Ind}(X \times Y) \leqq \operatorname{Ind} X + \operatorname{Ind} Y.$$

В случае метрических пространств обе формулы совпадают и хорошо известны, но уже в классе бикомпактов есть существенные различия. Здесь формула (***), как известно, выполнена и тем не менее существуют бикомпакты X_1, $\dim X_1 = \operatorname{ind} X_1 = \operatorname{Ind} X_1 = 1$, и X_2, $\dim X_2 = \operatorname{ind} X_2 = \operatorname{Ind} X_2 = 2$, для которых $\operatorname{ind}(X_1 \times X_2) \geqq 4$ (Филиппов [22]). Положительные результаты, аналогичные соответствующим результатам для dim, могут быть получены при добавлении условия выполнения теоремы суммы для Ind в сомножителях [21].

В связи с вопросом о монотонности размерности отметим два результата.

В [21] Филиппов строит наследственно нормальное пространство X, $\dim X = 0$, содержащее подпространство всех положительных размерностей, чем решаются проблемы Чеха и Даукера. Недавно В.П. Золотарев, показал, что если в нормальном пространстве X лежит вполне паракомпактное подпространство Y, то $\dim Y \leqq \dim X$.

Для построения универсальных пространств данного веса и данной размерности Б.А. Пасынков создал сильный метод, использующий факторизационные

теоремы [10], [12]. Этим методом им недавно было доказано существование универсального бикомпакта данного веса для размерности Ind. Для размерности dim отметим метод А.В. Архангельского [2] доказательства факторизационной теоремы Мардешича.

Цитированная литература

1. П.С. Александров и Б.А. Пасынков, *Введение в теорию размерности*, Москва, 1973.
2. А.В. Архангельский, Доклады АН СССР **174** (1967), 1243–1246.
3. _____, Czechoslovak Math. J. **18** (1968), 389–391.
4. W. Hurewicz and H. Wallman, *Dimension theory.*
5. В.П. Золотарев, Доклады АН СССР **195** (1970), 539–543.
6. K Morita, GTA, 3, no. 2, 1973, 125–133.
7. J. Nagata, *Modern dimension theory.* Vol. 6, Wiley, New York, 1965.
8. Б.А. Пасынков, Доклады АН СССР **192** (1970), 503–506.
9. _____, Доклады АН СССР **202** (1972), 1274–1276.
10. _____, Доклады АН СССР **154** (1964), 1042–1043.
11. _____, Доклады АН СССР **182** (1968), 268–271.
12. _____, Доклады АН СССР **201** (1971), 1049–1052.
13. Б.А. Пасынков и И.К. Лифанов, Доклады АН СССР **192** (1970), 276–278.
14. В.В. Федорчук, Доклады АН СССР **182** (1968), 275–277.
15. _____, Доклады АН СССР **213** (1973), 795–797.
16. В.В. Филиппов, Доклады АН СССР **184** (1969), 1050–1053.
17. _____, Мат. сборник **83** (125): 1(9), (1970), 42–60.
18. _____, Доклады АН СССР **192** (1970), 289–293.
19. _____, Доклады АН СССР **192** (1970), 516–519.
20. _____, Доклады АН СССР **205** (1972), 40–43.
21. _____, Доклады АН СССР **209** (1973), 805–807.
22. _____, Доклады АН СССР **202** (1972), 1016–1019.

Московский Государственный Университет
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Results and Independence Results in Set-Theoretical Topology

A. Hajnal

The lecture contained a survey of results obtained by the author and I. Juhász concerning the following problems:

(A) How large discrete subspaces can be found in a T_2 space of large cardinality?

(B) What cardinalities can a T_2 space have?

(C) Under what conditions do the following properties imply each other:

(1) R is hereditarily α-separable,

(2) R is hereditarily α-Lindelöf?

Here is a list of publications where most of the results stated can be found:

References

1. A. Hajnal and I. Juhász, *Some results in set-theoretic topology*, Dokl. Akad. Nauk SSSR **172** (1967), 541–542 = Soviet Math. Dokl. **8** (1967), 141–143. MR **35** #957.

2. ———, *Discrete subspaces of topological spaces*, Nederl. Akad. Wetensch. Proc. Ser. A **70** = Indag. Math. **29** (1967), 343–356. MR **37** #4769.

3. ———, *Some remarks on a property of topological cardinal functions*, Acta Math. Acad. Sci. Hungar. **20** (1969), 25–37. MR **39** #3437.

4. ———, *On hereditarily α-Lindelöf and hereditarily α-separable spaces*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **11** (1968), 115–124. MR **39** #2124.

5. ———, *Discrete subspaces of topological spaces*. II, Nederl. Akad. Wetensch. Proc. Ser. A 72 = Indag. Math. **31** (1969), 18–30. MR **41** #9177.

6. ———, *Discrete subspaces and de Groot's conjecture about the number of open subsets*, Proc. Internat. Sympos. on Topology and its Applications, Belgrad, 1969, p. 179.

7. ———, *A consequence of Martin's axiom*, Nederl. Akad. Wetensch. Proc. Ser. A **74** = Indag. Math. **33** (1971), 457–463. MR **46** #4474.

8. ———, *On discrete subspaces of product spaces*, General Topology and Appl. **2** (1972), 11–16. MR **45** #7664.

9. A. Hajnal and I. Juhász, *Two consistency results in topology*, Bull. Amer. Math. Soc. **78** (1972), 711. MR **47** #4195.

10. ———, *A consistency result concerning hereditarily α-separable spaces.* Indag. Math. **35** (1973), 301–307.

11. ———, *A consistency result concerning hereditarily α-Lindelöf spaces*, Acta Math. Acad. Sci. Hungar. **24** (1973), 307–312.

12. ———, *On hereditarily α-Lindelöf and α-separable spaces.* II, Fund. Math. **81** (1974), 147–158.

13. ———, *On the number of open sets*, Ann. Univ. Sci. Budapestiensis **16** (1973), 99–102.

14. ———, *On first countable non-Lindelöf S-spaces*, Colloq. Math. Soc. J. Bolyai, **10**, Infinite and Finite Sets, Proc. Keszthely Sympos., 1973, pp. 837–852.

MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST V, HUNGARY

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Topological Structures

Horst Herrlich

"New points of view on old subjects are needed. There are too many point set topologists, and not enough who pay attention to what a topological space ought to be."
Saunders MacLane, Notices Amer. Math. Soc. **21** (1974), 183.

The concept of a topological space has been a prime object of topological investigations. Unfortunately it suffers from certain deficiencies such as:

(a) The category **Top** of topological spaces and continuous maps is not as well behaved as one would like it to be; e.g., **Top** is not cartesian closed, i.e., it is not possible to supply for any pair (X, Y) of topological spaces the set X^Y of all continuous maps from Y to X with a topology such that $(X^Y)^Z$ is naturally isomorphic to $X^{Y\times Z}$. Also, in **Top** the product of quotient maps in general is no longer a quotient map.

(b) Several important concepts of a topological nature—such as uniform convergence, uniform continuity, and completeness—cannot be expressed in the framework of the theory of topological spaces.

There have been serious efforts by prominent mathematicians to remedy this situation. But none of the solutions offered is free from all the deficiencies mentioned above. The purpose of this note is to stimulate discussion on these matters among point set topologists.

A. General features of convenient topological categories. To remedy the deficiencies mentioned in (a) several substitutes for **Top** have been suggested, e.g.,

(α) suitable subcategories of **Top**, e.g.,

the category of Kelley spaces which is the coreflective hull of all compact Hausdorff spaces in the category **Haus** of Hausdorff spaces and continuous maps;

the category of compactly generated spaces which is the coreflective hull of all

compact Hausdorff spaces in **Top**;

the category of sequential spaces which is the coreflective hull of all (compact) metrizable spaces in **Top**.

(β) suitable supercategories of **Top**, e.g.,

the category of quasi-topological spaces and continuous maps;

the category of limit spaces and continuous maps;

the category of filter-generated merotopic spaces and merotopic maps.

Several of the above categories have some other deficiencies, e.g., we do not know how to describe them in a sufficiently elegant manner directly from suitable axioms. But the general features of convenient topological categories **A** become apparent:

A should be a cartesian closed concrete category with small fibres, just one structure on any one-point set, and such that initial structures for arbitrary sources exist. Equivalently, the following conditions should be satisfied:

(**Top** 1) **A** is (complete and) cocomplete,

(**Top** 2) **A** is (well-powered and) co-well-powered,

(**Top** 3) for any **A**-object A the functor $Ax\_\colon \mathbf{A} \to \mathbf{A}$ preserves colimits,

(**Top** 4) the terminal **A**-object T is a separator and the functor $\hom(T, -)\colon \mathbf{A} \to \mathbf{Set}$ preserves colimits.

A list of references concerning this subject is provided in H. Herrlich, *Cartesian closed topological categories*, Math. Colloq. Univ. Cape Town **9** (1974), 1–16.

B. Concrete topological structures. Many solutions have been offered for problem (b), e.g., (generalized) uniform structures, proximity structures, contiguity structures, syntopogeneous structures and merotopic structures. Recent investigations have shown that suitable axiomatizations of each of the following concepts give rise to equivalent(!) satisfactory solutions of problem (b):

(a) the collection μ of all "uniform covers" of a set X,

(b) the collection γ of all collections of subsets of X which "contain arbitrary small members",

(c) the collection ξ of all collections of subsets of X which are "near" to some "spot" inside or outside X.

Consider the following axioms on μ:

(N1) μ is a nonempty collection of covers of X.

(N2) A cover of X belongs to μ if it is refined by some member of μ.

(N3) $\emptyset \notin \mu$.

A pair (X, μ) may be called a preuniform space provided it satisfies the above axioms. A map $f\colon (X, \mu) \to (Y, \nu)$ may be called uniformly continuous provided the preimage of each member of ν belongs to μ. The category **P-Unif** of preuniform spaces and uniformly continuous maps satisfies (**Top** 1), (**Top** 2), (**Top** 4) but not (**Top** 3). The functor $Ax\_\colon$ **P-Unif** $\to$ **P-Unif** preserves coequalizers but in general not coproducts. Most of the above-mentioned "convenient" topological categories can be nicely embedded into **P-Unif**.

Before we define several familiar full subcategories of **P-Unif** by adding further

axioms, we define the structure $\gamma = \mathrm{mer}(\mu)$ and the structure $\xi = \mathrm{near}(\mu)$ associated with a preuniform structure μ on X:

$$\mathfrak{A} \in \gamma \Leftrightarrow \forall\, \mathfrak{U} \in \mu\, \exists\, U \in \mathfrak{U}\, \exists\, A \in \mathfrak{A},\, A \subset U,$$
$$\mathfrak{B} \in \xi \Leftrightarrow \{X - B \,|\, B \in \mathfrak{B}\} \notin \mu.$$

Then each of the structures γ, respectively ξ, contains complete information about μ since

$$\mathfrak{U} \in \mu \Leftrightarrow \{X - U \,|\, U \in \mathfrak{U}\} \notin \xi \Leftrightarrow \forall\, \mathfrak{A} \in \gamma\, \exists\, A \in \mathfrak{A}\, \exists\, U \in \mathfrak{U},\, A \subset U.$$

Further axioms:

(N4) $\mathfrak{U} \in \mu$ and $\mathfrak{B} \in \mu$ imply $\{U \cap V \,|\, U \in \mathfrak{U} \text{ and } V \in \mathfrak{B}\} \in \mu$.

(N5) $\mathfrak{U} \in \mu$ implies $\{\mathrm{int}_\mu\, U \,|\, U \in \mathfrak{U}\} \in \mu$ where $x \in \mathrm{int}_\mu(U) \Leftrightarrow \{U,\, X - \{x\}\} \in \mu$.

(N6) $\mathfrak{U} \in \mu$ implies $\{V \subset X \,|\, \exists\, U \in \mathfrak{U}\, \{U, X - V\} \in \mu\} \in \mu$.

(N7) Every $\mathfrak{U} \in \mu$ has a star-refinement in μ.

(N8) Every $\mathfrak{U} \in \mu$ is refined by some finite member of μ.

(N9) If $\{\mathrm{int}_\mu\, U \,|\, U \in \mathfrak{U}\}$ covers X then $\mathfrak{U} \in \mu$.

Some results. The full subcategory of **P-Unif** whose objects satisfy axiom (N4) is denoted by **Q-Unif**. It is bicoreflective in **P-Unif** and isomorphic to each of the following categories:

(a) the category of quasi-uniform spaces and uniformly continuous maps in the sense of J. R. Isbell,

(b) the category of merotopic spaces and merotopic maps in the sense of M. Katětov,

(c) the category of discrete structure spaces in the sense of D. Harris,

(d) the category of quasi-nearness spaces and nearness preserving maps in the sense of H. Herrlich.

(N4) and (N5) is denoted by **Near**. It is bireflective in **Q-Unif**.

(N4) and (N6) is denoted by **R-Near**. It is a bireflective subcategory of **Near** and equivalent to the category of regular T-uniform spaces in the sense of K. Morita, resp. to the category of semiuniform spaces in the sense of A. K. Steiner and E. F. Steiner.

(N4) and (N7) is denoted by **Unif**. It is a bireflective subcategory of **Near** and isomorphic to the category of uniform spaces in the sense of A. Weil and J. W. Tukey.

(N4), (N5) and (N8) is denoted by **Cont**. It is a bireflective subcategory of **Near** and isomorphic to the category of contiguity spaces in the sense of V. M. Ivanova, A. A. Ivanov, and W. L. Terwilliger.

(N4), (N5) and (N9) is denoted by **Top**. It is a bicoreflective subcategory of **Near** and isomorphic to the category of symmetric (= R_0-, = weakly regular) topological spaces and continuous maps.

(N4), (N6) and (N8) is denoted by **Prox**. It is bireflective in **Near** and isomorphic to the category of proximity spaces and δ-continuous maps in the sense of V. A. Efremovič and Yu. M. Smirnov.

Furthermore, if an object of **Near** is called complete provided each (nonempty)

maximal element of ξ contains an adherence point, then there is a canonical construction by means of which any **Near**-object can be embedded densely into a complete **Near**-object. This completion generalizes the uniform completion of uniform spaces and the various distinguished Hausdorff compactifications of suitable topological, uniform, and proximity spaces. Moreover any strict extension, hence especially any regular extension, of an arbitrary topological T_1-space can be obtained in this way.

Further, a topological R_0-space is compact iff—considered as a **Near**-object—it is contigual, i.e., iff it satisfies (N8). A topological R_0-space is fully normal (= regular paracompact) iff—considered as a **Near**-object—it is a uniform space, i.e., iff it satisfies (N7).

A topological R_0-space is regular iff—considered as **Near**-object—it is regular, i.e., iff it satisfies (N6).

A topological R_0-space is normal iff—considered as **Near**-object—its contigual reflection is regular (equivalently, uniform).

Further details and a reference list are contained in H. Herrlich, *Topological structures*, Math. Centrum Amsterdam **52** (1974), 59–122.

UNIVERSITY OF BREMEN
BREMEN, FEDERAL REPUBLIC OF GERMANY

Труды Международного Конгресса Математиков
Ванкувер, 1974

Некоторые Экстремальные Задачи Теории Аппроксимации

Н. П. Корнейчук

В докладе будут изложены некоторые результаты по точному решению экстремальных задач теории приближения, полученные в последние годы в Днепропетровском государственном университете.

В ряде случаев успех был достигнут благодаря введению специального оператора, с определения которого мы и начнем. Будем пока рассматривать 2π-периодические функции (непериодический случай будет оговариваться особо). Пусть V—множество функций с ограниченным изменением на периоде, а W^rV $(r = 1, 2, \cdots)$—множество r-х периодических интегралов от $f \in V$. Любую функцию $f(x) \not\equiv \text{const}$ из W^1V (и даже из более широкого множества) можно представить в виде

$$f(x) = \sum_k \varphi_k(x) + d \qquad (x_0 \leqq x \leqq x_0 + 2\pi), \tag{1}$$

где $d = f(x_0)$, $|f(x_0)| = \min_x |f(x)|$, причем

(a) $\varphi_k(x)$—абсолютно непрерывные финитные функции с носителями $[\alpha_k, \beta_k] \subset [x_0, x_0 + 2\pi]$, строго монотонные на интервалах (α_k, α_k') и (β_k', β_k) $(\alpha_k < \alpha_k' \leqq \beta_k' < \beta_k)$ и постоянные на $[\alpha_k', \beta_k']$, если $\alpha_k' < \beta_k'$;

(b) если $x \in (\alpha_k, \alpha_k') \cup (\beta_k', \beta_k)$, то $f(x) = \varphi_k(x) + \text{const}$;

(c) $\int_0^{2\pi} |f(x)|dx = \sum_k \int_{\alpha_k}^{\beta_k} |\varphi_k(x)|dx + 2\pi|d|$;

(d) $V_0^{2\pi}(f) = \sum_k V_{\alpha_k}^{\beta_k}(\varphi_k)$.

Положим для $f \in W^1V$ $(f \not\equiv \text{const})$

$$\Phi(f, x) = \sum_k \bar{\varphi}_k(x) + |d| \qquad (0 \leqq x \leqq 2\pi),$$

где $\bar{\varphi}_k(x)$—убывающая перестановка функции $|\varphi_k|$ на $[0, 2\pi]$. Из определения

оператора Φ и свойств (a)—(d) разложения (1) следует, что $\Phi(f, x)$ не возрастает и абсолютно непрерывна, $\Phi(f, x) > |d|$ для $0 \leqq x < \Delta = \sup_k(\beta_k - \alpha_k)$, $\Phi(f, x) = |d|$ для $\Delta \leqq x \leqq 2\pi$, причем

$$\int_0^{2\pi} \Phi(f, x)\, dx = \|f\|_L = \int_0^{2\pi} |f|\, dx,$$
$$\Phi(f, 0) = \tfrac{1}{2} V_0^{2\pi}(f) + |d| = \tfrac{1}{2}\|f'\|_L + |d|.$$

Весьма полезным оказалось далеко не тривиальное неравенство [1], [2]

$$|\Phi'(f, x)| \leqq \frac{1}{4} \int_0^x \Phi(f'', t)\, dt,$$

точное на множестве W^3V.

Следующие два факта обусловили эффективность использования оператора Φ.

(1) Оператор Φ позволил обнаружить и описать с помощью точных неравенств новые экстремальные свойства дифференцируемых периодических функций. Для $a > 0$ пусть $\eta_{a0}(x) = \frac{1}{2}$ $(0 \leqq x < a)$, $\eta_{a0}(x) = 0$ $(x \geqq a)$ и

$$\eta_{ar}(x) = \begin{cases} \dfrac{1}{2} \displaystyle\int_0^{a-x} \eta_{a,r-1}(t)\, dt & (0 \leqq x < a), \\ 0 & (x \geqq a),\ r = 1, 2, \cdots. \end{cases}$$

Легко проверить, что $\eta_{\pi/n,r}(x) = \Phi(g_{nr}, x)$ $(0 \leqq x \leqq 2\pi)$, где $g_{nr}(x)$—r-й периодический интеграл от $g_{n0}(x) = \operatorname{sign} \sin nx/4n$ с нулевым средним значением на периоде.

Теорема 1 [2]. *Если* $g \in W^rV$ $(r = 1, 2, \cdots)$, $V_0^{2\pi}(g^{(r)}) \leqq 1$ *и при некотором* $a > 0$

$$\|g\|_L \leqq \int_0^{2\pi} \eta_{ar}(t)\, dt,$$

то $|\Phi'(g, x)| \leqq |\eta'_{a,}(x)|$ *везде на* $(0, a)$, *где* $\Phi'(g, x)$ *существует, т.е. почти всюду.*

Следствие. *Если* $g \in W^rV$ $(r = 1, 2, \cdots)$, $V_0^{2\pi}(g^{(r)}) \leqq 1$ *и* $\|g\|_L \leqq \|g_{nr}\|_L$, *то* $|\Phi'(g, x)| \leqq |\Phi'(g_{nr}, x)|$ $(0 < x < \pi/n)$.

Теорема 2 [1]. *Пусть* $g \in W^rV$ $(r = 1, 2, \cdots)$, $|g^{(r)}(x)| \leqq 1$ *и* $\|g\|_C \leqq \|\psi_{nr}\|_C$, *где* $\psi_{nr}(x) = 4ng_{nr}(x)$. *Тогда почти всюду на* $(0, \pi/n)$ *выполняется, по меньшей мере, одно из неравенств*:

$$|\Phi'(g, x)| \leqq |\Phi'(\psi_{nr}, x)| \quad \textit{или} \quad \Phi(g, x) \leqq \Phi(\psi_{nr}, x),$$

а при $r \geqq 3$ *первое неравенство выполняется обязательно.*

Сформулированные утверждения можно рассматривать как аналоги теоремы сравнения А.Н. Колмогорова [3]: если $\|g\|_C \leqq \|\psi_{nr}\|_C$ и $|g^{(r)}(x)| \leqq 1$, то из равенства $g(a) = \psi_{nr}(\alpha)$ следует $|g'(a)| \leqq |\psi'_{nr}(\alpha)|$.

(2) С помощью оператора Φ можно в ряде случаев получить точную оценку для интеграла от произведения двух функций. Пусть H^ω—класс непрерывных

2π-периодических функций $f(x)$ таких, что $|f(x')-f(x'')| \leqq \omega(|x'-x''|)$, где $\omega(t)$—заданный модуль непрерывности, который везде ниже мы считаем выпуклым.

ТЕОРЕМА 3 [1], [2]. *Если* $g \in V$, $\int_0^{2\pi} g(t)\,dt = 0$, *то*

$$\sup_{f\in H^\omega} \int_0^{2\pi} g(t)f(t)\,dt \leqq \min_c \int_0^{2\pi} \Phi(G_c, t)\omega'(t)\,dt, \tag{2}$$

где $G_c(x) = \int_c^x g(t)\,dt$. *Существуют функции* $g \in V$, *для которых в* (2) *имеет место знак равенства.*

Перехожу к рассмотрению конкретных задач.

1. Наилучшее приближение тригонометрическими полиномами на классах дифференцируемых функций. Пусть X есть пространство C или L 2π-периодических функций $f(x)$ и

$$E_n(f)_X = \inf_{T\in F^T_{2n-1}} \|f - T\|_X,$$

где F^T_{2n-1}—подпространство тригонометрических полиномов порядка $n-1$ ($\dim F^T_{2n-1} = 2n-1$). Через W^rH^ω обозначим класс функций f из C^r, у которых $f^{(r)} \in H^\omega$ ($W^0H^\omega = H^\omega$). Если $f \in W^rH^\omega$, то в силу соотношений двойственности С.М. Никольского [4]

$$E_n(f)_C = \sup_{g\in W^r_{Ln}} \int_0^{2\pi} g(t)f^{(r)}(t)\,dt,$$

где $g \in W^r_{Ln}$ означает, что $g(x)$ ортогональна подпространству F^T_{2n-1}, $g^{(r-1)}(x)$ абсолютно непрерывна и $\|g^{(r)}\|_L \leqq 1$. Пусть $g \in W^r_{Ln}$ и $g_1(x)$—интеграл от g с нулевым средним значением на периоде. Тогда $g_1 \in W^{r+1}_{Ln}$ и, как следует из результатов С.М. Никольского [4], $\|g_1\|_L \leqq \|g_{nr}\|_L$. По Теореме 3

$$\int_0^{2\pi} g(t)f^{(r)}(t)\,dt \leqq \int_0^{2\pi} \Phi(g_1, t)\omega'(t)\,dt$$

$$= \int_0^{\pi/n} \eta_{\pi/n,r}(t)\omega'(t)\,dt - \int_0^{2\pi} [\eta_{\pi/n,r}(t) - \Phi(g_1, t)]\omega'(t)\,dt.$$

Последний интеграл по следствию из Теоремы 1 неотрицателен, поэтому для любой $f \in W^rH^\omega$

$$E_n(f)_C \leqq \int_0^{2\pi} \eta_{\pi/n,r}(t)\omega'(t)\,dt. \tag{3}$$

Легко проверить, что правая часть (3) совпадает с нормой в C функции $f_{nr}(x) = f_{nr}(\omega, x)$ периода $2\pi/n$ с нулевым средним значением на периоде, у которой

$$f^{(r)}_{nr}(x) = f_{n0}(x) = \begin{cases} \frac{1}{2}\omega(2x) & (0 \leqq x \leqq \pi/2n), \\ \frac{1}{2}\omega(2\pi/n - 2x) & (\pi/2n \leqq x \leqq \pi/n), \end{cases}$$

$$f_{n0}(-x) = -f_{n0}(x).$$

Так как $f_{nr} \in W^rH^\omega$ и $E_n(f_{nr})_C = \|f_{nr}\|_C$, то приходим к соотношениям [1]

$$\sup_{f\in W^rH^\omega} E_n(f)_C = \|f_{nr}\|_C = \int_0^{\pi/n} \eta_{\pi/n,r}(t)\omega'(t)\,dt$$

$$= \frac{1}{n^{r+1}} \int_0^{\pi} \eta_{\pi,r}(t)\omega'\left(\frac{t}{n}\right) dt \qquad (r = 0, 1, 2, \cdots; n = 1, 2, \cdots). \tag{4}$$

При $\omega(t) = Kt$ это есть известный результат Фавара-Ахиезера-Крейна для классов KW_M^{r+1} $(r = 0, 1, \cdots)$ функций $f \in C^r$, у которых $f^{(r)}(x)$ абсолютно непрерывна и vrai $\sup_x |f^{(r+1)}(x)| \leqq K$.

Аналогично, с помощью Теоремы 2, можно прийти к следующим равенствам [1]:

$$\sup_{f\in W^rH^\omega} E_n(f)_L = \|f_{nr}\|_L = \frac{4}{n^{r+1}} \int_0^{\pi} \eta_{\pi,r+1}(t)\omega'\left(\frac{t}{n}\right) dt.$$

Отметим еще результат А.А. Лигуна [5], который установил, что в неравенстве Джексона

$$E_n(f)_X \leqq \frac{c_r}{n^r}\,\omega\left(f^{(r)}, \frac{\pi}{n}\right)_X$$

($\omega(f, t)_X$—модуль непрерывности функции f в пространстве X) при $X = C$ или L и всех нечетных r наименьшее возможное значение константы c_r есть $2/\pi\ \sum_{i=0}^{\infty}(2i+1)^{-r-1}$. (Для $r = 1$ это ранее доказал В.В. Жук [6].) Существенно, что здесь указан линейный оператор из X в F_{2n-1}^T, осуществляющий приближение с этой константой на всем множестве функций f, у которых $f^{(r)} \in X$.

2. Наилучшее равномерное приближение класса W^rH^ω классом KW_M^{r+1}. Если $\varphi_0(x)$ из KW_M^{r+1} есть функция наилучшего равномерного приближения для $f \in W^rH^\omega$, то в силу общего критерия Иоффе-Тихомирова [7] существует функция $g_0 \in V$ с единичной вариацией и нулевым средним значением на периоде такая, что

$$\int_0^{2\pi} \varphi_0 dg_0 = \sup_{\varphi\in KW_M^{r+1}} \int_0^{2\pi} \varphi dg_0, \qquad \|f - \varphi_0\|_C = \int_0^{2\pi} (f - \varphi_0) dg_0.$$

Эти соотношения, а также Теорема 3 позволяют получить оценку

$$\rho(f, KW_M^{r+1})_C \overset{\text{def}}{=} \inf_{\varphi\in KW_M^{r+1}} \|f - \varphi\|_C \leqq \int_0^{2\pi} [\omega'(t) - K]\Phi(g_r, t)\,dt,$$

где $g_r(x)$—r-й периодический интеграл от g_0 такой, что $\int_0^{2\pi} \operatorname{sign} g_r(t)\,dt = 0$. Если выбрать $b > 0$ такое, чтобы было $\|\eta_{br}\|_L = \|\Phi(g_r)\|_L$, то по Теореме 1 $|\Phi'(g_r, t)| \leqq |\eta'_{br}(t)|$ $(0 < t < b)$ и мы приходим к неравенствам [2]

$$\rho(f, KW_M^{r+1})_C \leqq \int_0^{b} [\omega'(t) - K]\eta_{br}(t)\,dt \leqq \max_{0\leqq a\leqq\pi} \int_0^{a} [\omega'(t) - K]\eta_{ar}(t)\,dt.$$

Последняя оценка на классе W^rH^ω точна, когда $K = K_n$ выбрано так, что максимум в правой части достигается при $a = \pi/n$ $(n = 1, 2, \cdots)$.

3. Поперечники классов функций и экстремальные подпространства. Пусть F_N—подпространство размерности N нормированного пространства X и для $\mathfrak{M} \subset X$

$$E(\mathfrak{M}, F_N)_X = \sup_{f\in\mathfrak{M}} \inf_{\varphi\in F_N} \|f - \varphi\|_X.$$

N-мерным поперечником (по Колмогорову) множества $\mathfrak{M}$ в пространстве X называют величину (см., напр., [8])

$$d_N(\mathfrak{M}, X) = \inf_{F_N} E(\mathfrak{M}, F_N)_X.$$

Если $E(\mathfrak{M}, \tilde{F}_N)_X = d_N(\mathfrak{M}, X)$, то $\tilde{F}_N$ называется экстремальным подпространством для $\mathfrak{M}$ в X. Из приведенных в п. 1 результатов следует, что когда X есть C или L, то

$$d_{2n-1}(W^rH^\omega, X) \leqq \|f_{nr}\|_X = E(W^rH^\omega, F^T_{2n-1})_X. \tag{5}$$

Неравенство $d_{2n-1}(W^rH^\omega, X) \geqq \|f_{nr}\|_X$ для $X = C$ содержится в книге Lorentz'a [9], а для $X = L$ доказано В.П. Моторным и В.И. Рубаном [10], так что на самом деле в (5) нужно оставить знак равенства. Кроме того, В.И. Рубан установил [11], что $d_{2n}(W^rH^\omega, C) = d_{2n-1}(W^rH^\omega, C)$. Таким образом, тригонометрическое подпространство F^T_{2n-1} является экстремальным для классов W^rH^ω ($r = 0, 1, \cdots$) в C и в L, причем в пространстве C никакое подпространство размерности $2n$ не дает лучшего приближения.

Отметим, что в случае $\omega(t) = Kt$, т.е. для классов KW^r_M ($r = 1, 2, \cdots$) эти результаты в пространстве C были ранее получены В. М. Тихомировым [8], [12], а в пространстве L—Ю.И. Маковозом [13].

Используя Теоремы 1 и 3 можно доказать [14], что если при некотором $r = 0, 1, 2, \cdots$ подпространство F_N является экстремальным в C для W^rH^ω при $\omega(t) = t$, то F_N будет экстремальным в C для класса W^rH^ω при любом выпуклом $\omega(t)$. В частности, подпространство S^r_{2n} полиномиальных сплайнов порядка r и дефекта 1 по равномерному разбиению $k\pi/n$ экстремально в C не только для класса W^{r+1}_M, как показал В.М. Тихомиров [12], но и для класса W^rH^ω. Заметим однако, что хотя поперечники класса W^{r+1}_M в C реализуют интерполяционные сплайны из S^r_{2n} [12], на классе W^rH^ω в случае нелинейности $\omega(t)$ на $[0, \pi/n]$ этот факт по крайней мере при четных r не имеет места.

4. Приближение интерполяционными сплайнами на классах функций. Здесь библиография предшествующих исследований особенно обширна, но мы, как и в предыдущих пунктах, ограничимся тем, что приведем несколько наиболее интересных, на наш взгляд, окончательных результатов.

При оценке погрешности приближения сплайнами на классах дифференцируемых функций решающую роль имело исследование экстремальных свойств ядра интегрального представления этой погрешности и получение утверждений типа Теоремы 3. Будем обозначать через $\sigma_{mn}(f, x)$ интерполяционные эрмитовы сплайны порядка $2m + 1$ и дефекта $m + 1$ по разбиению $\Delta = \{0 = x_0 < x_1 < \cdots < x_n = 1\}$ такие, что $\sigma^{(i)}_{mn}(f, x_k) = f^{(i)}(x_k)$ ($i = 0, 1, \cdots, m$; $k = 0, 1, \cdots, n$). Множества заданных на $[0, 1]$ функций, вообще говоря непериоди-

ческих, условимся отмечать чертой сверху. Предполагая что $\bar{\mathfrak{M}} \subset \bar{C}^s$, но $\bar{\mathfrak{M}} \not\subset \bar{C}^{s+1}$, положим

$$\mathcal{E}(\bar{\mathfrak{M}}, \Delta)_{L_p} = \sup_{f \in \bar{\mathfrak{M}}} \|f - \sigma_{[s/2], n}(f)\|_{L_p(0,1)} \qquad (1 \leqq p \leqq \infty).$$

Для классов $\bar{W}^r_{L_p}$ $(r = 1, 2, \cdots)$ функций f, у которых $f^{(r-1)}$ абсолютно непрерывна на $[0,1]$ и $\|f^{(r)}\|_{L_p(0,1)} \leqq 1$ $(\bar{W}^r_{L_\infty} = \bar{W}^r_M)$ В. Л. Великин доказал [15], что

$$\mathcal{E}(\bar{W}^r_M, \Delta)_C = |\Delta|^r r^{-1} [(r-1)!!]^{-2} 2^{-r} \quad (r \text{ нечетно}),^1$$
$$\mathcal{E}(\bar{W}^r_L, \Delta)_L = |\Delta|^r (r!)^{-1} 2^{-r} \quad (r \text{ четно}),$$

где $|\Delta| = \max_{1 \leqq k \leqq n}(x_k - x_{k-1})$, а также при всех четных r вычислил значения величин $\mathcal{E}(\bar{W}^r_M, \Delta)_{L_p}$, $\mathcal{E}(\bar{W}^r_{L_p}, \Delta)_L$ и $\mathcal{E}(\bar{W}^r_L, \Delta)_C$. В работе [15] дана также неулучшаемая на множестве $\bar{C}^r$ $(r = 1, 3, 5, \cdots)$ оценка

$$|f(x) - \sigma_{(r-1)/2, n}(f, x)| \leqq |\Delta|^r r^{-1} [(r-1)!!]^{-2} 2^{-r-1} \omega(f^{(r)}, |\Delta|).$$

Что касается классов $\bar{W}^r H^\omega$, то мы можем указать на полученные в статье [17] равенства

$$\mathcal{E}(\bar{W}^2 H^\omega, \Delta)_C = \frac{1}{4} \int_0^{|\Delta|/2} t\omega(t)\, dt; \qquad \mathcal{E}(\bar{W}^3 H^\omega, \Delta)_C = \frac{1}{32} \int_0^{|\Delta|} t(|\Delta| - t)\omega(t)\, dt,$$

которые интересно сопоставить с соотношениями (4) п. 1 при $r = 2$ и 3.

Отметим еще, что в работах В. Л. Великина [18] и А. А. Женсыкбаева [19], [20] получен ряд окончательных результатов по оценке приближения интерполяционными сплайнами дефекта 1 на классах непрерывных и дифференцируемых периодических функций.

5. Наилучшие квадратурные формулы для классов функций. Рассматриваются квадратурные формулы вида

$$\int_a^b f(x)\, dx = \sum_{k=1}^{n} \sum_{i=0}^{\rho} p_{ki} f^{(i)}(x_k) + R(f). \tag{6}$$

Если $\mathfrak{M}$—некоторый класс заданных на $[a, b]$ и достаточное число раз дифференцируемых функций, то задача состоит в отыскании величины

$$\inf_{x_k, p_{ki}} \sup_{f \in \mathfrak{M}} |R(f)|$$

и в указании узлов $\bar{x}_k$ и коэффициентов $\bar{p}_{ki}$, реализующих точную нижнюю грань. Для классов функций с ограниченной по норме в L_p r-ой производной эта задача редуцируется к эквивалентной задаче минимизации нормы некоторого моносплайна в сопряженной метрике [21], [22]. Здесь получен ряд окончательных результатов (как в периодическом, так и в непериодическом случаях), причем наиболее далеко исследования в этом направлении продвинуты в работах Н. Е. Лушпая (см., напр., [23]). (Подробнее об этом сказано в добавлении автора к находящемуся в печати второму изданию книги [21].)

Базируясь на совершенно иных идеях и используя тонкие факты топологического характера, а также рассуждения, связанные с применением оператора

[1] Для четных r величина $\mathcal{E}(\bar{W}^r_L, \Delta)_C$ вычислена ранее в [16].

Φ, В.П. Моторный доказал [24], что среди квадратурных формул (6) при $[a, b] = [0, 2\pi]$ и $\rho = 0,1$ наилучшей для определенных в п. 1 периодических классов W_M^r $(r = 1, 2, \cdots)$ и $W^r H^\omega$ $(r = 1, 3, 5, \cdots)$ является формула прямоугольников с $\bar{x}_k = 2k\pi/n$, $\bar{p}_{k0} = 2\pi/n$, $\bar{p}_{k1} = 0$. Верхняя грань остатка этой формулы на указанных классах равна соответственно $2\pi \| \phi_{nr} \|_C$ и $2\pi \| f_{nr} \|_C$.

Литература

1. Н.П. Корнейчук, *Экстремальные значения функционалов и наилучшее приближение на классах периодических функций*, Известия АН СССР, серия матем. **35** (1971), 93–124.

2. _____, *Неравенства для дифференцируемых периодических функций и наилучшее приближение одного класса функций другим*, Известия АН СССР, серия матем. **36** (1972), 423–434.

3. А.Н. Колмогоров, *О неравенствах между верхними гранями последовательных производных произвольной функции на бесконечном интервале*, Учен. зап. Москов. ун-та, вып. 30, "Математика" **3** (1939), 3–13.

4. С.М. Никольский, *Приближение функций тригонометрическими полиномами в среднем*, Известия АН СССР, серия матем. **10** (1946), 207–256.

5. А.А. Лигун, *О точных константах приближения дифференцируемых периодических функций*, Мат. заметки **14** (1973), 21–30.

6. В.В. Жук, *Некоторые точные неравенства между равномерными приближениями периодических функций*, Доклады АН СССР **201** (1967), 263–266.

7. А.Д. Иоффе и В.М. Тихомиров, *Двойственность выпуклых функций и экстремальные задачи*, Успехи матем. наук **27**, вып. **6** (144), (1968), 51–116.

8. В.М. Тихомиров, *Поперечники множеств в функциональных пространствах и теория наилучших приближений*, Успехи матем. наук **15**, вып. **3** (93), (1960), 81–120.

9. G.G. Lorentz, *Approximation of functions*, Holt, Rinehart and Winston, New York, 1966.

10. В.П. Моторный и В.И. Рубан, *Поперечники некоторых классов дифференцируемых периодических функций в пространстве L*, Мат. заметки **16** (1974) (в печати).

11. В.И. Рубан, *Четные поперечники классов $W^{(r)}H^\omega$ в пространстве $C_{2\pi}$*, Мат. заметки **15** (1974), 387–392.

12. В.М. Тихомиров, *Наилучшие методы приближения и интерполирования дифференцируемых функций в пространстве $C[-1, 1]$*, Мат. сборник **80** (1969), 290–304.

13. Ю.И. Маковоз, *Об одном приеме оценки снизу поперечников множеств в банаховых пространствах*, Мат. сборник **87** (1972), 136–142.

14. Н.П. Корнейчук, *О равномерном приближении периодических функций подпространствами конечной размерности*, Доклады АН СССР **213** (1973), 525–528.

15. В.Л. Великин, *Точные значения приближения эрмитовыми сплайнами на классах дифференцируемых функций*, Известия АН СССР, серия матем. **37** (1973), 165–185.

16. P.G. Ciarlet, M.H. Schultz and R.S. Varga, *Numerical methods of high-order accuracy for nonlinear boundary value problems*, I. *One dimensional problem*, Numer. Math. **9** (1967), 394–430.

17. В.Л. Великин и Н.П. Корнейчук, *Точные оценки приближения сплайн-функциями на классах дифференцируемых функций*, Мат. заметки **9** (1971), 483–494.

18. В.Л. Великин, *Приближение кубическими сплайнами на классах непрерывно дифференцируемых функций*, Мат. заметки **11** (1972), 215–226.

19. А.А. Женсыкбаев, *Точные оценки равномерного приближения непрерывных периодических функций сплайнами r-го порядка*, Мат. заметки **13** (1973), 217–228.

20. _____, *Приближение дифференцируемых периодических функций сплайнами по равномерному разбиению*, Мат. заметки **13** (1973), 807–816.

21. С.М. Никольский, *Квадратурные формулы*, Физматгиз, Москва, 1958.

22. Н.П. Корнейчук и Н.Е. Лушпай, *Наилучшие квадратурные формулы для классов*

дифференцируемых функций и кусочно-полиномиальное приближение, Известия АН СССР, серия матем. **33** (1969), 1416–1437.

23. Н.Е. Лушпай, *Наилучшая квадратурная формула на классе $W^r_* L_2$ периодических функций*, Мат. заметки **16** (1974), 193–204.

24. В.П. Моторный, *О наилучшей квадратурной формуле вида $\sum_{k=1}^n p_k f(x_k)$ для некоторых классов периодических дифференцируемых функций*, Известия АН СССР, серия матем. **38** (1974), 583–614.

Днепропетровский Университет
Днепропетровск, СССР

Actes du Congrès International des Mathématiciens
Vancouver, 1974

Quelques Problèmes de Factorisation d'Opérateurs Linéaires

Bernard Maurey

En 1956, A. Grothendieck a démontré le résultat suivant: Tout opérateur linéaire continu d'un espace $C(K)$ dans un espace L^1 se factorise par un espace de Hilbert. Plus précisément, il existe une constante K_G telle que pour tout compact K, et pour tout opérateur linéaire u de norme $\leqq 1$ de $C(K)$ dans un espace L^1, il existe une probabilité de Radon μ sur K telle que u admette la factorisation: $C(K) \xrightarrow{j} L^2(K, \mu) \xrightarrow{\bar{u}} L^1$, où j désigne l'application naturelle de $C(K)$ dans $L^2(K, \mu)$, et où $\|\bar{u}\| \leqq K_G$ (voir [2], [5]).[1]

Dans le cas d'un opérateur de c_0 dans un espace L^1, le résultat prend la forme suivante: Tout opérateur linéaire u de norme $\leqq 1$ de c_0 dans un espace L^1 admet la factorisation $c_0 \xrightarrow{\alpha} l^2 \xrightarrow{\bar{u}} L^1$, où α est un opérateur diagonal $(\beta_n) \to (\alpha_n \beta_n)$, avec $\sum|\alpha_n|^2 \leqq 1$, et où $\|\bar{u}\| \leqq K_G$. Cette dernière propriété peut être interprétée comme une propriété des séries inconditionnellement convergentes dans un espace L^1 : si $\sum x_n$ est une série inconditionnellement convergente dans un espace L^1, avec $\|\sum \varepsilon_n x_n\| \leqq 1$ pour toute suite de nombres $\varepsilon_n = \pm 1$, on peut trouver une décomposition $x_n = \alpha_n y_n$, où (α_n) est une suite scalaire telle que $\sum|\alpha_n|^2 \leqq 1$, et où (y_n) vérifie : $\forall\, (\beta_n)$, $\|\sum \beta_n y_n\| \leqq K_G(\sum|\beta_n|^2)^{1/2}$.

Du point de vue dual, on peut dire que tout opérateur linéaire continu d'un espace L^∞ dans l^1 admet la factorisation $L^\infty \to l^2 \xrightarrow{\alpha} l^1$. Plus généralement Grothendieck a démontré que tout opérateur linéaire u de norme $\leqq 1$ d'un espace $C(K)$ dans $L^1(\Omega, \mu)$ admet la factorisation:

$$C(K) \xrightarrow{\bar{u}} L^2(\Omega, \mu) \xrightarrow{T_\cdot} L^1(\Omega, \mu),$$

[1] Cette constante n'est pas égale à la constante K_G de [5].

où T_g désigne un opérateur de multiplication $T_g(f) = fg$, avec $\int |g|^2 \, d\mu \leqq 1$, et $\|\tilde{u}\| \leqq K_G$.

Nous étudierons dans ce qui suit deux problèmes de factorisation suggérés par le théorème de Grothendieck. L'un concerne simplement la possibilité de factoriser un opérateur donné à travers un espace de Hilbert. L'autre problème est l'étude des espaces de Banach E vérifiant la propriété suivante:

(D) Tout opérateur linéaire continu d'un espace $C(K)$ dans E admet la factorisation $C(K) \xrightarrow{i} L^2(K, \mu) \to E$, où μ est une probabilité de Radon sur K.

Comme dans le cas du théorème de Grothendieck, la propriété (D) admet deux propriétés équivalentes: la décomposition des suites (x_n) telles que la série $\sum x_n$ soit inconditionnellement convergente, sous la forme $x_n = \alpha_n y_n$, $\sum |\alpha_n|^2 < \infty$, $\sup\{\|\sum \beta_n y_n\| ; \sum |\beta_n|^2 \leqq 1\} < \infty$, et la factorisation des opérateurs linéaires continus de E' (dual de E) dans un espace $L^1(\Omega, \mu)$ sous la forme $E' \to L^2(\Omega, \mu) \xrightarrow{T_g} L^1(\Omega, \mu)$, avec $g \in L^2(\Omega, \mu)$.

A partir de cette dernière formulation, il est naturel de poser une question plus générale: trouver un critère pour qu'un opérateur linéaire continu u d'un espace de Banach F dans un espace $L^p(\Omega, \mu)$, $0 < p < \infty$, admette la factorisation:

$$(F_q) \qquad F \xrightarrow{\tilde{u}} L^q(\Omega, \mu) \xrightarrow{T_g} L^p(\Omega, \mu)$$

avec $p \leqq q \leqq \infty$, $\tilde{u}$ continu et $g \in L^r(\Omega, \mu)$, $1/p = 1/q + 1/r$.

THÉORÈME 1 ([13] ET [6, CHAPITRE I]). *Pour que $u: F \to L^p(\Omega, \mu)$ admette la factorisation (F_q) il faut et il suffit qu'il existe une constante C telle que*

$$\left(\int \left(\sum |u(x_n)|^q\right)^{p/q} d\mu\right)^{1/p} \leqq C\left(\sum \|x_n\|^q\right)^{1/q}, \quad \forall (x_n) \in F.$$

La démonstration utilise des arguments de convexité, et en particulier une forme du théorème du minimax. Ce théorème a été généralisé dans un cadre plus abstrait par J. L. Krivine [12, Exposés 22—23].

La condition (D) est en général difficile à vérifier directement. Nous allons introduire une condition sur les suites "presque inconditionnellement convergentes", c'est-à-dire les suites (x_n) telles que la série $\sum \varepsilon_n x_n$ converge pour presque tout choix de signes $\varepsilon_n = \pm 1$. Nous dirons qu'un espace de Banach E est de cotype 2 si'il existe une constante K telle que:

$$(C) \qquad \left(\sum \|x_n\|^2\right)^{1/2} \leqq K\left(\int \left\|\sum x_n \varepsilon_n(t)\right\|^2 dP(t)\right)^{1/2}, \quad \forall (x_n) \in E,$$

où $(\varepsilon_n(t))$ désigne une suite de variables de Bernoulli sur un espace de probabilité (X, P), c'est-à-dire une suite de variables aléatoires indépendantes, prenant les valeurs ± 1 avec probabilité 1/2.

On vérifie facilement que $L^p(\Omega, \mu)$ est de cotype 2 pour $1 \leqq p \leqq 2$ (en utilisant les inégalités de Khintchine). Plus généralement, si E est de cotype 2, $L^p(\Omega, \mu, E)$ est de cotype 2 pour $1 \leqq p \leqq 2$. Si A est une C^*-algèbre (non nécessairement commutative!), son dual A' est de cotype 2; ce résultat nettement plus difficile est dû à N. Tomczak [14].

THÉORÈME 2. (C) ⇒ (D).

(Ce théorème contient le théorème de Grothendieck, puisque L^1 est de cotype 2.)

La démonstration se fait en deux étapes: la première, assez facile, consiste à montrer que lorsque E est de cotype 2, tout opérateur linéaire continu de E' dans un espace $L^p(\Omega, \mu)$ admet la factorisation (F_2), lorsque $1 < p \leqq 2$. Dans la deuxième étape, plus difficile, on passe de $p > 1$ à $p = 1$. Cela peut être fait d'au moins quatre façons différentes: [**6**, Chapitre VIII], [**11**, Exposé XXII, Théorème 1 bis], [**9**] ou [**10**].

On peut généraliser la propriété (D) de la façon suivante: Notons $L^0(\Omega, \mu)$ l'espace vectoriel des classes de fonctions mesurables scalaires sur un espace de probabilité (Ω, μ), muni de la topologie de la convergence en probabilité. On a

THÉORÈME 2 BIS. *Supposons que E vérifie* (C) *et que E' vérifie l'hypothèse d'approximation bornée. Tout opérateur linéaire continu de E' dans un espace $L^0(\Omega, \mu)$ admet la factorisation: $E' \to L^2(\Omega, \mu) \xrightarrow{T_g} L^0(\Omega, \mu)$, avec $g \in L^0(\Omega, \mu)$.*

Cet énoncé s'applique en particulier aux opérateurs d'un espace L^∞ dans un espace $L^0(\Omega, \mu)$. En fait on étend facilement le résultat aux espaces $\mathscr{L}^\infty$ de [**5**], donc aux espaces $C(K)$: le Théorème 2 bis donne dans ce cas une généralisation du théorème de Grothendieck aux opérateurs de $C(K)$ dans $L^0(\Omega, \mu)$. (C'est bien une généralisation dans la mesure où tout espace L^1 peut se plonger dans un espace $L^0(\Omega, \mu)$.)

Le Théorème 2 bis permet de réduire certaines questions concernant des fonctions mesurables quelconques au cas de fonctions de carré intégrable. Par exemple,

COROLLAIRE 1 [**9**]. *Soit (X_n) une suite d'éléments de $L^0(\Omega, \mu)$ telle que $\sum X_n$ soit inconditionnellement convergente en probabilité (c'est-à-dire que $\sum X_n$ converge en probabilité pour toute suite de signes $\varepsilon_n = \pm 1$). Il existe $Z \in L^0(\Omega, \mu)$ tel que $Z^{-1}X_n \in L^2(\Omega, \mu)$ pour tout n, et que $\sum Z^{-1}X_n$ converge inconditionnellement dans $L^2(\Omega, \mu)$.*

A partir de ce corollaire, on peut obtenir en utilisant un théorème classique de Menchov:

COROLLAIRE 2 [**6**, COROLLAIRE 96]. *Soit (X_n) une suite dans $L^0(\Omega, \mu)$, telle que $\sum X_n$ converge inconditionnellement en probabilité. La série $\sum X_n/\log(n+1)$ converge presque sûrement.*

Le résultat est le meilleur possible dans le sens suivant: Si (α_n) est une suite décroissante vers zéro telle que $\sum \alpha_n X_n$ converge presque sûrement dès que $\sum X_n$ converge inconditionnellement en probabilité, on a $\alpha_n = O(1/\log n)$ [**8**].

On peut se demander si le Théorème 2 admet une réciproque, à savoir:

Question 1. Est-il vrai que (D) ⇒ (C)?

D'après [**1**], la réponse est oui si l'espace E est un treillis normé, ou plus généralement un espace muni d'une structure locale inconditionnelle.

Dans le cas général, on montre dans [10] que: si E vérifie (D), il existe pour tout $q > 2$ une constante K_q telle que

$$(\sum \|x_n\|^q)^{1/q} \leqq K_q(\int \|\sum x_n \varepsilon_n(t)\|^2 \, dP(t))^{1/2}, \quad \forall (x_n) \in E.$$

Dans une seconde partie, nous allons nous intéresser au problème général de la factorisation d'un opérateur linéaire à travers un espace de Hilbert. Une condition nécessaire et suffisante a été donnée par Lindenstrauss et Pełczyński: Si (x_n) et (y_n) sont deux suites d'éléments de E, nous noterons $(x_n) < (y_n)$ si l'on a

$$\sum |\langle x_n, \xi\rangle|^2 \leqq \sum |\langle y_n, \xi\rangle|^2, \quad \forall \xi \in E'.$$

THÉORÈME 3 [5]. *Un opérateur linéaire u entre deux espaces de Banach E et F se factorise par un espace de Hilbert si et seulement s'il existe une constante C telle que*

$$(x_n) < (y_n) \Rightarrow \sum \|u(x_n)\|^2 \leqq C^2 \sum \|y_n\|^2, \quad \forall (x_n), (y_n) \in E.$$

Ce critère n'est pas facile à vérifier en général. Il permet néanmoins de déduire un résultat de S. Kwapien, que nous allons énoncer après quelques définitions.

Un espace E est dit de type 2 s'il existe une constante K telle que

$$(\int \|\sum x_n \varepsilon_n(t)\|^2 \, dP(t))^{1/2} \leqq K(\sum \|x_n\|^2)^{1/2}, \quad \forall (x_n) \in E.$$

On voit facilement que L^p, $2 \leqq p < \infty$, est de type 2. Plus généralement $L^p(E)$ est de type 2 si E est de type 2, $2 \leqq p < \infty$. Si E est de type 2, il en est de même de ses sous-espaces et de ses quotients, et le dual E' est de cotype 2. (Par contre l^1 est de cotype 2, mais l^∞ n'est pas de type 2.)

THÉORÈME 4 ([4] ET [11], EXPOSÉ VIII]). *Si E est de type 2 et F de cotype 2, tout opérateur linéaire continu de E dans F se factorise par un espace de Hilbert. En particulier, si E est de type 2 et de cotype 2, il est isomorphe à un espace de Hilbert.*

Ce théorème est à rapprocher du Théorème 2. En fait, on peut poser une question, dont la solution par l'affirmative impliquerait à la fois les Théorème 2 et 4:

Question 2. Si E' et F sont de cotype 2, est-il vrai que tout opérateur linéaire continu de E dans F se factorise par un espace de Hilbert?

Citons un cas particulier de ce problème: $E = A$, $F = B'$, où A et B sont deux C^*-algèbres. Plus particulièrement $E = \mathscr{L}(H)$, algèbre des opérateurs d'un espace de Hilbert, et $F = N(H)$, espace des opérateurs nucléaires d'un espace de Hilbert.

En modifiant très légèrement les arguments qui démontrent les Théorèmes 3 et 4, on obtient le résultat suivant:

THÉORÈME 5 [7]. *Soient E un espace de type 2, E_0 un sous-espace de E, F un espace de cotype 2. Tout opérateur linéaire continu u de E_0 dans F admet un prolongement linéaire continu $\tilde{u}$ de E dans F, tel que $\|\tilde{u}\| \leqq K\|u\|$, où K ne dépend que de E et F.*

COROLLAIRE. *Soit E un espace de type 2. Si F est un sous-espace de E isomorphe à un espace de Hilbert, il existe une projection π de E sur F, telle que $\|\pi\| \leqq Kd(F, H)$, où H désigne un espace de Hilbert de même dimension (finie ou infinie) que F, et où K ne dépend que de E.*

Ce dernier résultat généralise un théorème de Kadéc et Pełczynski [3], obtenu pour $E = L^p$, $2 \leqq p < \infty$.

Bibliography

1. E. Dubinsky, A. Pełczynski and H.P. Rosenthal, *On Banach spaces X for which* $\pi_2(\mathscr{L}_\infty, X) = B(\mathscr{L}_\infty, X)$, Studia Math. **44** (1972), 617–648.

2. A. Grothendieck, *Résumé de la théorie métrique des produits tensoriels topologiques*, Bol. Soc. Mat. São Paul **8** (1953), 1–79 (1956). MR **20** #1194.

3. M.I. Kadéc and A. Pełczynski, *Bases, lacunary sequences and complemented subspaces in the space* L_p, Studia Math. **21** (1961/62), 161–176. MR **27** #2851.

4. S. Kwapien, *Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients*, Studia Math. **44** (1972).

5. J. Lindenstrauss and A. Pełczynski, *Absolutely summing operators in* $\mathscr{L}_p$*-spaces and their applications*, Studia Math. **29** (1968), 275–326. MR **37** #6743.

6. B. Maurey, *Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces* L^p, Astérisque No. 11 (1973).

7. ———, *Un théorème de prolongement*, C. R. Acad. Sci. Paris **24** (1974), 329–332.

8. B. Maurey and A. Nahoum, *Applications radonifiantes dans l'espace des séries convergentes*, C.R. Acad. Sci. Paris Sér. A-B **276** (1973), A751–A754. MR **47** #9329.

9. B. Maurey and G. Pisier, *Un théorème d'extrapolation et ses consequénces*, C. R. Acad. Sci. Paris Ser. **277** (1974), 39–42.

10. ———, *Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach* (to appear).

11. *Séminaire Maurey-Schwartz* 1972–73, École Polytechnique, Paris.

12. *Séminaire Maurey-Schwartz* 1973–74, École Polytechnique, Paris.

13. H.P. Rosenthal, *On subspaces of* L^p, Ann. of Math. (2) **97** (1973), 344–373. MR **47** #784.

14. N. Tomczak-Jaegermann, *On the moduli of smoothness and the Rademacher averages of the trace classes* S_p, Studia Math. **50** (1974), 163–182.

CENTRE DE MATHÉMATIQUES DE L'ECOLE POLYTECHNIQUE
PARIS, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Normality of Products

Mary Ellen Rudin

By a *space*, we mean a *nondiscrete* Hausdorff space, and by a *map* we mean a continuous function. References to early papers can be found in the papers listed in the bibliography, particularly in [3].

Until paracompactness was defined, we essentially knew nothing about the normality of products. Sorgenfrey's half-open interval topology on the line was the first example of a normal (paracompact) space whose square is not normal. Michael gave a similar example of a metric space and a normal space whose product is not normal. Nonnormal products with one compact and one normal factor have also been known for many years. It was known that $\omega_1 \times (\omega_1 + 1)$ is not normal when Tamano proved that, for completely regular X, $X \times \beta X$ is normal if and only if X is paracompact. Dowker and Katětov independently proved that if I is the closed unit interval (or any compact metric space), then $I \times Y$ is normal if and only if Y is normal and countably paracompact. Extending this theorem, Morita proved that, for infinite cardinals λ, $I^\lambda \times Y$ is normal if and only if Y is normal and λ-paracompact. The problem of finding a Dowker space (normal but not countably paracompact) seemed important partly because of Borsuk's homotopy extension theorem which had been proved to hold for spaces Y where $I \times Y$ is normal.

Closed maps preserve normality, paracompactness, collectionwise normality, λ-paracompactness, and λ-collectionwise normality. For a given space X, let $\mathcal{N}(X)$ denote the class of all spaces Y such that $X \times Y$ is normal. Let $\mathcal{N}$ be the class of all spaces X such that $\mathcal{N}(X)$ is closed under closed maps. The class $\mathcal{N}(X)$ is trivially closed under perfect maps since, if $f: X \to Z$ is perfect, then $(f \times \mathrm{id}_Y) : (X \times Y) \to (Z \times Y)$ is perfect. Morita asked if all metric and all compact spaces belong to $\mathcal{N}$.

Four years ago in Nice, Nagami spoke on the normality of products [5]. He stressed the beautiful work which had been done in space classification, particularly

in the discovery of a number of useful classes which are preserved under countable products. He called for answers to Morita's questions as well as to others almost all of which have now been answered.

We know that there is a Dowker space [6]; this Dowker space is a subset of a box product of $\{\omega_n\}_{n\in\omega}$ and its cardinal functions are basically bounded below by $\aleph_\omega$. If there is a Souslin line, we know that there is a Dowker space of cardinality $\aleph_1$ which is hereditarily separable. A Souslin tree of cardinality λ, where λ is the successor of a regular cardinal, can be used to construct a Dowker space most of whose cardinal functions are $\leqq \lambda$.

QUESTION 1. *Is the existence of a separable (ccc, 1st countable, cardinality $\aleph_1$) Dowker space independent of the usual axioms for set theory*?

QUESTION 2. *Is there a Dowker space $X \times Y$ such that neither X nor Y is a Dowker space*?

Normality does not imply countable paracompactness nor vice versa. However [8] if X is a metric space and Y is a normal space, then $X \times Y$ is normal if and only if $X \times Y$ is countably paracompact; in addition if $X \times Y$ is normal, then $X \times Y$ is λ-paracompact (λ-collectionwise normal) if and only if Y is. In fact,

THEOREM 1 [8]. *Suppose that X is metric, C is compact, and Y is normal and λ-paracompact. Then the following are equivalent:*

(a) *$X \times Y$ and $X \times C$ are both normal.*

(b) *$X \times Y \times C$ is normal and countably paracompact.*

(c) *$X \times Y \times C$ is normal and λ-paracompact.*

This allows us to answer Morita's questions. Both the class of all metric spaces and the class of all compact spaces are contained in $\mathscr{N}$. In fact,

THEOREM 2 [8]. *If X is a metric space and C is a compact space, $X \times C \times Y$ is normal, and Z is the image of Y under a closed map, then $X \times C \times Z$ is normal.*

All of the questions about normality in products with a metric factor are tied to the countable paracompactness of the *product*. However, when one looks at normality questions for products with a compact factor X, the basic requirement lies between the $w(X)$-collectionwise normality and the $w(X)$-paracompactness; neither condition is both necessary and sufficient in all cases. By $w(X)$ we mean the weight of X or the minimal cardinality of a basis for X.

THEOREM 3 [7], [10]. *Suppose that X is a compact space and that Y is a normal space. If $X \times Y$ is normal, then Y is $w(X)$-collectionwise normal. If Y is $w(X)$-paracompact, then $X \times Y$ is normal.*

Necessary and sufficient conditions for the product of a compact space and a normal space to be normal must of necessity be complicated, but one can give such conditions which together with the following basic lemma are sufficient to prove that all compact spaces belong to $\mathscr{N}$.

THEOREM 4 [7]. *Assume that λ is a cardinal, that Y is a space which is normal and α-collectionwise normal for all $\alpha < \lambda$, that $\mathscr{G}$ is an open cover of Y of cardinality λ,*

and that $\mathscr{H}$ is a hereditarily closure-preserving closed refinement of $\mathscr{G}$. Then $\mathscr{G}$ has a locally finite refinement.

Starbird proves [8] that $\mathscr{N}$ is not the class of all spaces. However the following basic questions remain unanswered:

QUESTION 3. *Is there a paracompact (or collectionwise normal) space not in $\mathscr{N}$?*

QUESTION 4. *Is there a paracompact p-space not in $\mathscr{N}$?*

The behavior of products with a compact factor leads one to a theory of *test spaces*; a space X is a test space for property P provided a space Y has property P if and only if $X \times Y$ is normal. Besides I^λ, $\lambda + 1$ is a test space for λ-paracompactness. The one-point compactification of a discrete set of cardinality λ is a test space for λ-collectionwise normality [10].

Starbird [9] and Morita [4] independently discovered the following remarkable theorem.

THEOREM 5. *If C is a closed subset of a normal space X, A is any compact or metrizable absolute neighborhood retract, and $f: (C \times I) \cup (X \times \{0\}) \to A$ is a map, then there exists a map extending f to $X \times I$.*

Thus the binormal hypothesis in Borsuk's homotopy extension theorem is unnecessary! Both Starbird and Morita also discovered a related extension theroem:

THEOREM 6. *If X is a compact space, then any map from a closed subset of $C(X)$ into a $w(X)$-collectionwise normal space has an extension to $C(X)$.*

Still unsolved is:

QUESTION 5. *If X is normal, C is a closed subset of X, and $f:(C \times I) \cup (X \times \{0\}) \to Y$ is continuous, can f be extended to $X \times I$ if Y is an* ANR *(normal)?*

The example of a Dowker space in [6] is a subset of a box product. Four years ago we knew nothing about the normality or paracompactness of any box product of infinitely many spaces. For the next theorems we assume that X is a box product of a family $\{X_n\}_{n\in\omega}$ of (nondiscrete) topological spaces. The fact that all of the positive theorems are consistency results and that we have no positive theorems for box products of uncountably many spaces is unfortunate. However, the progress is tremendous even so.

THEOREM 7 [1]. *If X_0 is the set of all irrational numbers with the usual topology, and $X_n = \omega_0 + 1$ for all $n > 0$, then X is not normal.*

So having all factors metric is not sufficient to ensure normality.

THEOREM 8 [2]. *If each $X_n = 2^{(c^+)}$, then X is not normal.*

So having all factors compact is not sufficient to ensure normality. However:

THEOREM 9 [2], [11]. *The continuum hypothesis implies that X is paracompact if each X_n is a compact space which is either scattered or of weight $\leqq c$.*

Compact in this theorem can be replaced by σ-compact and paracompact.

If one assumes the generalized continuum hypothesis, one can decide whether

any given box product of ordinals is normal [3], [12]. Assuming Martin's axiom rather than the continuum hypothesis, one can still prove Theorem 9 if *1st countable* is added to the hypothesis [3].

QUESTION 6. *Is the box product of uncountably many copies of I normal?*

QUESTION 7. *Can the set theoretic assumptions be removed from Theorem 9?*

I conjecture that the answer to Questions 6 and 7 is *no*, but that the answer to the remaining questions is *yes*.

Bibliography

1. E. K. van Douwen, *The box product of countably many metrizable spaces need not be normal*, Fund. Math. (to appear).

2. K. Kunen, *Box products of compact spaces* (to appear).

3. ———, *On normality of box products of ordinals* (to appear).

4. K. Morita, *On generalizations of Borsuk's homotopy extension theorem*, Fund. Math. (to appear).

5. K. Nagami, *Normality of products*, Proc. Internat. Congress Math. (Nice, 1970), vol. 2, Gauthier-Villars, Paris, 1971, pp. 33–37.

6. M. E. Rudin, *A normal space X for which $X \times I$ is not normal*, Fund. Math. 73 (1971/72), no. 2, 179–186. MR **45** #2660.

7. ———, *The normality of products with one compact factor* (to appear).

8. M. E. Rudin and M. Starbird, *Products with a metric factor*, General Topology and Appl. (to appear).

9. M. Starbird, *The Borsuk homotopy extension theorem without the binormality condition*, Fund. Math. (to appear).

10. ———, *The normality of products with a compact factor* (to appear).

11. M. E. Rudin, *The box product of countably many compact metric spaces*, General Topology and Appl. **2** (1972), 293–298.

12. ———, *Countable box products of ordinals*, Trans. Amer. Math. Soc. **192** (1974), 121–128.

UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706, U.S.A.

Section 10

Operator Algebras, Harmonic Analysis and Representation of Groups

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Structure Theory for Type III Factors

A. Connes

Motivated by the study of the regular representation of nonunimodular locally compact groups J. Dixmier, in 1952, introduced quasi-Hilbert algebras. M. Tomita proved in 1967 that any von Neumann algebra M arises from a modular Hilbert algebra.[1] As shown then by F. Combes and M. Takesaki, each weight[2] φ on M gives rise to a modular Hilbert algebra giving back M, and in particular to a positive operator Δ_φ (the modular operator) and a one-parameter group of automorphisms of M: σ^φ (the modular automorphism group). See [7].

The formulas

(A) $$r_\infty(M) = \bigcap \operatorname{Sp} \Delta_\varphi,^3$$

(B) $$\rho(M) = \{e^{-2\pi/T},\ \exists\, \varphi \text{ with } \sigma_T^\varphi = 1\}$$

relating the above objects to the Powers, Araki, Woods, Krieger classification [1] if M is an infinite tensor product of factors of type I drove us to study the two invariants:

$$S(M) = \bigcap \operatorname{Sp} \Delta_\varphi, \qquad T(M) = \{T \in \boldsymbol{R},\ \exists\, \varphi,\ \sigma_T^\varphi = 1\}$$

for arbitrary type III factors.

It was essential, in this respect, to determine how the modular automorphism group σ^φ depends on the choice of φ. The answer [2] is summarised in:

(1°) For any weight φ on the von Neumann algebra M there exists a unique uni-

[1] Hence a quasi-Hilbert algebra.

[2] We mean a faithful semifinite normal weight.

[3] V. Ya. Golodets reached a formula very close to (A); see [2] for bibliography.

tary cocycle[4] $t \to \mu_t$ (μ denotes $(D\psi:D\varphi)$ because it is a "Radon-Nikodym" derivative) such that:

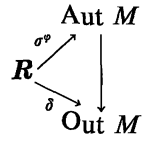
$$\sigma_t^{\psi}(x) = \mu_t \sigma_t^{\varphi}(x)\mu_t^* \qquad \forall\, x \in M, \forall\, t \in \boldsymbol{R}$$

and

$$\psi(x) = \varphi(x\mu_{-i}) \qquad \forall\, x \in M_+.^5$$

(2°) For any unitary cocycle $t \to \mu_t$ there exists a unique weight ψ on M such that $(D\psi:D\varphi)_t = \mu_t\ \forall\, t \in \boldsymbol{R}$.

Hence there exists an abstract kernel δ, homomorphism from $\boldsymbol{R}$ to the center of Out M = Aut M/Int M which characterises the automorphism groups σ^{φ} by the commutativity of the diagram:



Moreover $T(M) = \operatorname{Ker}\delta$ and $S(M) \cap \boldsymbol{R}_+^*$ is the spectrum (in the sense of [2]) of δ provided M is a factor. It follows that both are groups,[6] that when $S(M) \neq \{0, 1\}$ $T(M)$ is the orthogonal of $S(M) \cap \boldsymbol{R}_+^*$, while when $S(M) = \{0, 1\}$, it can be any denumerable subgroup of $\boldsymbol{R}$.

Also T and S become easy to compute; for instance, when M is the factor arising from an ergodic action of a group, they are related by formulas (A), (B) to the invariants r and ρ introduced in ergodic theory by W. Krieger.

In particular this showed that property L_λ of Power is not equivalent to L'_λ of Araki [1], so that in general S and r_∞ are different invariants. But let us enter in the details of the classification: Any factor of type III belongs to one of the following three classes:

III_λ $\lambda \in]0, 1[$ meaning $S(M) = \{\lambda^n, n \in \boldsymbol{Z}\}^-$,

III_0 i.e., $S(M) = \{0, 1\}$, and

III_1 i.e., $S(M) = [0, +\infty[$.

For factors M of type III_λ, $\lambda \in]0, 1[$, we have:

There exist maximal subalgebras[7] of type II_∞ of M. Let N be a maximal II_∞ subalgebra of M; then it is a factor and M is generated by N and a unitary U in the normaliser of N such that $\tau(UxU^*) = \lambda\tau(x)$, $\forall\, x \in N_+$, $\forall\, \tau$ normal trace on N. Let $M = N_1(U_1) = N_2(U_2)$ be two decompositions of M as above; then there is an inner automorphism ϕ of M such that $\phi(N_1) = N_2, \phi(U_1) = U_2$. This decomposition allows us to translate most of the problems on M in terms of N which is a simpler object. For instance any normal state φ on M is unitarily equivalent to a state $\psi \circ E$

[4]For each $t \in \boldsymbol{R}$, μ_t is a unitary in M, the map $t \to \mu_t$ is continuous and satisfies $\mu_{t_1+t_2} = \mu_{t_1}\sigma_{t_1}^{\varphi}(\mu_{t_2})$ $\forall\, t_1, t_2 \in \boldsymbol{R}$.

[5]For a precise statement see [2]

[6]For S this property was proven in collaboration with van Daele.

[7]Here by "subalgebra" we mean a von Neumann subalgebra range of a normal conditional expection of M.

where ϕ is a normal state on N and E is the unique expectation of M onto N.

The cross product $N(\theta)$ of a factor N of type II_∞ by an automorphism θ multiplying traces by λ is a factor of type III_λ, and any factor of type III_λ is obtained this way, with $N_1(\theta_1)$ isomorphic to $N_2(\theta_2)$, if and only if there exists[8] an isomorphism π of N_1 on N_2 such that $\pi\theta_1\pi^{-1} = \theta_2$. There are factors N of type II_∞ and automorphisms θ_1, θ_2 of N multiplying the traces by the same number $\lambda \in]0, 1[$ though they do not belong to the same conjugacy class in Aut N.

Factors of type III_0 appear as a limiting case of the type II_∞. Let M be of type III_0; then it is the cross product of a von Neumann algebra of type II_∞ by the infinite coproduct of groups of two elements. Moreover any subalgebra[9] of type II_∞ of M is the first element of an increasing sequence N_j of von Neumann subalgebras of type II_∞ with $\overline{\bigcup N_j} = M$. Let N be a von Neumann algebra of type II_∞, and $\theta \in \mathrm{Aut}\ N$ be a strict contraction with respect to some trace (see [2]) and strictly ergodic on the center C of N. Then the cross product $N(\theta)$ is a factor of type III_0. Any factor of type III_0 arises this way and $N_1(\theta_1)$ is isomorphic to $N_2(\theta_2)$ iff there are nonzero projections $e_j \in C_j$ such that the automorphisms θ_j, e_j induced by θ_j on e_j in the sense of Kakutani are the same.

Using this and the previous work of W. Krieger on automorphisms which are not of infinite product type, one gets a hyperfinite factor which is not an infinite tensor product of type I factors [2]. Now starting from a discrete decomposition $M = N(\theta)$ as above of a factor of type III_0 and building the flow on θ under the function $d\tau/d\tau\circ\theta$,[10] one obtains a one-parameter group $(\alpha_\lambda)_{\lambda\in\mathbf{R}^*_+}$ of automorphisms of a von Neumann algebra p of type II_∞ yielding the decomposition of M as a continuous cross product given by M. Takesaki [8] which this time is unique.

His duality technique allowed him to prove the following final results [8]:

Factors of type III_1. Let N be a factor of type II_∞, $(\theta_t)_{t\in\mathbf{R}}$ a one-parameter group of automorphisms of N, with $\tau\circ\theta_t = e^{-t}\tau$ for any normal trace τ; then the continuous cross product of N by $(\theta_t)_{t\in\mathbf{R}}$ is a factor of type III_1. Any factor of type III_1 arises this way and the decomposition is unique (as for factors of type III_λ).

However the appearance of continuous cross products complicates the study of M. For instance, though on factors of type III_λ, $\lambda \neq 1$, any normal state has a centraliser containing a maximal abelian subalgebra of M, it fails for factors of type III_1. Also, using the closure of the range of the modular homomorphism δ in Aut $M/\overline{\mathrm{Int}}\ M$ as a Galois group of M, one can exhibit factors of type III_1 which admit no decompositions as cross products of semifinite von Neumann algebras by discrete abelian groups [4].

Insofar as factors of type III_λ, $\lambda \in]0, 1[$, are very simple to analyse (for instance they exhibit strongly their nonnormalcy: the relative commutant of a maximal II_∞ subalgebra is reduced to the scalars), it is often helpful, when trying to prove a general property of type III factors to begin by the case III_λ, $\lambda \in]0, 1[$, and then

[8]The theory of factors of type III_λ as presented in [2] was complete in the spring of 1972; the improvement on [2] on uniqueness was obtained in collaboration with M. Takesaki [5].

[9]Here by "subalgebra" we mean a von Neumann subalgebra range of a normal conditional expectation of M.

[10]τ is a trace contracted by θ; for details see [2].

consider the case III_1 as a limit when $\lambda \to 1$. The final proof is independent of the classification (cf. [2] for the nonnormalcy of the III's).

One of the important effects of Tomita's theory is to yield the right generalisations to type III of notions which existed only for semifinite algebras. For instance, the only Hilbert-Schmidt operator on a factor of type III being 0, it seemed difficult to have an interesting generalisation of the cone of positive Hilbert-Schmidt operators. Using the modular operator, one can construct pre-Hilbert space structures s on M for which the completion of M_+ is a self-dual cone; the cone obtained is then independent of s,[11] and reduces to the Hilbert-Schmidt cone when M is semifinite.[12]

Moreover the class of cones thus obtained is exactly the class of convex self-dual cones V in Hilbert space H satisfying the two following conditions: V is *complex*, i.e., the quotient by its center of the Lie algebra $D(V) = \{\delta,\ \delta \in \mathscr{L}(H),\ e^{t\delta}V = V,\ \forall\, t\}$ is a *complex* Lie algebra for some complex structure I. V is *facially homogeneous*, i.e., for each face F of V the operator $P_F - P_{F^\perp}$ belongs to $D(V)$ where P_F means the orthogonal projection on the linear span of F and $F^\perp$ the face orthogonal to F.

There is a striking analogy between the relations type II and type III and the relations between unimodular and nonunimodular locally compact groups. Compare for instance what is said above for the case III_λ, $\lambda \in]0, 1[$, with the elementary description of a locally compact group G whose module Δ_G has range $\{\lambda^n, n \in \boldsymbol{Z}\}$, as a cross product of a unimodular group (the kernel of Δ_G) by a single automorphism.

If one enlarges the notion of locally compact group to include Mackey's virtual groups, one has still a notion of module Δ of G and of left regular representation of G; it generates a von Neumann algebra $U(G)$. Then the closure of the range of the module is a virtual subgroup of $\boldsymbol{R}_+^*$ and the following is true when G is a *principal* virtual group:[13]

If $\overline{\Delta(G)}$ is a usual subgroup of $\boldsymbol{R}_+^*$, it coincides with $S(U(G))$ (and in particular with $r(G)$, the ratio set of Krieger); otherwise $U(G)$ is of type III_0, and the strictly ergodic action of $\boldsymbol{R}_+^*$ corresponding to $\overline{\Delta(G)}$ is nothing but the flow described above for factors of type III_0.

In [5] we study the virtual modular spectrum $S_V(M)$ for arbitrary factors and show, for instance, by the formula $S_V(M_1 \otimes M_2) = \overline{S_V(M_1) \cdot S_V(M_2)}$ how much it behaves like the closure of the range of a "module" of M.

We cannot end without quoting the beautiful result of W. Krieger on weak equivalence [6]. With the above terminology it shows that the virtual modular spectrum S_V is a complete invariant for the class of factors which arise from ergodic transformations,[14] and can assume as value any virtual subgroup of $\boldsymbol{R}_+^*$.

[11] For a precise statement see [3].

[12] This generalisation was obtained independently by S. L. Woronowicz, H. Araki and myself. See [3] for bibliography.

[13] $U(G)$ is then a factor.

[14] By the group measure space construction.

This shows how important it is to decide whether any hyperfinite factor arises from an ergodic transformation and in particular whether the hyperfinite factor of type II_∞ is unique.

References

1. H. Araki, *Some topics in the theory of operator algebras*, Proc. Internat. Congress Math. (Nice, 1970), vol. 2, Gauthier-Villars, Paris, 1971, pp. 379–382.

2. A. Connes, *Une classification des facteurs de type* III, Ann. Sci. École Norm. Sup. (4) **6** (1973), 133.

3. ———, *Caractérisation des espaces vectoriels ordonnés sous jacents aux algèbres de von Neumann*, Ann. Inst. Fourier (Grenoble) **24** (1974), fasc. 4.

4. ———, *Almost periodic states and factors of type* III_1, J. Functional Analysis June 1974.

5. A. Connes and M. Takesaki, *The flow of weights of type* III *factors* (to appear).

6. W. Krieger, *On the weak equivalence problem of ergodic groups* (to appear).

7. M. Takesaki, *One parameter automorphism groups and states of operator algebras*, Proc. Internat. Congress Math. (Nice, 1970), vol. 2, Gauthier-Villars, Paris, 1971, pp. 427–432.

8. ———, *Duality in cross products and the structure of von Neumann algebras of type* III, Acta Math. **131** (1973), 249–310.

CNRS
VAL D'OISE, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Inversion Formula and Invariant Differential Operators on Solvable Lie Groups

Michel Duflo

In this article, I show some applications of Kirillov's method of orbits to the harmonic analysis of bi-invariant differential operators on Lie groups. For instance, we obtain the Plancherel formula for an exponential solvable group, and the fact that on such a group a bi-invariant differential operator has a fundamental solution. We prove that, on a solvable group, a bi-invariant differential operator is locally solvable. This is a joint work with Mustapha Raïs.

1. Notations. We consider a connected Lie group G with Lie algebra $\mathfrak{g}$. We choose a Haar measure dX on $\mathfrak{g}$ and denote by dg the associated left Haar measure on G. We define a positive function j on $\mathfrak{g}$ by $j(X)^2 = d(\exp X)/dX$. We denote by Δ the modular function on G: $\Delta(g) = d(hg)/dh$. We denote by $U(\mathfrak{g})$ the complex enveloping algebra of $\mathfrak{g}$, by $Z(\mathfrak{g})$ its center, by $Z'(\mathfrak{g})$ the set of semi-invariants in $U(\mathfrak{g})$. We denote by $S(\mathfrak{g})$ the complex symmetric algebra of $\mathfrak{g}$, by $I(\mathfrak{g})$ the subalgebra of invariants.

Let $\mathfrak{g}^*$ be the dual space of $\mathfrak{g}$. It is identified to the dual group of $\mathfrak{g}$ by the pairing $\exp(i\langle f, X\rangle)$. If m is a bounded measure on $\mathfrak{g}$, we put $\hat{m}(f) = \langle m, e^{if}\rangle$. If n is a bounded measure on $\mathfrak{g}^*$, we put $\check{n}(X) = \langle n, e^{iX}\rangle$. We denote by df the dual Haar measure on $\mathfrak{g}^*$.

Let $f \in \mathfrak{g}^*$. We denote by G_f the stabilizer of f in G, and by $\mathfrak{g}_f$ its Lie algebra. Then $\mathfrak{g}_f$ is the kernel of the 2-form B_f on $\mathfrak{g}$ defined by f. Let $\Omega \subset \mathfrak{g}^*$ be a G-orbit. We denote by β_Ω the canonical invariant measure on Ω, normalized as in [1, p. 20].

2. Definition of the mapping a. We fix a connected G-invariant open neighborhood V of 0 in $\mathfrak{g}$ such that exp is a diffeomorphism of V onto $W = \exp(V)$. Let ϕ be a distribution on W. We define a distribution $a(\phi)$ on V by the formula

$$\langle a(\varphi), j\varphi \circ \exp \rangle = \langle \phi, \varphi \rangle \quad \text{for all } \varphi \in C_c^\infty(W).$$

When G is semisimple (and V completely invariant), Harish-Chandra [5] proved the following result:

(1) $a(u * \phi) = a(u) * a(\phi)$ for all $u \in Z(\mathfrak{g})$ and all central distributions ϕ on W.

This paper would be greatly simplified if we knew this result for all Lie groups. Suppose G is solvable. Then I proved in [3] that (1) is true if ϕ is in $Z(\mathfrak{g})$. We shall prove below that (1) is also true for some distributions associated with orbits of G in $\mathfrak{g}^*$.

3. Definition of some distributions. We fix a G-orbit Ω in $\mathfrak{g}^*$ of maximal dimension, a character χ of G with values in $\mathbf{C}^*$, and a function ψ on Ω such that $\psi(g^{-1}f) = \chi(g)\psi(f)$ for all $f \in \Omega$ and $g \in G$. We shall say that ψ has weight χ. We assume that the following is verified:

There exist a positive N and a norm $\|\cdot\|$ on $\mathfrak{g}^*$ such that

$$\int_\Omega (1 + \|f\|)^{-N} |\psi(f)| \beta_\Omega(df) < \infty. \tag{2}$$

For $\varphi \in C_c(W)$, we put

$$\langle \phi(\Omega, \psi), \varphi \rangle = \int_\Omega (j\varphi \circ \exp)^\wedge(f)\psi(f)\beta_\Omega(df). \tag{3}$$

Thus, $\phi(\Omega, \psi)$ is a distribution on W, and $a(\phi(\Omega, \psi))$ is the restriction to V of the Fourier transform $(\psi\beta_\Omega)^\sim$ of $\psi\beta_\Omega$. Under the adjoint representation, $\phi(\Omega, \psi)$ is semi-invariant with weight $\Delta\chi$.

4. The distributions $\phi(\Omega, \varphi)$ on solvable groups. We assume that G is a solvable connected simply connected Lie group. We consider an integral G-orbit $\Omega \subset \mathfrak{g}^*$: This means there exists a unitary character η of G_f (f is some point of Ω) whose differential is the restriction of if to $\mathfrak{g}_f$. Fix such a character η. We denote by $\pi_{\Omega,\eta}$ the irreducible unitary class of representations of G associated to these data by Auslander and Kostant, with the normalization of [1, p. 217].

Let χ be a character of G with values in $(0, \infty)$, and ψ a positive function of weight χ on Ω. There is a canonical way to associate to ψ a selfadjoint positive (in general unbounded) operator A_ψ in the space of $\pi_{\Omega,\eta}$ such that

$$\pi_{\Omega,\eta}(g)A_\psi\pi_{\Omega,\eta}(g^{-1}) = \chi(g)A_\psi \quad \text{for all } g \in G.$$

Suppose moreover that (2) is verified and that Ω has maximal dimension. Then, if $\varphi \in C_c^\infty(G)$, the operator $A_\psi^{1/2}\pi_{\Omega,\eta}(\varphi)A_\psi^{1/2}$ can be extended to a trace class operator, and we have

$$\operatorname{tr}(A_\psi^{1/2}\pi_{\Omega,\eta}(\varphi)A_\psi^{1/2}) = \langle \phi(\Omega, \psi), \varphi \rangle \quad \text{for all } \varphi \in C_c^\infty(W). \tag{4}$$

(This is an extension of [1, Chapter IX].)

Let $u \in Z(\mathfrak{g})$. Then $a(u)^\wedge$ is an invariant polynomial on $\mathfrak{g}^*$. We denote by $a(u)^\wedge(\Omega)$ its constant value on Ω. We have

$$\pi_{\Omega,\eta}(u) = a(u)^\wedge(\Omega)\mathrm{Id}. \tag{5}$$

If ϕ is a distribution on G, put $d\check{\phi}(g) = d\phi(g^{-1})$. Formulas (4) and (5) imply

$$\check{u} * \phi(\Omega, \psi) = a(u)^{\wedge}(\Omega)\phi(\Omega, \psi) \tag{6}$$

which is equivalent to

$$a(\check{u} * \phi(\Omega, \psi)) = a(\check{u}) * a(\phi(\Omega, \psi)). \tag{7}$$

When (2) is verified, but Ω not integral, there is probably a formula analogous to (4) involving instead of $\pi_{\Omega,\eta}$ one of the factorial normal representations constructed by Pukańszky [7].

5. Harmonic analysis of the distributions $u^{(s)}$ ($u \in Z(\boldsymbol{g})$). We do not suppose for the moment that G is solvable. We suppose there is a G-invariant subset $\boldsymbol{g}_r^*$ of $\boldsymbol{g}^*$ which is the union of locally closed orbits of maximal dimension, and such that $\boldsymbol{g}^* - \boldsymbol{g}_r^*$ is of Lebesgue measure 0. Choose a nonzero rational function θ' on $\boldsymbol{g}^*$, semi-invariant with weight Δ^{-1} (cf. [2]) and put $\theta = |\theta'|$. The measure $\theta(f)^{-1}df$ on $\boldsymbol{g}^*$ is G-invariant. The quotient space $X = \boldsymbol{g}_r^*/G$ is a standard Borel space. There exists a positive Borel measure m_θ on X such that

$$\int_{\boldsymbol{g}^*} h(f)\, df = \int_X m_\theta(d\Omega) \int_\Omega h(f)\theta(f)\beta_\Omega(df) \tag{8}$$

for all positive Borel functions on $\boldsymbol{g}^*$.

Apply (8) to the function $h(f) = (1 + \|f\|)^{-N}$, where $N = 1 + \dim \boldsymbol{g}$. Fubini's theorem shows that we have

$$\int_\Omega (1 + \|f\|)^{-N}\theta(f)\beta_\Omega(df) < \infty, \tag{9}$$

for almost all $\Omega \in X$. Then, the distribution $\phi(\Omega, \theta)$ is defined on W for almost all $\Omega \in X$. Notice that these distributions are central. We shall use them to decompose other interesting central distributions on W.

Let $p \in S(\boldsymbol{g})$ be such that $\hat{p}$ is a positive function on $\boldsymbol{g}^*$. For $s \in \boldsymbol{C}$ with $\mathrm{Re}(s) \geqq 0$, we denote by p^s the tempered distribution on $\boldsymbol{g}$ whose Fourier transform is $\hat{p}^s$. Then, by the Atiyah-Bernstein theorem, p^s can be extended to a meromorphic function of s defined in $\boldsymbol{C}$ with values in the space of tempered distributions. Let $u \in Z(\boldsymbol{g})$ such that $a(u)^{\wedge}$ is positive. We shall denote by $u^{(s)}$ the distribution on W such that $a(u^{(s)}) = a(u)^s$. Thus $u^{(s)}$ is a meromorphic function of $s \in \boldsymbol{C}$ with values in the space of central distributions on W.

From (8) and (3), we get the following formula: Let $u \in Z(\boldsymbol{g})$ such that $a(u)^{\wedge}$ is positive. Suppose $\mathrm{Re}(s) \geqq 0$. Let $\varphi \in C_c(W)$. We have

$$\langle u^{(s)}, \varphi\rangle = \int_X m_\theta(d\Omega)a(\check{u})^{\wedge}(\Omega)^s\langle\phi(\Omega, \theta), \varphi\rangle. \tag{10}$$

When $u = \delta$, (10) gives a local inversion formula:

$$\varphi(1) = \int_X m_\theta(d\Omega)\langle\phi(\Omega, \theta), \varphi\rangle. \tag{11}$$

6. Solvability of bi-invariant differential operators on solvable Lie groups. In this section, G is a simply connected solvable Lie group.

We have the following result. Let $u \in Z(\mathfrak{g})$ be such that $a(u)^\wedge$ is positive, and let $v \in Z(\mathfrak{g})$. We have the equality of meromorphic functions:

$$a(v * u^{(s)}) = a(v) * a(u^{(s)}). \tag{12}$$

To prove (12), we imbed G in a simply connected group $\tilde{G}$ corresponding to an algebraic envelope $\tilde{\mathfrak{g}}$ of $\mathfrak{g}$, and remark that we can assume $G = \tilde{G}$. Then all orbits of G in $\mathfrak{g}^*$ are integral and locally closed. Then (12) follows from (7) and (10).

It follows from (12) that the constant term of the Laurent expansion of $u^{(s)}$ at $s = -1$ is a distribution E on W which verifies $u * E = E * u = \delta$ (the Dirac mass at 1). From this, it is easy to obtain the following theorem.

THEOREM 1. *Let G be a connected solvable Lie group. Keep the notations of* §§1 *and* 2. *Let $u \in Z'(\mathfrak{g})$. There exists a distribution E on W such that $u * E = E * u = \delta$, and such that, for all $v \in Z'(\mathfrak{g})$, $a(v * E) = a(v) * a(E)$.*

When G is an exponential solvable group, we may choose $W = G$. In this case u has a fundamental solution defined in all of G. For G simply connected nilpotent, this is due to Raïs [8].

We do not know any example of a simply connected solvable group G, and $u \in Z'(\mathfrak{g})$, without a global fundamental solution.

7. The Plancherel formula for an exponential solvable group. In this section, G is an exponential solvable group. In this case, it is known that all orbits in $\mathfrak{g}^*$ are locally closed and simply connected. To each orbit is associated one unitary irreducible class π_Ω of representations of G. Choose the function θ on $\mathfrak{g}^*$ as in §5. Denote by $A_{\theta,\Omega}$ the positive selfadjoint operator in the space of π_Ω associated with the restriction of θ to Ω. Let $\mathfrak{g}_r^*$ be the union of orbits of maximal dimension in $\mathfrak{g}^*$. Define X and m_θ as in §5. From (11) and (4), we get the following inversion formula.

For almost all $\Omega \in X$, the operator $A_{\theta,\Omega}^{1/2} \pi_\Omega(\varphi) A_{\theta,\Omega}^{1/2}$ extends to a trace class operator for all $\varphi \in C_c^\infty(G)$, and we have

$$\varphi(1) = \int_X m_\theta(d\Omega) \operatorname{tr}(A_{\theta,\Omega}^{1/2} \pi_\Omega(\varphi) A_{\theta,\Omega}^{1/2}). \tag{13}$$

For each $\Omega \in X$, realize π_Ω in some Hilbert space H_Ω, and denote by $L_2(H_\Omega)$ the Hilbert space of Hilbert-Schmidt operators on H_Ω. From π_Ω, we get an irreducible representation of $G \times G$ in $L_2(H_\Omega)$.

THEOREM 2. *Let $\varphi \in L_1(G) \cap L_2(G)$. For almost all $\Omega \in X$, the operator $\pi_\Omega(\varphi) A_{\theta,\Omega}^{1/2}$ extends to an element $[\pi_\Omega(\varphi) A_{\theta,\Omega}^{1/2}]$ of $L_2(H_\Omega)$. The mapping $\varphi \to \{[\pi_\Omega(\varphi) A_{\theta,\Omega}^{1/2}]\}_{\Omega \in X}$ extends to an isometry U of $L_2(G)$ onto $\int_X L_2(H_\Omega) m_\theta(d\Omega)$. The isometry U intertwines the representation of $G \times G$ in $L_2(G)$ (the double regular representation) and the representation of $G \times G$ in $\int_X L_2(H_\Omega) m_\theta(d\Omega)$.*

Theorem 2 is the Plancherel theorem for G. If G is unimodular, it is a well-known application of formula (13).

8. Square-integrable representations of exponential solvable groups. In this section, G is an exponential solvable group with center Z. We choose a left Haar measure on G/Z. Let π be a unitary irreducible representation of G in a Hilbert space H. It is square-integrable (mod Z) if and only if it occurs discretely in $L_2(G, \eta)$, where η is the restriction of π to Z, and $L_2(G, \eta)$ the space of the representation induced by η to G. Suppose π is square-integrable (mod Z). There is on H a positive selfadjoint operator K, called the formal degree, which verifies $\pi(g)K\pi(g)^{-1} = \Delta(g)^{-1}K$ for all $g \in G$, and such that

$$\int_{G/Z} |(k, \pi(g)h)|^2 \, dg = \|k\|^2 \, \|K^{-1/2}h\|^2$$

for all $k \in H$ and $h \in \operatorname{dom} K^{-1/2}$ (cf. [4]).

The extension of Theorem 2 to $L_2(G, \eta)$ gives the following result (which, if G is nilpotent, is due to Moore and Wolf [6]):

Let $\Omega \subset \mathfrak{g}^*$. The representation π_Ω is square-integrable (mod Z) if and only if $G_f = Z$ for all $f \in \Omega$. Let $e_1, \cdots, e_{2d}$ be a basis of $\mathfrak{g}/\mathfrak{z}$ such that the unit cube has volume 1. If $f \in \mathfrak{g}^*$, we denote by $\operatorname{discr}(f)$ the discriminant of the 2-form on $\mathfrak{g}/\mathfrak{z}$ deduced from B_f, relative to the basis $e_1, \cdots, e_{2d}$. Suppose π_Ω is square-integrable (mod Z). If $f \in \Omega$, put $\psi(f) = (2\pi)^{-d}|\operatorname{discr}(f)|$. Then the formal degree of π_Ω is the operator A_ψ (cf. §4).

References

1. P. Bernat et al., *Représentations des groupes de Lie résolubles*, Dunod, Paris, 1972.
2. J. Dixmier, M. Duflo and M. Vergne, *Sur la représentation coadjointe d'une algèbre de Lie* Composito Math. (1975) (to appear).
3. M. Duflo, *Caractères des groupes et des algèbres de Lie résolubles*, Ann. Sci. Ecole Norm. Sup. (4) **3** (1970), 23–74. MR **42** #4672.
4. M. Duflo and C. C. Moore, *On the regular representation of a non unimodular locally compact group* (to appear).
5. Harish-Chandra, *Invariant eigendistributions on a semisimple Lie group*, Trans. Amer. Math. Soc. **119** (1965), 457–508. MR **31** #4862d.
6. C. C. Moore and J. A. Wolf, *Square integrable representations of nilpotent groups*, Trans. Amer. Math, Soc. **185** (1973), 445–462.
7. L. Pukańszky, *Unitary representations of solvable Lie groups*, Ann. Sci. Ecole Norm. Sup. (4) **4** (1971), 457–608.
8. M. Raïs, *Solutions élémentaires des opérateurs différentiels bi-invariants sur un groupe de Lie nilpotent*, C. R. Acad. Sci. Paris Ser. A-B **273** (1971), A495-A498. MR **44** #6908.

UNIVERSITÉ DE PARIS VII
PARIS, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

A Szegö Kernel for Discrete Series

A. W. Knapp*

The Szegö kernel for the unit ball in C^m is a reproducing kernel that gives a formula for holomorphic functions in the ball in terms of their boundary values, namely

$$F(z) = \frac{(m-1)!}{2\pi^m} \int_{|\zeta|=1} \frac{f(\zeta)\, d\sigma(\zeta)}{(1 - \langle z, \zeta \rangle)^m}, \tag{1}$$

where f is the boundary function for F and $d\sigma$ is Lebesgue measure on the sphere. When $m = 1$, (1) easily transforms into the Cauchy integral formula. In dimension m, the formula extends to be defined on all f in L^2, always yielding holomorphic functions. If we identify holomorphic functions with their boundary values, the extended operator can be regarded as the orthogonal projection from L^2 to the holomorphic functions in L^2. This projection property characterizes the kernel.

In terms of semisimple Lie groups, functions on the sphere suggest nonunitary principal series representations and holomorphic functions on the ball suggest discrete series representations, and the Szegö kernel should suggest a map from the one to the other. Actually formula (1) will not arise with discrete series but with so-called limits of discrete series. For ordinary discrete series representations, we shall use operators that are more analogous to the formula for the $(n-1)$st complex derivative in the disc,

$$F(z) = \frac{(n-1)!}{2\pi i} \oint \frac{f(\zeta)\, d\zeta}{(\zeta - z)^n}, \qquad n > 1. \tag{2}$$

*Supported by NSF grant GP-42459. The author is pleased to acknowledge the extensive influence of conversations with K. Okamoto and early collaboration with E. M. Stein upon the developments in this paper.

Such operators are not projections but do carry general boundary functions into good functions with a known relationship to the original function.

We examine first the role of this mapping for the holomorphic discrete series of $G = \mathrm{SU}(1, 1) = \{\left(\begin{smallmatrix}\alpha & \beta \\ \bar{\beta} & \bar{\alpha}\end{smallmatrix}\right), \ |\alpha|^2 - |\beta|^2 = 1\}$. The holomorphic discrete series is a sequence of square-integrable unitary representations parametrized by an integer $n \geqq 2$. The space is

(3) $\{F(z) \text{ analytic in } \boldsymbol{C}^1 \text{ for } |z| < 1 \, | \int_{|z|<1} |F(z)|^2 (1 - |z|^2)^{n-2} \, dx \, dy < \infty\}$

with group action

$$U(g)F(z) = (\bar{\beta}z + \alpha)^{-n} F\Big(\frac{\beta + z\bar{\alpha}}{\alpha + z\bar{\beta}}\Big).$$

For the nonunitary principal series, let

$$A = \begin{pmatrix}\cosh r & \sinh r \\ \sinh r & \cosh r\end{pmatrix}, \quad N = \begin{pmatrix}1 + ix & -ix \\ ix & 1 - ix\end{pmatrix}, \quad K = \begin{pmatrix}e^{i\theta} & 0 \\ 0 & e^{-i\theta}\end{pmatrix}.$$

For a the indicated matrix in A, let $\lambda_s(a) = e^{r(s+1)}$ for complex s. Pick some complex s and one of "even" or "odd", say "even", for example. The space of a representation is the space of f defined on the circle K with just "even"-numbered Fourier coefficients. These functions extend uniquely to G by $f(ank) = \lambda_s(a)f(k)$, and the group action $U(g)f(x) = f(xg)$ preserves the space.

Fix $n \geqq 2$, use "even" or "odd" according to what n is, and put $s = \pm (n - 1)$. To fix the ideas, let us take n even. Then the nonunitary principal series representation is reducible. In terms of Fourier coefficients, the situation is as follows: At $s = -(n - 1)$, holomorphic discrete series number n arises as a quotient from coefficients $\{\cdots, -(n + 2), -n\}$, a finite-dimensional representation appears as an invariant subspace from coefficients $\{-(n - 2), -(n - 4), \cdots, (n - 2)\}$, and an antiholomorphic discrete series appears from coefficients $\{n, (n + 2), \cdots\}$. At $s = +(n - 1)$, the numbers are the same, but the roles of quotient and subspace are reversed.

The intertwining operators first constructed by Kunze and Stein [4] are equivariant maps of the representation at s to the one at $-s$, but at $s = -(n - 1)$ the operator has a simple pole. Work by Sally [8] yields an explicit formula for the residue operator in terms of Fourier coefficients, showing that the residue operator is equivariant and maps the nonunitary principal series representation to the sum of the two discrete series. If we compose with the projection to the holomorphic discrete series and reinterpret the principal series as suitable functions on $K/\boldsymbol{Z}_2$ and the discrete series in its form (3), we obtain (2) as the formula for the composition intertwining operator.

We shall attempt to extend these matters to arbitrary (not necessarily holomorphic) discrete series representations of the automorphism groups of noncompact hermitian symmetric spaces. Let G be a connected semisimple Lie group with finite center, let K be a maximal compact subgroup, and assume that G/K is hermitian. Let $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ be the corresponding Cartan decomposition of the Lie algebra of

G, and let $\mathfrak{h} \subseteq \mathfrak{k}$ be a maximal abelian subspace. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, and we let $\mathfrak{g}^C = \mathfrak{h}^C + \sum \mathfrak{g}_\alpha$ be the root-space decomposition. Here $\mathfrak{g}_\alpha = \mathbf{C}E_\alpha$, and we take the E_α to be normalized as in [1]. Introduce an ordering on the roots such that no sum of two noncompact positive roots is a root.

A parametrization of the discrete series was supplied by Harish-Chandra [3]. If we let ρ be half the sum of the positive roots and if, as we may, we assume that ρ is integral, then the parameter space is the set of integral forms $\Lambda + \rho$ on $\mathfrak{h}$ such that $\langle \Lambda + \rho, \alpha \rangle > 0$ for $\alpha > 0$ compact and $\langle \Lambda + \rho, \alpha \rangle \neq 0$ for $\alpha > 0$ noncompact. Since G/K is hermitian, it follows that every such Λ satisfies $\langle \Lambda, \alpha \rangle \geqq 0$ for $\alpha > 0$ compact.

The Langlands conjecture says that the above representation is realized in a certain $\bar{\partial}$-cohomology space, described as follows. Let $Q = \{\alpha > 0 \text{ noncompact} \mid \langle \Lambda + \rho, \alpha \rangle > 0\}$ and $q = |Q|$. To avoid bundles, introduce a symbol $\omega^{-\alpha}$ for each $\alpha > 0$, to be thought of as a $d\bar{z}$-type differential form. If $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_q)$ is any ordered q-tuple of arbitrary positive roots, let $\omega^{-\boldsymbol{\alpha}} = \omega^{-\alpha_1} \wedge \cdots \wedge \omega^{-\alpha_q}$. Consider expressions

$$F = \sum_{|\boldsymbol{\alpha}|=q} F_{\boldsymbol{\alpha}} \omega^{-\boldsymbol{\alpha}},$$

where each $F_{\boldsymbol{\alpha}}$ is a smooth scalar-valued function on G satisfying

$$F_{\boldsymbol{\alpha}}(hx) = \xi_{\Lambda+\alpha_1+\cdots+\alpha_q}(h) F_{\boldsymbol{\alpha}}(x) \quad \text{for } h \in \exp \mathfrak{h},\ x \in G.$$

(Here ξ denotes a character.) The $\bar{\partial}$-operator is given on functions (0-forms) by

$$\bar{\partial} F(x) = \sum_{\alpha>0} E_{-\alpha} F(x) \omega^{-\alpha},$$

where $E_{-\alpha}$ denotes the right-invariant differentiation computed as the real part plus i times the imaginary part. Extend $\bar{\partial}$ by making $\bar{\partial}^2 = 0$, and let $\bar{\partial}^*$ be the formal adjoint relative to $L^2(G)$. Take $\ker \bar{\partial} \cap \ker \bar{\partial}^*$ in dimension q as a dense subspace of a representation, and use the norm $(\sum \int_G |F_{\boldsymbol{\alpha}}|^2\, dx)^{1/2}$. The group operation is a right translation of the coefficient functions. The Langlands conjecture is the statement that the completion of this constructed representation is the discrete series representation with parameter $\Lambda + \rho$. The conjecture is known for $G = \mathrm{SU}(1, 1)$ by easy computation, for general G and $q = 0$ by work of Harish-Chandra [2], for $|\langle \Lambda + \rho, \alpha \rangle| > c$ whenever α is noncompact by Narasimhan and Okamoto [5], and for $\langle \Lambda + \rho, \alpha \rangle > 0$ whenever $\alpha > 0$ is noncompact by Parthasarathy [7]. The proofs by Narasimhan and Okamoto and Parthasarathy are not constructive.

For the nonunitary principal series of G, let $G = ANK$ be an Iwasawa decomposition, let $M = Z_K(A)$, let σ be an irreducible representation of the compact group M, let λ be a linear functional on the Lie algebra of A, and let ρ^+ be half the sum of the positive restricted roots counted with multiplicities. The space consists of vector-valued functions f on K with $f(mk) = \sigma(m) f(k)$, which are then extended to G by the definition $f(ank) = \exp((\lambda + \rho^+) \log a) f(k)$. The group G acts by right translation.

Problem. Give an integral formula for passing from appropriate nonunitary

principal series, realized as spaces of functions on K, to discrete series realized in the format of the Langlands conjecture.

Solving the problem would consist of four steps: (1) existence—producing an integral formula so that the members of the image satisfy $\bar{\partial} F = 0$ and $\bar{\partial}^* F = 0$, (2) finiteness—establishing the square-integrability of the K-finite members of the image, (3) uniqueness—proving irreducibility of the image, and (4) identification—computing the character of the image representation. We shall do (1) in general and (2) in some special cases. Steps (3) and (4) follow whenever (1) and (2) and the Langlands conjecture are known.

To define the Szegö kernel, let $\Lambda + \rho$ and Q and q be as before. The linear functional λ is yet to be specified, but we extend arbitrary *scalar-valued* functions f on K to G by $f(ank) = \exp((\lambda + \rho^+) \log a) f(k)$. We map a smooth f into $F = \sum_{|\alpha|=q} F_\alpha \omega^{-\alpha}$ with

$$(4) \qquad F_\alpha(x) = \int_K \overline{S_\alpha(k)}\, f(kx)\, dk = \int_K \overline{S_\alpha(\kappa(kx^{-1}))}\, \exp((\rho^+ - \lambda) H(kx^{-1}))\, f(k)\, dk$$

and $S_\alpha(k) = (\tau_\Lambda(k)\phi_\Lambda, \phi_\Lambda)(\mathrm{Ad}(k)E_\alpha, E_Q)$. Here ϕ_Λ is a highest weight vector, and E_α and E_Q denote alternating tensors. This is a group-theoretic generalization of the formula in SU(1, 1) except that the Ad factor was not present in dealing with 0-forms. The first part of (4) shows the map is equivariant, but the second is more useful in computations. The formula for F_Q has been considered by Okamoto [6].

Two comments are in order before we state precise theorems. First, no special M-dependence of f is assumed in (4). However expansion of f in Fourier series on M and a change of variables show that only one particular representation of M plays a role. Thus the domain can be regarded as a single nonunitary principal series, but with a finite multiplicity. Second, the choice of A in the Iwasawa decomposition is not arbitrary. [In fact, assume the Langlands conjecture. Then the infinitesimal character of the nonunitary principal series is determined in one way by that of the discrete series and in another way by the character of A and the placement of M in K. These cannot match unless M is placed in K properly.]

Thus we must define A. Any two A's are conjugate by a member of K. Define a standard A_0 by $\mathfrak{a}_0 = \sum \boldsymbol{R}(E_{\alpha_i} + E_{-\alpha_i})$, where α_i is defined inductively as the largest noncompact positive root orthogonal to $\alpha_1, \cdots, \alpha_{i-1}$, and use the basis $\{E_{\alpha_i} + E_{-\alpha_i}\}$ to define an ordering. Let $A = pA_0p^{-1}$, where p is a member of the Weyl group of K yet to be specified.

THEOREM 1. *If p and Q satisfy a compatibility condition* (∗), *then to each Λ corresponds a unique $\lambda = \lambda(\Lambda)$ such that the image $F = \sum F_\alpha \omega^{-\alpha}$ satisfies $\bar{\partial} F = 0$ and $\bar{\partial}^* F = 0$ for each f in $C^\infty(K)$.*

We shall state (∗) shortly. Our finiteness theorem is as follows.

THEOREM 2. *Let $G = \mathrm{SU}(1, 1)$, or let G be general but $q = 0$. If* (∗) *holds and if $\lambda = \lambda(\Lambda)$, then the image of the trigonometric polynomials under* (4) *is exactly the space of K-finite elements in the cohomology space, and the cohomology space is the discrete series representation with parameter $\Lambda + \rho$.*

Also, by a lengthy computation, we have established square-integrability of the image of $f = S_Q$ for SU(2, 1) when $q = 1, \lambda = \lambda(\Lambda)$, and p is trivial.

To state condition (*), let $\alpha(\gamma)$ be the function from positive noncompact roots to the α_i's defined above, given as the first α_i such that γ is not orthogonal to α_i. Then $\alpha(\gamma) - \gamma$ is always a positive compact root or 0. Hence $p\alpha(p^{-1}\gamma) - \gamma$ is always a compact root or 0.

Condition (*). Every positive noncompact root γ not in Q satisfies $p\alpha(p^{-1}\gamma) - \gamma \geqq 0$. Also every positive noncompact root γ in Q satisfies $p\alpha(p^{-1}\gamma) - \gamma \leqq 0$.

One can show case-by-case that to each Q corresponds some p such that p and Q satisfy (*).

EXAMPLE. $G = \mathrm{SU}(m, 1)$. One can arrange that the positive compact roots are $e_i - e_j$, $i < j \leqq m$, and the positive noncompact roots are $e_i - e_{m+1}$, $i \leqq m$. The element p is a permutation of $\{1, \cdots, m\}$. Then A_0 is built out of $E_{e_1 - e_{m+1}} + E_{-(e_1 - e_{m+1})}$, and pA_0p^{-1} necessarily is built out of $E_{e_l - e_{m+1}} + E_{-(e_l - e_{m+1})}$ for some l. The set Q has the form $\{e_i - e_{m+1}, i \leqq q\}$. If $0 < q < m$, two choices of l (namely $l = q$ and $q + 1$) are such that p and Q satisfy (*), and generally these lead to really distinct nonunitary principal series. If $q = 0$ or m, there is only one such choice (namely $l = 0$ or m, respectively). However, when $m = 2$ and $q = 1$, it is known that there are three nonunitary principal series, not two, with a given discrete series as quotient.

We conclude with the formula for $\lambda(\Lambda)$. Let

$$n_i^+ = \left|\{\gamma > 0 \text{ noncompact} \mid \gamma \in Q, \gamma \neq p\alpha_i, p\alpha(p^{-1}\gamma) = p\alpha_i\}\right|,$$

$$n_i^- = \left|\{\gamma > 0 \text{ noncompact} \mid \gamma \notin Q, \gamma \neq p\alpha_i, p\alpha(p^{-1}\gamma) = p\alpha_i\}\right|.$$

Then λ is determined by the values of $\lambda' = \rho^+ - \lambda$ on all $\mathrm{Ad}(p)(E_{\alpha_i} + E_{-\alpha_i})$. If $p\alpha_i$ is not in Q,

$$\lambda'(\mathrm{Ad}(p)(E_{\alpha_i} + E_{-\alpha_i})) = -\left|\alpha_i\right|^{-1}\sqrt{2}\,\langle \Lambda + Q - n_i^+ p\alpha_i, p\alpha_i\rangle.$$

If $p\alpha_i$ is in Q,

$$\lambda'(\mathrm{Ad}(p)(E_{\alpha_i} + E_{-\alpha_i})) = \left|\alpha_i\right|^{-1}\sqrt{2}\,\langle \Lambda + Q + n_i^- p\alpha_i, p\alpha_i\rangle.$$

References

1. P. Griffiths and W. Schmid, *Locally homogeneous complex manifolds*, Acta Math. **123** (1969), 253–302. MR **41** #4587.

2. Harish-Chandra, *Representations of semisimple Lie groups.* VI. *Integrable and square-integrable representations*, Amer. J. Math. **78** (1956), 564–628. MR **18**, 490.

3. ———, *Discrete series for semisimple Lie groups.* II. *Explicit determination of the characters*, Acta Math. **116** (1966), 1–111. MR **36** #2745.

4. R. A. Kunze and E. M. Stein, *Uniformly bounded representations and harmonic analysis of the* 2×2 *real unimodular group*, Amer. J. Math. **82** (1960), 1–62. MR **29** #1287.

5. M. S. Narasimhan and K. Okamoto, *An analogue of the Borel-Weil-Bott theorem for hermitian symmetric pairs of non-compact type*, Ann. of Math. (2) **91** (1970), 486–511. MR **43** #419.

6. K. Okamoto, *Harmonic analysis on homogeneous vector bundles*, Conference on Harmonic

Analysis (College Park, Maryland, 1971), Lecture Notes in Math., vol. 266, Springer-Verlag, Berlin and New York, 1972, pp. 255–271.

7. R. Parthasarathy, *A note on the vanishing of certain L^2-cohomologies*, J. Math. Soc. Japan **23** (1971), 676–691.

8. P. J. Sally, Jr., *Analytic continuation of the irreducible unitary representations of the universal covering group of* SL(2, *R*), Mem. Amer. Math. Soc. No. 69 (1967). MR **38** #3380.

CORNELL UNIVERSITY
ITHACA, NEW YORK 14853, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Operator Algebras and Their Abelian Subalgebras

J. R. Ringrose

1. Introduction. From time to time in the study of operator algebras a situation arises in which it is desirable to deduce a certain piece of information about an algebra $\mathfrak{A}$ from the same (or similar) information about every maximal abelian * subalgebra $\mathscr{A}$ of $\mathfrak{A}$. After a very brief description of several problems of this type, and of the progress that has been made towards their solution, I shall discuss two of them in somewhat more detail.

1.1. *Measures on projections.* By a *measure* on the set $\mathscr{P}$ of projections in a von Neumann algebra $\mathfrak{A}$ we mean a mapping $\mu: \mathscr{P} \to \boldsymbol{R}^+$, such that $\mu(\sum P_j) = \sum \mu(P_j)$ whenever (P_j) is an orthogonal family of projections in $\mathfrak{A}$. When ω is a positive normal linear functional on $\mathfrak{A}$, the restriction $\omega \mid \mathscr{P}$ is a measure on $\mathscr{P}$.

1.1.1. *Problem* (Mackey). Does every measure on $\mathscr{P}$ arise in this way?

For abelian algebras, the question is easily answered in the affirmative. From this, and since a completely additive positive linear functional on $\mathfrak{A}$ is normal, the question (for general $\mathfrak{A}$) can without difficulty be rephrased in the following form.

1.1.2. *Problem.* If a mapping $\nu: \mathfrak{A} \to \boldsymbol{C}$ satisfies

(i) $\nu(H + iK) = \nu(H) + i\nu(K)$ for all selfadjoint H and K in $\mathfrak{A}$,

(ii) for each maximal abelian * subalgebra $\mathscr{A}$ of $\mathfrak{A}$, the restriction $\nu \mid \mathscr{A}$ is a positive normal linear functional on $\mathscr{A}$,

does it follow that ν is linear on $\mathfrak{A}$? When μ is the dimension function in a factor $\mathfrak{A}$ of type II_1, the above questions reduce to the problem of "additivity of the trace", solved affirmatively by Murray and von Neumann [**14**, Chapter II]. An ingenious argument due to Gleason [**6**] gives a positive answer when $\mathfrak{A}$ is $\mathscr{B}(\mathscr{H})$, the algebra of all bounded operators acting on a Hilbert space $\mathscr{H}$. From this, and by establishing continuity properties of measures, Gunson [**8**] proved the same result for hyperfinite factors. Analogous questions for certain special C^*-algebras have been considered by Aarnes [**1**], [**2**].

1.2. *Closure problems.* Let $\mathfrak{A}$ be a * algebra of operators acting on a Hilbert space $\mathscr{H}$.

Problem. If each maximal abelian * subalgebra $\mathscr{A}$ of $\mathfrak{A}$ is norm closed in $\mathscr{B}(\mathscr{H})$, does it follow that $\mathfrak{A}$ itself is norm closed?

This question remains open. The analogous problem, with weak-operator closure in place of norm closure, has been solved affirmatively by Pedersen [**15**], after earlier work by Laison [**12**].

1.3. *Continuity of linear mappings.* Suppose that $\mathfrak{A}$ and $\mathscr{B}$ are C^*-algebras, acting on Hilbert spaces $\mathscr{H}$ and $\mathscr{K}$ respectively, and $\xi : \mathfrak{A} \to \mathscr{B}$ is a linear mapping.

1.3.1. *Problem.* If $\xi \mid \mathscr{A}$ is norm continuous, for every maximal abelian * subalgebra $\mathscr{A}$ of $\mathfrak{A}$, does it follow that ξ is norm continuous on $\mathfrak{A}$?

1.3.2. *Problem.* The same question for ultraweak continuity.

1.3.3. *Problem.* The same question for ultrastrong * continuity.

Easy examples, in which ξ is a * isomorphism or a * anti-isomorphism between von Neumann algebras, give negative answers to the corresponding questions about the weak, strong and ultrastrong topologies.

Problem 1.3.1 arose in connexion with a question about continuity of certain derivations (which was later answered by different methods [**9**], [**16**]). By the principle of uniform boundedness, it is equivalent to the corresponding question about linear functionals on $\mathfrak{A}$. It has an affirmative answer when $\mathfrak{A}$ is a von Neumann algebra, a uniformly hyperfinite C^*-algebra [**17**] or a postliminal C^*-algebra [**4**]. For general C^*-algebras, the problem remains open, and is equivalent [**4**] to the corresponding question in which $\mathscr{A}$ runs through all C^*-subalgebras with a single selfadjoint generator. In the negative direction, Barnes and Duncan [**5**, Proposition 5.1] give an example of a Banach algebra with a discontinuous linear functional which is continuous on each abelian subalgebra. The analogous problem, for homomorphisms from a C^*-algebra $\mathfrak{A}$ into a Banach algebra, has recently been solved affirmatively by Sinclair [**18**, Corollary 4.3], following earlier work by Stein [**19**, Corollary II.1] in the von Neumann algebra case. The existence or otherwise of a discontinuous homomorphism, from an abelian C^*-algebra into a Banach algebra, is a well-known unsolved problem; for some recent results, see [**10**].

Problem 1.3.2 reduces at once to the corresponding question about a linear functional ρ on $\mathfrak{A}$. Takesaki [**20**, Corollary 1] gave an affirmative answer for the case in which $\mathfrak{A}$ is a von Neumann algebra and ρ is norm continuous on $\mathfrak{A}$; in view of later work on Problem 1.3.1, the assumption of norm continuity is seen to be redundant. In the context of C^*-algebras, Problem 1.3.2 remains open. By taking $\mathfrak{A}$ in its universal representation, an affirmative answer for Problem 1.3.2 (in full generality) would imply the same for Problem 1.3.1.

For a linear functional on $\mathfrak{A}$, ultraweak continuity is equivalent to ultrastrong * continuity, so Problems 1.3.2 and 1.3.3, for linear functionals, are equivalent.

1.3.4. *Problem.* Does ultraweak continuity of a linear mapping $\xi : \mathfrak{A} \to \mathscr{B}$ imply ultrastrong * continuity?

It is known that ultraweak continuity of ξ on $\mathfrak{A}$ is equivalent to ultrastrong * continuity *on bounded subsets of* $\mathfrak{A}$ ([**3**, Corollary III. 10]; the extension from von

Neumann algebras to represented C^*-algebras is straightforward). If ξ is positive, or if either of the algebras $\mathfrak{A}$, $\mathscr{B}$ is abelian, Problem 1.3.4 has an affirmative solution; in general, it remains open. This problem, for represented C^*-algebras, is equivalent to the same question for von Neumann algebras only; and both are equivalent to Problem 1.3.3 for von Neumann algebras.

2. Norm continuity of linear functionals. Our work on Problem 1.3.1 starts with the following simple result.

2.1. LEMMA. *Suppose that $\mathscr{F}$ is a family of C^*-algebras with the following property:*

($\bigstar$) *If $\mathfrak{A} \in \mathscr{F}$, τ is an unbounded linear functional on $\mathfrak{A}$ and $K > 0$, there is a self-adjoint element A in the unit ball of $\mathfrak{A}$, and a C^*-subalgebra $\mathscr{B}$ of $\mathfrak{A}$, such that $\mathscr{B} \in \mathscr{F}$, $\tau \mid \mathscr{B}$ is unbounded,*

$$|\tau(A)| > K, \qquad AB = BA = 0 \ (B \in \mathscr{B}).$$

*If ρ is a linear functional on a C^*-algebra $\mathfrak{A}$ in $\mathscr{F}$, and $\rho \mid \mathscr{A}$ is bounded for each maximal abelian * subalgebra $\mathscr{A}$ of $\mathfrak{A}$, then ρ is bounded.*

PROOF. If ρ is unbounded, successive applications of ($\bigstar$) yield a sequence $(\mathfrak{A}_j)$ of C^*-subalgebras of $\mathfrak{A}$, and a sequence (A_j) of selfadjoint elements in the unit ball of $\mathfrak{A}$, such that

$$|\rho(A_j)| > j, \ A_jB = BA_j = 0 \ (B \in \mathscr{B}_j), \ A_{j+1} \in \mathscr{B}_j, \mathscr{B}_j \in \mathscr{F},$$
$$\mathscr{B}_1 \in \mathscr{B}_2 \supseteq \mathscr{B}_3 \supseteq \cdots, \ \rho \mid \mathscr{B}_j \text{ is unbounded.}$$

Since $A_jA_k = 0$ when $j \neq k$, there is a maximal abelian * subalgebra $\mathscr{A}$ of $\mathfrak{A}$ whose unit ball contains each A_j; and $\rho \mid \mathscr{A}$ is unbounded, contrary to our assumption. Thus ρ is bounded.

We next sketch a proof that the condition ($\bigstar$) is satisfied when $\mathscr{F}$ consists *either* of all continuous von Neumann algebras *or* of all properly infinite von Neumann algebras. In either case, each $\mathfrak{A}$ in $\mathscr{F}$ has a * subalgebra $\mathscr{M}$, with the same unit as $\mathfrak{A}$, * isomorphic to the algebra of 2×2 complex matrices. With $\{F_{jk}: j, k = 1, 2\}$ a selfadjoint system of matrix units in $\mathscr{M}$, and τ an unbounded linear functional on $\mathfrak{A}$, at least one of the four linear functionals τ_{jk} ($j, k = 1, 2$) on $\mathfrak{A}$ defined by $\tau_{jk}(A) = \tau(F_{j1}AF_{1k})$ is unbounded. If E ($\neq 0, I$) is a projection in $\mathscr{M}$, and

$$\rho(A) = \tau(EF_{11}AF_{11}E), \tag{1}$$

then ρ can be expressed as a linear combination of the τ_{jk}'s. Given four such projections $E_1, \cdots, E_4$, and the corresponding $\rho_1, \cdots, \rho_4$, computation of the appropriate determinant shows that it is, in general, possible to "invert" and express each τ_{jk} as a linear combination of $\rho_1, \cdots, \rho_4$. Hence at least one of $\rho_1, \cdots, \rho_4$ is unbounded. This shows that the linear mapping ρ in (1), and hence also the restriction $\tau \restriction E\mathfrak{A}E$, is unbounded for "most" projections E in $\mathfrak{A}$. It turns out that one can choose E so that both $\tau \mid E\mathfrak{A}E$ and $\tau \mid (I - E)\mathfrak{A}(I - E)$ are unbounded. With K a positive constant, there is a selfadjoint A in the unit ball of $(I - E)\mathfrak{A}(I - E)$ such that $|\tau(A)| > K$; so the conclusions of condition ($\bigstar$) are satisfied, with $\mathscr{B}$ defined as $E\mathfrak{A}E$.

The preceding paragraph, together with Lemma 2.1, gives an affirmative answer

to Problem 1.3.1 for continuous or properly infinite von Neumann algebras. With slight variation, it yields the same result for all von Neumann algebras and for certain other C^*-algebras. For the details, we refer to [17].

3. Ultrastrong * continuity of linear mappings. Suppose that $\mathfrak{A}$ and $\mathscr{B}$ are C^*-algebras acting on Hilbert spaces $\mathscr{H}$ and $\mathscr{K}$ respectively, $\eta : \mathfrak{A} \to \mathscr{B}$ is an ultraweakly continuous linear mapping and $\bar{\eta} : \mathfrak{A}^- \to \mathscr{B}^-$ is its ultraweakly continuous extension to the weak-operator closure $\mathfrak{A}^-$ of $\mathfrak{A}$. It is not difficult to prove

3.1. LEMMA. *η is ultrastrong * continuous if and only if the same is true of $\bar{\eta}$.*

By exploiting the fact that the given representation of $\mathfrak{A}$ on $\mathscr{H}$ is quasi-equivalent to a subrepresentation of the universal representation π of $\mathfrak{A}$, we can deduce from Lemma 3.1 that

3.2. LEMMA. *η is ultrastrong * continuous if and only if the same is true of $\eta \circ \pi^{-1}$: $\pi(\mathfrak{A}) \to \mathscr{B}$.*

Since every positive linear functional on $\pi(\mathfrak{A})$ is ultraweakly continuous, it is easy to check that $\eta \circ \pi^{-1}$ is ultrastrong * continuous if and only if the following condition is satisfied: Given any normal state ω of $\mathscr{B}$, there is a positive linear functional ρ on $\mathfrak{A}$ such that

$$\omega(\eta(A)^*\eta(A) + \eta(A)\eta(A)^*) \leqq \rho(A^*A + AA^*) \qquad (A \in \mathfrak{A}). \tag{2}$$

Let $\mathscr{S}$ ($\subseteq \mathscr{B}$) be the set of operators

$$\sum_1^n [\eta(A_r)^*\eta(A_r) + \eta(A_r)\eta(A_r)^*],$$

where $A_1, \cdots, A_n \in \mathfrak{A}$ and $\left\| \sum_1^n [A_r^*A_r + A_rA_r^*] \right\| \leqq 1$. If $\eta \circ \pi^{-1}$ is ultrastrong * continuous, it follows from (2) that each normal state ω on $\mathscr{B}$ is bounded on $\mathscr{S}$; and from the principle of uniform boundedness, $\mathscr{S}$ is bounded. By homogeneity, if $M = \sup \{ \|S\| : S \in \mathscr{S} \}$,

$$\left\| \sum_{r=1}^n [\eta(A_r)^*\eta(A_r) + \eta(A_r)\eta(A_r)^*] \right\| \leqq M \left\| \sum_{r=1}^n [A_r^*A_r + A_rA_r^*] \right\|, \tag{3}$$

whenever $A_1, \cdots, A_n \in \mathfrak{A}$. Conversely, if (3) is satisfied for some constant M, and ω is a normal state of $\mathscr{B}$, it is easily verified that the convex hull $\mathscr{S}_1$ ($\subseteq \mathfrak{A}^+$) of the set

$$\{A^*A + AA^* : A \in \mathfrak{A},\ \omega(\eta(A)^*\eta(A) + \eta(A)\eta(A)^*) = 1\}$$

does not contain $\{0\}$. Upon separating $\mathscr{S}_1$ and $\{0\}$ by a hyperplane, we find a positive linear functional ρ on $\mathfrak{A}$ which satisfies (2).

From the preceding paragraph, together with Lemma 3.2, we have

3.3. THEOREM. *An ultraweakly continuous linear mapping $\eta : \mathfrak{A} \to \mathscr{B}$ is ultrastrong * continuous if and only if there is a constant K such that*

$$\left\| \sum_{r=1}^n [\eta(A_r)^*\eta(A_r) + \eta(A_r)\eta(A_r)^*] \right\| \leqq K \|\eta\|^2 \left\| \sum_{r=1}^n [A_r^*A_r + A_rA_r^*] \right\| \tag{4}$$

for every finite set $A_1, \cdots, A_n$ of elements of $\mathfrak{A}$.

From Theorem 3.3, and by taking $\mathfrak{A}$ in its universal representation, it follows that Problem 1.3.4 has an affirmative answer if and only if every *bounded* linear mapping η between C^*-algebras satisfies an inequality of the type (4). It suffices to consider only *hermitian* mappings η and, for these, (4) can be replaced by the simpler inequality

$$\left\| \sum_{r=1}^{n} \eta(A_r)^2 \right\| \leqq K \|\eta\|^2 \left\| \sum_{r=1}^{n} A_r^2 \right\|, \tag{5}$$

for every finite set $A_1, \cdots, A_n$ of *selfadjoint* elements of $\mathfrak{A}$.

When $\mathfrak{B}$ is abelian, (5) is automatically satisfied with $K = 1$ and can be proved by straightforward estimation of $\rho(\sum \eta(A_r)^2)$, where ρ is a multiplicative linear functional on $\mathfrak{B}$. It follows at once from the generalized Schwarz inequality [**11**] that (5) is again satisfied with $K = 1$ when η is a positive linear mapping.

Suppose next that $\mathfrak{A}$ is abelian. By taking $\mathfrak{A}$ in its universal representation, extending η to an ultraweakly continuous hermitian linear mapping from $\mathfrak{A}^-$ into $\mathfrak{B}^-$ and approximating each A_r by a finite linear combination of its spectral projections, the proof of (5) is quickly reduced to the case in which each A_r has the form

$$A_r = \sum_{j=1}^{m} \lambda_{jr} E_j,$$

where $\{E_1, \cdots, E_m\}$ is an orthogonal family of projections with sum I. With x a unit vector in $\mathfrak{K}$, and $a_{jk} = \operatorname{Re}\langle \eta(E_j)x, \eta(E_k)x \rangle$,

$$\left| \sum_{j,k=1}^{m} a_{jk} s_j t_k \right| = \left| \operatorname{Re}\langle \eta(\textstyle\sum s_j E_j)x, \eta(\sum t_k E_k)x \rangle \right| \leqq \|\eta\|^2,$$

whenever $s_1, \cdots, s_m, t_1, \cdots, t_m$ are real numbers with modulus not greater than 1. A well-known inequality of Grothendieck ([7], see also [**13**, p. 279]) now implies that

$$\left| \sum_{j,k=1}^{m} a_{jk} \langle u_j, v_k \rangle \right| \leqq K \|\eta\|^2 \sup_j \|u_j\| \sup_k \|v_k\| \tag{6}$$

whenever $u_1, \cdots, u_m, v_1, \cdots, v_m$ are vectors in a real inner product space, K being a constant ($\leqq \sinh \frac{1}{2}\pi$). With $u_j (= v_j)$ the vector $(\lambda_{j1}, \cdots, \lambda_{jn})$ in $\boldsymbol{R}^n$, (6) reduces to

$$\left\langle \left(\sum_r \eta(A_r)^2 \right) x, x \right\rangle \leqq K \|\eta\|^2 \, \|A_1^2 + \cdots + A_n^2\|.$$

Thus (5) is automatically satisfied, with $K = \sinh \frac{1}{2}\pi$, when $\mathfrak{A}$ is abelian.

This concludes my positive information concerning (5). In the negative direction, one can construct von Neumann algebras $\mathfrak{A}$ (abelian) and $\mathfrak{B}$, an ultraweakly closed selfadjoint subspace $\mathfrak{M}$ of $\mathfrak{A}$ which contains the identity of $\mathfrak{A}$, and an ultraweakly continuous hermitian linear mapping $\eta: \mathfrak{M} \to \mathfrak{B}$, such that (5) fails for every $K > 0$.

References

1. J. F. Aarnes, *Physical states on a C*-algebra*, Acta Math. **122** (1969), 161–172. MR **40** #747.
2. ———, *Quasi-states on C*-algebras*, Trans. Amer. Math. Soc. **149** (1970), 601–625. MR **43** #8311.

3. C. A. Akemann, *The dual space of an operator algebra*, Trans. Amer. Math. Soc. **126** (1967), 286–302. MR **34** #6549.

4. B. A. Barnes, *Linear functionals continuous on Abelian * subalgebras of a B*-algebra* (preprint).

5. B. A. Barnes and J. Duncan, *The Banach algebra $l^1(S)$* (preprint).

6. A. M. Gleason, *Measures on the closed subspaces of a Hilbert space*, J. Math. Mech. **6** (1957), 885–894. MR **20** #2609.

7. A. Grothendieck, *Résumé de la théorie métrique des produits tensoriels topologiques*, Bol. Soc. Math. São Paulo **8** (1956), 1–79. MR **20** #1194.

8. J. Gunson, *Physical states on quantum logics*. I, Ann. Inst. Henri Poincaré **17** (1972), 295–311.

9. B. E. Johnson and S. K. Parrott, *Operators commuting with a von Neumann algebra modulo the set of compact operators*, J. Functional Analysis **11** (1972), 39–61.

10. B. E. Johnson, *Norming $C(\Omega)$ and related algebras* (preprint).

11. R. V. Kadison, *A generalized Schwarz inequality and algebraic invariants for operator algebras*, Ann. of Math. (2) **56** (1952), 494–503. MR **14**, 481.

12. D. Laison, *On W*-embeddings of AW*-algebras*, Proc. Amer. Math. Soc. **35** (1972), 499–502. MR **46** #6049.

13. J. Lindenstrauss and A. Pełczyński, *Absolutely summing operators in $\mathscr{L}_p$-spaces and their applications*, Studia Math. **29** (1968), 275–328. MR **37** #6743.

14. F. J. Murray and J. von Neumann, *On rings of operators*. II, Trans. Amer. Math. Soc. **41** (1937), 208–248.

15. G. K. Pedersen, *Operator algebras with weakly closed abelian subalgebras*, Bull. London Math. Soc. **4** (1972), 171–175. MR **47** #9300.

16. J. R. Ringrose, *Automatic continuity of derivations of operator algebras*, J. London Math. Soc. **5** (1972), 432–438.

17. ———, *Linear functionals on operator algebras and their abelian subalgebras*, J. London Math. Soc. **7** (1973), 553–560.

18. A. M. Sinclair, *Homomorphisms from C*-algebras* (preprint).

19. J. D. Stein, *Continuity of homomorphisms of von Neumann algebras*, Amer. J. Math. **91** (1969), 153–159. MR **39** #6095.

20. M. Takesaki, *On the conjugate space of an operator algebra*, Tôhoku Math. J. (2) **10** (1958), 194–203. MR **20** #7227.

THE UNIVERSITY
NEWCASTLE UPON TYNE, ENGLAND NE1 7RU

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Some Aspects of Ergodic Theory in Operator Algebras

Erling Størmer

1. Introduction. There seems to be a strong tendency towards a closer relationship between the theory of von Neumann algebras and ergodic theory. This will be apparent from the work of Connes and Krieger to be discussed at this Congress, and I believe that further study of automorphisms of von Neumann algebras will be most fruitful. In the present report I shall give a survey of what is known in the case when there exist normal states invariant under a group of automorphisms, and show how many results are direct generalizations of theorems on finite invariant measures in ergodic theory.

The main progress on this part of the theory has been made during the last eight years after mathematical physicists became seriously interested in C^*-algebras and mainly in the more special ones called asymptotically abelian C^*-algebras, discussed by Ruelle at the Nice Congress; see, e.g., [6], [15], [10]. One of the more special definitions of this concept states that a C^*-algebra $\mathfrak{A}$ is asymptotically abelian with respect to a group G of *-automorphisms if there is a sequence $\{g_n\}$ in G such that $\lim_n \|[g_n(A), B]\| = 0$ for all $A, B \in \mathfrak{A}$. Then the G-invariant states form a Choquet simplex, so it is natural to study the extreme points, called ergodic states. Using the GNS-representation one quickly reduces the analysis to the case of ergodic groups of automorphisms of von Neumann algebras such that there exists a normal invariant state, where a group G is said to be ergodic if $g(A) = A$ for all $g \in G$ implies A is a scalar operator.

2. Generalizations of abelian theory. I shall first review the part of the theory which is analogous to the abelian case, and then in the next section discuss results which are purely nonabelian in character. Throughout these notes we let $\mathscr{R}$ denote a von Neumann algebra, G a group of *-automorphisms usually considered as a discrete group, and $\mathscr{R}^G$ the fixed point algebra in $\mathscr{R}$ under G. Note that if $(X, \mathscr{B}, \mu)$

is a measure space with μ a probability measure and T a nonsingular measurable transformation of X leaving μ invariant, then T induces a *-automorphism of the abelian von Neumann algebra $L^\infty(X, \mu)$ via $f \to f \circ T^{-1}$, and $f \to \int f d\mu$ is a normal invariant state on $L^\infty(X, \mu)$. Thus the connection with ergodic theory should be clear.

The ergodic theorem known [12] states that $\mathscr{R}$ is G-finite, i.e., the normal G-invariant states separate points of $\mathscr{R}^+$, if and only if there is a faithful normal G-invariant projection map (i.e., conditional expectation) Φ of $\mathscr{R}$ onto $\mathscr{R}^G$. Furthermore, the normal G-invariant states are exactly those of the form $\omega \circ \Phi$ with ω a normal state on $\mathscr{R}^G$. For $A \in \mathscr{R}$, $\Phi(A)$ is the unique operator in $\text{conv}(g(A): g \in G)^- \cap \mathscr{R}^G$, and there exists a fixed net in $\text{conv}(g: g \in G)$ converging pointwise strongly to Φ [7]. In particular, if G is cyclic generated by θ, then

$$\Phi(A) = \text{strong limit}_n \frac{1}{n} \sum_1^n \theta^j(A).$$

As in the abelian case compactness conditions on G are equivalent to G-finiteness. Let ω be a normal state such that its orbit $\{\omega \circ g: g \in G\}$ is weakly relatively compact in the predual $\mathscr{R}_*$. Then its convex hull is weakly relatively compact, so the Ryll-Nardzewski theorem gives rise to a normal invariant state. Using a result of Akemann on weak relative compactness in $\mathscr{R}_*$ we have [17]:

THEOREM 1. *Suppose ω is a faithful normal state on $\mathscr{R}$. Then there exists a faithful normal G-invariant state if and only if given $\varepsilon > 0$ there exists $\delta > 0$ such that if E is a projection in $\mathscr{R}$ with $\omega(E) < \delta$ then $\omega(g(E)) < \varepsilon$ for all $g \in G$.*

It is also possible to characterize G-finiteness in terms of relative compactness of G in the space of linear operators on $\mathscr{R}$ [17].

About 30 years ago Hopf (see [11]) introduced an equivalence relation on the measurable sets. In the language of von Neumann algebras it said that two projections E and F in $\mathscr{R}$ (now abelian) are equivalent if for each $g \in G$ there is a projection $E_g \in \mathscr{R}$ such that $E = \sum E_g$ and $F = \sum g(E_g)$. Hopf showed that finiteness in the partial ordering obtained was equivalent to the existence of a finite invariant measure, and later on Kawada and Halmos (see [11]) did a similar thing in the semifinite case. In the nonabelian case the analogous equivalence relation is the following: If E and F are projections in $\mathscr{R}$ then $E \sim_G F$ if for each $g \in G$ there is $T_g \in \mathscr{R}$ such that $E = \sum T_g^* T_g$ and $F = \sum g(T_g T_g^*)$ [18]. Note that if $T_g = 0$ whenever $g \neq e$ then $E \sim F$ in the usual sense, and if $G = \{e\}$, $E \sim F$ if and only if $E \sim_G F$. Since it can be shown that under an assumption of countable decomposability the relation $\sim_G$ restricts to that of Hopf in the abelian case we have unified the latter and the usual equivalence relation in von Neumann algebras. Let $\mathfrak{M}$ be the crossed product of $\mathscr{R}$ with G, and let Φ be the imbedding of $\mathscr{R}$ into $\mathfrak{M}$. Then it can be shown that $E \sim_G F$ if and only if $\Phi(E) \sim \Phi(F)$, and furthermore that E is finite in the ordering $\prec_G$ defined by $\sim_G$ if and only if $\Phi(E)$ is finite in $\mathfrak{M}$ [14]. Thus it is not difficult to show the following [18]:

THEOREM 2. *$\mathscr{R}$ is semifinite with respect to the ordering $\prec_G$ if and only if there exists a faithful normal semifinite G-invariant trace on $\mathscr{R}$.*

THEOREM 3. *If $\mathscr{R}$ is countably decomposable, $\mathscr{R}$ is finite with respect to $\prec_G$ if and only if there exists a faithful normal finite G-invariant trace on $\mathscr{R}$.*

From the point of view of this report very little has been done on spectral theory for unitary operators implementing automorphisms. A very powerful technique is available in the theory of spectral subspaces [2], [3], [4], which is applicable to the case when G is a locally compact abelian group and the automorphisms are implemented by a strongly continuous unitary representation of G. A result along these lines is the following [20]:

THEOREM 4. *Suppose $\mathscr{R}$ is not both abelian and finite dimensional. Suppose U is a unitary operator implementing an ergodic automorphism of $\mathscr{R}$. Then the spectrum of U is the unit circle.*

When $\mathscr{R}$ is hyperfinite it appears that entropy will be a useful tool in the study of automorphisms. In work being done by Connes and the author at the time of this writing two possible candidates for entropy appear naturally when $\mathscr{R}$ has a faithful normal finite trace τ. The simplest concept of entropy for an automorphism θ is the abelian entropy $H(\theta) = \sup H(\theta \mid \mathscr{A})$, where the sup is taken over all abelian von Neumann subalgebras $\mathscr{A}$ of $\mathscr{R}$ invariant under θ, and $H(\theta \mid \mathscr{A})$ is the entropy of the restriction of θ to $\mathscr{A}$ defined in the abelian theory. For the analogues of Bernoulli shifts it is possible to show that the entropy is as expected from the abelian case; in particular we have the following [5]:

THEOREM 5. *Let $\mathscr{R}$ be the hyperfinite II_1-factor and θ the n-shift, i.e., there exists a I_n-subfactor F of $\mathscr{R}$ such that F and $\theta^j(F)$ commute whenever $j \neq 0$, and $(\bigcup_{j\in Z} \theta^j(F))'' = \mathscr{R}$. Then $H(\theta) = \log n$.*

3. Purely nonabelian theory. This theory is concerned with the algebraic structure of $\mathscr{R}$ and special properties of the invariant states, e.g., how close they are to being traces. The first general result is due to Hugenholtz [10] and was afterwards generalized in several directions. In one direction we have [16]:

THEOREM 6. *Suppose $\mathscr{R}$ is semifinite and G acts ergodically on the center of $\mathscr{R}$. Suppose ω is a faithful normal G-invariant state. Then there exists up to a scalar multiple a unique faithful normal semifinite G-invariant trace τ on $\mathscr{R}$. Furthermore there exists a unique positive selfadjoint operator H affiliated with $\mathscr{R}^G$ such that $\omega(A) = \tau(HA)$ for $A \in \mathscr{R}$.*

In particular if G is ergodic, ω is a trace. If separability conditions are added, it is possible to extend the theorem to the case without ergodic action on the center [8], [13].

In order to study the type III case it seems that some assumptions of asymptotic abelianness are necessary. The most easily formulated result is this [19], [1], [9]:

THEOREM 7. *Let ω be a faithful vector state on $\mathscr{R}$ obtained from the GNS-representation of an ergodic state of an asymptotically abelian C*-algebra. Then the spectrum of the modular operator defined by ω equals the invariant $S(\mathscr{R})$ defined by Connes* [4].

It should be remarked that the assumption that ω be faithful is unnecessary. If E denotes the support projection in $\mathscr{R}$ of ω and $\Delta_\omega = \Delta E + (1 - E)$ with Δ the modular operator of ω restricted to $\mathscr{R}_E$, then the following possibilities occur for ω obtained as in the theorem.

(a) $\mathscr{R}$ is of type I_∞ and $\omega|\mathscr{R}_E$ is a trace.

(b) $\mathscr{R}$ is finite and ω is a trace.

(c) $\mathscr{R}$ is of type II_∞ and $\omega|\mathscr{R}_E$ is a trace.

(d) $\mathscr{R}$ is of type III_λ, $0 < \lambda < 1$, and $\operatorname{Sp} \Delta_\omega = \{\lambda^n : n \in \boldsymbol{Z}\}^-$.

(e) $\mathscr{R}$ is of type III_1.

Note that $\mathscr{R}$ cannot be of type III_0, and that we obtain no new information in the type III_1 case.

References

1. H. Araki, *Remarks on spectra of modular operators of von Neumann algebras*, Comm. Math. Phys. **28** (1972), 267–277.

2. W. Arveson, *On groups of automorphisms of operator algebras*, J. Functional Analysis **15** (1974), 217–243.

3. H. J. Borchers, *Über Ableitungen von C*-algebren*, Nach. der Göttingen Akad. **2** (1973).

4. A. Connes, *Une classification des facteurs de type* III, Ann. Sci. École Norm. Sup. **6** (1973), 133–252.

5. A. Connes and E. Størmer, *Entropy of automorphisms of* II_1 *von Neumann algebras*, Acta Math. (to appear).

6. S. Doplicher, D. Kastler and W. D. Robinson, *Covariance algebras in field theory and statistical mechanics*, Comm. Math. Phys. **3** (1966), 1–28. MR **34** #4930.

7. S. Doplicher, D. Kastler and E. Størmer, *Invariant states and asymptotic abelianness*, J. Functional Analysis **3** (1969), 419–434. MR **39** #4686.

8. A. Guichardet, *Sur les systèmes dynamiques non commutatifs* (to appear).

9. R. Herman, *Spectra of automorphism groups of operator algebras*, Duke Math. J. **41** (1974), 667–673.

10. N. M. Hugenholtz, *On the factor type of equilibrium states in quantum statistical mechanics*, Comm. Math. Phys. **6** (1967), 189–193. MR **37** #772.

11. K. Jacobs, *Neuere Methoden und Ergebnisse der Ergodentheorie*, Ergebnisse der Mathematik und ihrer Grenzegebiete, N. F., Heft 29, Springer-Verlag, Berlin, 1960. MR **23** #A1000.

12. I. Kovács and J. Szücs, *Ergodic type theorems in von Neumann algebras*, Acta Sci. Math. (Szeged) **27** (1966), 233–246. MR **35** #753.

13. R. Nest, *Invariant weights of operator algebras satisfying the KMS condition* (to appear).

14. G. K. Pedersen and E. Størmer, *Automorphisms and equivalence in von Neumann algebras*. II, Indiana Univ. Math. J. **23** (1973/74), 121–129. MR **47** #9301.

15. D. Ruelle, *States of physical systems*, Comm. Math. Phys. **3** (1966), 133–150. MR **34** #3360.

16. E. Størmer, *Automorphisms and invariant states of operator algebras*, Acta Math. **127** (1971), 1–9.

17. ———, *Invariant states of von Neumann algebras*, Math. Scand. **30** (1972), 253–256. MR #7455.

18. ———, *Automorphisms and equivalence in von Neumann algebras*, Pacific J. Math. **44** (1973), 371–383. MR **47** #4756.

19. ———, *Spectra of states and asymptotically abelian C*-algebras*, Comm. Math. Phys. **28** (1972), 279–294.

20. ———, *Spectra of ergodic transformations*, J. Functional Analysis **15** (1974), 202–215.

INSTITUTE OF MATHEMATICS, BLINDERN
OSLO, NORWAY

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Homotopy Invariants for Banach Algebras*

Joseph L. Taylor

Our topic is the relationship between the structure of a commutative Banach algebra A with identity and various homotopy invariants of its maximal ideal space Δ.

The subject began with Shilov's idempotent theorem [**11**] which can be formulated as follows: If $Q(A)$ denotes the additive subgroup of A generated by the idempotents of A, then the Gel'fand transform induces an isomorphism $Q(A) \simeq H^0(\Delta, Z)$ where $H^p(\Delta, Z)$ denotes the pth integral Čech cohomology group of Δ.

Arens [**1**] and Royden [**10**] later proved that the Gel'fand transform induces an isomorphism $A^{-1}/\exp(A) \simeq H^1(\Delta, Z)$, where A^{-1} is the invertible group of A and $\exp(A)$ is the subgroup consisting of elements with logarithms in A.

At this point it seemed natural to try to characterize all the Čech groups of Δ in terms of the structure of A. However, the next result led in a new direction. This was the Arens theorem [**3**] which asserts that the Gel'fand transform yields an isomorphism

$$[\mathrm{GL}_n(A)] \to [\Delta, \mathrm{GL}_n(\boldsymbol{C})].$$

Here, $\mathrm{GL}_n(A)$ is the group of invertible $n \times n$ matrices over A, $[\mathrm{GL}_n(A)]$ is the factor group modulo the identity component, and $[\Delta, \mathrm{GL}_n(\boldsymbol{C})]$ is the set of homotopy classes of maps from Δ to $\mathrm{GL}_n(\boldsymbol{C})$.

Ĕidlin [**4**] and Novodvorskiĭ [**8**] pointed out that the Arens result could be used to characterize the group $K^{-1}(\Delta)$ of Atiyah-Hirzebruch K-theory in terms of the structure of A. Novodvorskiĭ [**8**] also established an isomorphism $\mathrm{Proj}(A) \simeq \mathrm{Vect}(\Delta)$ where $\mathrm{Proj}(A)$ is the semigroup (under $\oplus$) of isomorphism classes of finitely generated projective A-modules and $\mathrm{Vect}(\Delta)$ is the semigroup of isomorphism

* Research supported by the National Science Foundaion under NSF grant GP-43114X.

classes of complex vector bundles on Δ. This leads immediately to an isomorphism $K_0(A) \simeq K^0(\Delta)$ between the Grothendieck group $K_0(A)$ for A and the group $K^0(\Delta)$ of Atiyah-Hirzebruch K-theory for Δ.

Module tensor product induces a multiplication on Proj(A) which makes it into a semiring with identity. The invertible group in this semiring is the Picard group Pic(A). Forster [5] pointed out that the Gel'fand transform induces an isomorphism $\mathrm{Pic}(A) \simeq H^2(\Delta, Z)$, thus carrying the program of characterizing the Čech groups of Δ one step further.

The second generation results along these lines (those of Arens, Novodvorskiĭ, Forster) all derive from fundamental work of Grauert [6] on holomorphic fiber bundles. Grauert's work makes possible a very general theorem (Theorem 1 of this paper) which can be used to generate the preceding results and many more. A weak version of this general theorem appears in Novodvorskiĭ [8]. The general result is more or less implicit in a paper of Lin [7]. We shall give a detailed proof and a wide variety of applications in [14]. Here we shall simply state the result and list some unsolved problems.

Since our interest in this subject stems from our study of measure algebras on locally compact groups, we include a brief discussion of the connection between the two subjects.

1. A general theorem. In what follows let F be a closed complex submanifold of a domain $U \subset \boldsymbol{C}^n$, and let A be a commutative Banach algebra with identity and with maximal ideal space Δ.

We define a subset A_F of A^n—the set of n-tuples of elements of A. If A is semisimple, it suffices to define A_F to be the set of all $\alpha \in A^n$ such that

(a) $$\sigma(\alpha) \subset F,$$

where $\sigma(\alpha)$ is the joint spectrum of the tuple α. If A is not semisimple we must be more subtle. If $\alpha = (a_1, \cdots, a_n)$ satisfies (a) and $a_{n+1}, \cdots, a_m$ are arbitrary elements of A, then the tuple $\alpha' = (a_1, \cdots, a_m)$ has its spectrum in the set $F \times \boldsymbol{C}^{m-n} \subset \boldsymbol{C}^m$. Given a domain $V \subset \boldsymbol{C}^m$ with $\sigma(\alpha') \subset V$ and a function $f \in \mathcal{O}(V)$, the holomophic functional calculus for Banach algebras yields an element $f(\alpha') \in A$. We define A_F to the set of tuples $\alpha \in A^n$ such that α satisfies (a) and

(b) For arbitrary $m \geqq n$, $a_{n+1}, \cdots, a_m$, V, and f, as above, if $f = 0$ on $V \cap (F \times \boldsymbol{C}^{m-n})$ then $f(\alpha') = 0$.

This is complicated looking condition. However, in most applications it turns out to be equivalent to a very simple condition. Suppose there exist functions $\varphi_1, \cdots, \varphi_k \in \mathcal{O}(U)$ which vanish on F and are such that the complex Jacobian matrix $(\partial\varphi_i/\partial z_j)$ at each $z \in F$ has kernel equal to the complex tangent space to F at z. Then (b) is equivalent to

(b′) $$\varphi_i(\alpha) = 0 \quad \text{for } i = 1, \cdots, k$$

(cf. [14, §2.8]). This is a condition that is easily checked in practice.

We can now state the main theorem:

THEOREM 1. *If F is a discrete union of complex homogeneous spaces, then A_F is locally path connected and the Gel'fand transform induces a bijection $[A_F] \simeq [\Delta, F]$, where $[A_F]$ is the set of components of A_F and $[\Delta, F]$ is the set of homotopy classes of maps from Δ to F.*

In the case where F is open in $\boldsymbol{C}^n$ and A is semisimple, this result is due to Novodvorskiĭ [8]. The result is implicit in work of Lin [7] and is developed in detail in [14].

Each of the theorems mentioned in the introduction can be easily deduced from this result. For example, the Arens theorem follows directly by choosing $F = \mathrm{GL}_n(\boldsymbol{C})$. The results of Novodvorskiĭ [8] and Forster [5] on $\mathrm{Proj}(A)$, $K_0(A)$, and $\mathrm{Pic}(A)$ follow by choosing F to be appropriate spaces of idempotent complex matrices. Other choices for F lead to results relating topological invariants from real and symplectic K-theory to the structure of Banach algebras. For details see [14].

The point is that Theorem 1 can be used to relate homotopy invariants of the form $[\Delta, F]$ to the structure of A whenever F is a space which has the homotopy type of a complex homogeneous space or a direct limit of complex homogeneous spaces. The invariants that arise in K-theory (the study of real, complex, or symplectic vector bundles) are all of this form. On the other hand, while the functors $H^p(\ , Z)$ of Čech cohomology have the form $\Delta \to [\Delta, K(Z, p)]$ for spaces $K(Z, p)$ (the Eilenberg-MacLane spaces), only for $p = 0, 1, 2$ can these spaces be approximated by complex homogeneous spaces. Hence we are still faced with the following problem:

Problem 1. Find a direct characterization of the Čech groups $H^p(\Delta, Z)$ $(p > 2)$ in terms of the structure of A.

Of course the cohomology and K-theory of a space determine each other up to torsion *via* the Chern character. Since the analogous problem for K-theory has been solved, Problem 1 has been solved up to torsion (i.e., for rational rather than integral coefficients). However, a solution to the full problem would be very interesting.

Complex homogeneous spaces play the role they do because Theorem 1 is based on a result of Grauert [6] and Ramspott [9]: If F is a complex homogeneous space and X is a Stein space, then the inclusion yields a bijection $[X, F]_a \to [X, F]$, where $[X, F]_a$ is the set of homotopy classes of holomorphic maps from X to F. This raises the following question:

Problem 2. For what class of complex manifolds F does the conclusion of the Grauert-Ramspott theorem hold for all Stein spaces X? Is it nececssary that F be complex homogeneous?

We close with a discussion of a particular class of Banach algebras where further progress in this area might lead to very interesting results.

2. Measure algebras on a locally compact group. Let G be a locally compact abelian group and let $M(G)$ be the Banach algebra of all finite regular Borel measures on G under convolution multiplication. We denote the maximal ideal space of $M(G)$ by Δ.

It is well known that $M(G)$ is a very complicated algebra and that Δ is a large and complicated space which defies explicit description. Hence, it is rather surprising that certain kinds of problems in $M(G)$ have rather simple and elegant answers. For example, Cohen's computation of the idempotents of $M(G)$ is such a result (cf. [3]).

In view of Shilov's idempotent theorem, Cohen's result can be thought of as a computation of $H^0(\Delta, Z)$.

In fact, all of the Čech groups of Δ can be computed. Let $M_1(G) = \bigoplus_\tau L(G_\tau)$, where τ ranges over topologies on G which dominate the original topology and for which G_τ is still a locally compact topological group. Here $L(G_\tau)$ is the algebra of measures absolutely continuous with respect to Haar measure on G_τ. Now $M_1(G)$ is a relatively simple subalgebra of $M(G)$. If Δ_1 is its maximal ideal space, then the inclusion $M_1(G) \to M(G)$ induces a continuous map $\Delta \to \Delta_1$. The basic theorem is the following (cf. [12]):

THEOREM 2. *The map $\Delta \to \Delta_1$ induces an isomorphism $H^p(\Delta_1, Z) \to H^p(\Delta, Z)$ of Čech cohomology for each $p \geqq 0$.*

For $p = 0$ this leads directly to Cohen's idempotent theorem. For $p = 1$ it leads, together with the Arens-Royden theorem, to a characterization of the invertible elements of $M(G)$ modulo exponentials (cf. [12]). For $p = 2$ it leads, together with Forster's theorem, to a characterization of $\mathrm{Pic}(M(G))$ (cf. [13]). In each of these cases it leads to an explicit and useful result concerning measures on G. What significance does it have for $p > 2$? This is not yet clear in view of the lack of a solution to Problem 1.

If we could prove that the map $\Delta \to \Delta_1$ induces a bijection $[\Delta, F] \to [\Delta_1, F]$ whenever F is a complex homogeneous space, then we could use Theorem 1 to generate a wide variety of results on $M(G)$. A map $X \to Y$ of compact Hausdorff spaces is called a shape equivalence if the induced map $[Y, F] \to [X, F]$ is bijective whenever F is an ANR. Thus, what we really need is a positive solution to the following problem:

Problem 3. Is the map $\Delta \to \Delta_1$ a shape equivalence?

In conjunction with Theorem 1, a positive solution to Problem 3 would yield solutions to a variety of problems in harmonic analysis. Among them are

Problem 4. Characterize (up to exponential factors) the invertible $n \times n$ matrices over $M(G)$.

Problem 5. Characterize (up to similarity) the $n \times n$ idempotent matrices over $M(G)$.

Problem 6. Characterize the finitely generated projective $M(G)$ modules.

References

1. R. F. Arens, *The group of invertible elements of a commutative Banach algebra*, Studia Math. (Ser. Specjalna) Zeszyt **1** (1963), 21–23. MR **26** #4198.

2. ———, *To what extent does the space of maximal ideals determine the algebra*?, Function

Algebras (Proc. Internat. Sympos. on Function Algebras, Tulane Univ., 1965), Scott-Foresman, Chicago, Ill., 1966, pp. 164–168. MR **33** #3126.

3. P. J. Cohen, *On a conjecture of Littlewood and idempotent measures*, Amer. J. Math. **82** (1960), 191–212. MR **24** #A3231.

4. V. L. Ĕĭdlin, *The topological characteristics of the space of maximal ideals of a Banach algebra*, Vestnik Leningrad. Univ. **22** (1967), no. 13, 173–174. MR **36** #1982.

5. O. Forster, *Functiontheoretische Hilfsmittel im der Theorie der kommutativen Banach-Algebren* (in manuscript).

6. H. Grauert, *Analytische Faserungen über holomorph-vollstandigen Räumen*, Math. Ann. **135** (1958), 263–273. MR **20** #4661.

7. V. Ja. Lin, *Holomorphic fiberings and multivalued functions of elements of a Banach algebra*, Funkcional. Anal. i Priložen. **7** (1973), no. 2, 43–51 = Functional Anal. Appl. **7** (1973), 122–128. MR **47** #7444.

8. M. E. Novodvorskiĭ, *Certain homotopic invariants of the space of maximal ideals*, Mat. Zametki **1** (1967), 487–494. (Russian) MR **35** #742.

9. K. J. Ramspott, *Stetige und holomorphe Schnitte in Bündeln mit homogener Faser*, Math. Z. **89** (1965), 234–246. MR **31** #4932.

10. H. L. Royden, *Function algebras*, Bull. Amer. Math. Soc. **69** (1963), 281–298. MR **26** #6817.

11. G. E. Šilov, *On decomposition of a commutative normed ring in a direct sum of ideals*, Mat. Sb. **32** (**74**) (1953), 353–364; English transl., Amer. Math. Soc. Transl. (2) **1** (1955), 37–48. MR **14**, 884; **17**, 512.

12. J. L. Taylor, *Measure algebras*, CBMS Regional Conference Ser. in Math., no. 16, Amer. Math. Soc., Providence, R. I., 1972.

13. ———, *Banach algebras and topology*, Proc. of "Algebras in Analysis" Conference, Birmingham, England, 1973 (to appear).

14. ———, *Topological invariants of the maximal ideal space of a Banach algebra*, Advances in Math. (to appear).

University of Utah
Salt Lake City, Utah 84112, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Harmonic Analysis on Real Semisimple Lie Groups

V. S. Varadarajan*

This report is a survey of some of the main results concerning harmonic analysis on real semisimple Lie groups. Due to limitations of space and time only the bare outlines of the theory are sketched; the reader interested in a more detailed treatment should consult the articles cited at the end. The bibliography is not intended to be exhaustive and serves only to indicate some of the main sources of the subject.

Let G be a connected semisimple Lie group having finite center and θ a Cartan involution of G. We write K for the maximal compact subgroup of G corresponding to θ; $\mathfrak{g}$ (resp. $\mathfrak{k}$) for the Lie algebra of G (resp. K); and $\mathfrak{G}$ for the universal enveloping algebra of the complexification $\mathfrak{g}_c$ of $\mathfrak{g}$. Most of the progress in the subject has centered around two broad themes:

(I) Explicit construction of the irreducible representations of G.

(II) Development of a Fourier transform theory for functions and distributions on G, in particular, explicit determination of the Plancherel formula for G.

The most fundamental representations of G are those that occur as irreducible direct summands of the regular representation of G in $L^2(G)$. These are precisely the irreducible unitary representations of G having square-integrable matrix coefficients. The set of equivalence classes of these is called the *discrete series* of G and is denoted by $\mathscr{E}_2(G)$. Of course G need not always have a discrete series. It was proved by Harish-Chandra [3] that $\mathscr{E}_2(G) \neq \emptyset$ if and only if $\mathrm{rk}(G) = \mathrm{rk}(K)$, rk being the rank, or equivalently if and only if G has a compact Cartan subgroup (CSG). We shall now describe briefly Harish-Chandra's construction of the discrete series of G [3].

Let $\mathscr{D}_0$ be the algebra of all differential operators on G and $\mathscr{D}_l$ (resp. $\mathscr{D}_r$) the subalgebra of all elements of $\mathscr{D}_0$ that commute with the left (resp. right) transla-

*Supported in part by NSF grant GP 33696X.

tions of G; let $\mathscr{D}$ be the subalgebra of $\mathscr{D}_0$ generated by $\mathscr{D}_l \cup \mathscr{D}_r$. Let $G = KAN$ be an Iwasawa decomposition of G. If $\|\cdot\|$ is any K-invariant norm on $\mathfrak{g}$, there is a unique function σ on G which is invariant under right and left translations by elements of K and is given on A by $\sigma(\exp H) = \|H\|$, $H \in \mathfrak{a}$, $\mathfrak{a}$ being the Lie algebra of A. By the *Schwartz space* $\mathscr{C}(G)$ of G we then understand the space of all $f \in C^\infty(G)$ for which $(1 + \sigma)^r(Ef) \in L^2(G)$ $\forall\, r \geqq 0$, $E \in \mathscr{D}$. $\mathscr{C}(G)$ is a Fréchet space under the collection of seminorms $f \mapsto \|(1 + \sigma)^r(Ef)\|_2$ ($\|\cdot\|_2$ is the norm in $L^2(G)$). It is closed under convolution ($*$), and $*$ converts it into a topological algebra. $C_c^\infty(G) \subset \mathscr{C}(G)$, the natural inclusion being continuous; moreover, $C_c^\infty(G)$ is dense in $\mathscr{C}(G)$. Distributions on G that extend (necessarily uniquely) to continuous linear functionals on $\mathscr{C}(G)$ are called *tempered.*

Suppose π is an irreducible unitary representation of G. Then its character Θ_π is a distribution on G that is invariant under all inner automorphisms of G and is an eigendistribution for every $z \in \mathfrak{Z}$ where $\mathfrak{Z}$ is the center of $\mathfrak{G}$ and is identified as usual with $\mathscr{D}_l \cap \mathscr{D}_r$. Distributions on G with these two properties are called *invariant eigendistributions* and it was proved by Harish-Chandra [2] that any such distribution is a locally integrable function on G which coincides with an invariant analytic function on the dense open subset G' of regular points of G. Given an invariant eigendistribution Θ on G we write Θ also for the corresponding analytic function on G'. We now have the following theorem, where, for simplicity of formulation, we assume that G is contained in the simply connected complex analytic group G_c with Lie algebra $\mathfrak{g}_c$.

THEOREM. *$\mathscr{E}_2(G) \neq \emptyset$ if and only if* $\mathrm{rk}(G) = \mathrm{rk}(K)$. *Let* $\mathrm{rk}(G) = \mathrm{rk}(K)$ *and $B \subset K$ a compact CSG. Let $\mathfrak{b} \subset \mathfrak{k}$ be the Lie algebra of B and λ an integral linear form on $\mathfrak{b}_c$ which is regular in the sense that $\langle \lambda, \alpha \rangle \neq 0$ for all roots α of $(\mathfrak{g}_c, \mathfrak{b}_c)$. Let Δ as usual be the function on B such that*

$$\Delta(\exp H) = \prod_{\alpha \in P} (e^{\alpha(H)/2} - e^{-\alpha(H)/2}) \qquad (H \in \mathfrak{b})$$

where P is a positive system of roots. Then there exists exactly one invariant eigendistribution Θ_λ on G such that

(i) *Θ_λ is tempered,*

(ii) $\Theta_\lambda(\exp H)\Delta(\exp H) = \sum_{s \in W(G/B)} \varepsilon(s) e^{s\lambda(H)}$ *($H \in \mathfrak{b}$, $\exp H \in B \cap G'$),*

where $W(G/B)$ is the subgroup of the Weyl group of $(\mathfrak{g}_c, \mathfrak{b}_c)$ coming from G, and $\varepsilon(s)$ has the usual meaning. Further, let $\hat{b} = \hat{b}_\lambda$ be the character of B such that $\hat{b}(\exp H) = e^{\lambda(H)}$ ($H \in \mathfrak{b}$) and write $\Theta_{\hat{b}} = \Theta_\lambda$, $\varepsilon(\hat{b}) = \mathrm{sgn} \prod_{\alpha \in P} \langle \lambda, \alpha \rangle$ and [1] $q = \frac{1}{2}\dim(G/K)$. *Then $(-1)^q \varepsilon(\hat{b})\Theta_{\hat{b}}$ is the character of a class $\omega(\hat{b}) \in \mathscr{E}_2(G)$; and the map $\hat{b} \mapsto \omega(\hat{b})$ induces a bijection of $\hat{B}'/W(G/B)$ onto $\mathscr{E}_2(G)$, $\hat{B}'$ being the set of all $\hat{b}_\lambda \in \hat{B}$ with λ regular and integral.*

With appropriate modifications this theorem remains valid even for a possibly disconnected reductive group G provided it satisfies the following conditions: (i) G and center(G) have finitely many connected components, (ii) Ad(G) is con-

[1] q is an integer $\geqq 0$.

tained in the connected complex adjoint group of $\mathfrak{g}_c$, (iii) the analytic subgroup of G defined by $[\mathfrak{g}, \mathfrak{g}]$ is closed in G. Further, one can associate an invariant eigendistribution $\Theta_{\hat{b}}$ which is tempered and is given on the compact CSG B by a formula similar to (ii) of the theorem, not only for $\hat{b} \in \hat{B}'$ but also for all $\hat{b} \in \hat{B}$. It is no longer true that such an eigendistribution is determined by its restriction to $B \cap G'$, but one can make canonical choices and arrange matters so that the $\Theta_{\hat{b}}$ satisfy the following symmetry conditions: If σ is an automorphism of G with $\sigma B = B$ and if $\varepsilon(\sigma) = \pm 1$ is such that $\sigma \cdot \Delta = \varepsilon(\sigma)\Delta$, then $\sigma \cdot \Theta_{\hat{b}} = \varepsilon(\sigma)\Theta_{\hat{b}}$ $(\hat{b} \in \hat{B})$ (cf. Harish-Chandra [3]). It is however not known whether $\Theta_{\hat{b}}$ is, for $\hat{b} \in \hat{B} \backslash \hat{B}'$, a scalar multiple of the character of an irreducible unitary representation of G.

In view of the great importance of the discrete series it is natural to seek to obtain the corresponding representations of G more directly. It turns out that one can realize "almost all" of the discrete series in certain "L^2 cohomology" spaces associated with G. For details the reader is referred to the papers of Narasimhan and Okamoto [1], Parthasarathy [1], and Schmid [1].

Following Harish-Chandra we now associate with each conjugacy class Γ of CSG's of G a set $\mathscr{E}_\Gamma(G)$ of equivalence classes of irreducible unitary representations of G. If the CSG's in B are compact, $\mathscr{E}_\Gamma(G) = \mathscr{E}_2(G)$; otherwise, we proceed as follows. Let $A \in \Gamma$ be θ-stable and write $A = A_I A_R$ where $A_I = A \cap K$, $A_R = \exp \mathfrak{a}_R$, $\mathfrak{a}_R$ being the subspace of the Lie algebra of A on which $\theta = -\mathrm{id}$. We choose a parabolic subgroup (psgrp) P of G whose Langlands decomposition is of the form $P = MA_RN$ (cf. Harish-Chandra [4] for details); M is reductive, $K_M = K \cap M$ is a maximal compact subgroup of M, A_I is a CSG of M (so that $\mathrm{rk}(M) = \mathrm{rk}(K_M)$) and A_R lies in the center of MA_R. Then $\mathscr{E}_2(M) \neq \emptyset$. For any $\sigma \in \mathscr{E}_2(M)$, $\xi \in \hat{A}_R$, we can now introduce the unitary representation $\pi_{\sigma,\xi}$ of G induced by the representation $man \mapsto \xi(a)\eta(m)$ of P where η is any member of σ. The equivalence class of $\pi_{\sigma,\xi}$ depends only on σ and ξ and not on P, while, for fixed σ, $\pi_{\sigma,\xi}$ is irreducible for almost all ξ and has a finite composition series for all ξ. The set of equivalence classes of the irreducible constituents of the $\pi_{\sigma,\xi}$ is $\mathscr{E}_\Gamma(G)$; it depends only on Γ and not on the choice of $A \in \Gamma$. If $\omega \in \mathscr{E}_\Gamma(G)$, ω is tempered in the sense that the corresponding character and matrix coefficients are tempered distributions on G; conversely, if π is an irreducible unitary representation of G whose matrix coefficients are tempered, then the equivalence class of π belongs to $\mathscr{E}_\Gamma(G)$ for some Γ (cf. Langlands [1], Trombi [1]).

If the CSG's of Γ are of the Iwasawa type, i.e., if $\dim(A_R)$ has its maximum possible value, then the associated psgrp's are the minimal ones. Then the representations $\pi_{\sigma,\xi}$ are the so-called principal series representations of G. Since M is compact in this case, one does not need the theory of the discrete series to construct them. So it is not surprising that historically these were studied first (cf. Gel'fand and Naĭmark [1], Harish-Chandra [1], Bruhat [1]).

For fixed A, Γ, it is natural to ask for the structure of the ring of bounded operators in the Hilbert space of $\pi_{\sigma,\xi}$ commuting with $\pi_{\sigma,\xi}$; in particular, one would like to determine the set $\Omega(A)$ of all $(\sigma, \xi) \in \mathscr{E}_2(M) \times \hat{A}_R$ for which $\pi_{\sigma,\xi}$ is irreducible. Although these questions have not yet been fully answered, substantial results have

been obtained; for A not of Iwasawa type, see Harish-Chandra [5]. Suppose A is of the Iwasawa type. Then Bruhat [1] has shown that $\pi_{\sigma,\xi}$ is irreducible whenever (σ, ξ) is in "general position" in the following sense: If $x \in G$ normalizes MA_R but does not centralize it, then the representations $ma \mapsto \xi(a)\eta(m)$ and $ma \mapsto \xi(a^x)\eta(m^x)$ of MA_R are inequivalent; in the case when $\mathrm{rk}(G/K) = 1$, Knapp and Stein [1], [2] have determined $\Omega(A)$ explicitly. If we suppose further that σ is the trivial class of M, then the representations $\pi_{\sigma,\xi}$ are spherical, i.e., they contain the trivial class of K exactly once. One can analyse them in great detail using infinitesimal methods. This was done by K. R. Parthasarathy and others [1] when G was complex and by Kostant [1] for arbitrary G, and led to the result that the representations $\pi_{\sigma,\xi}$ (σ trivial, $\xi \in \hat{A}_R$) are all irreducible. In the case of a complex G, there is only one conjugacy class of CSG's; the CSG's in this class are of the Iwasawa type and the corresponding representations $\pi_{\sigma,\xi}$ are all irreducible (cf. Wallach [1], Želebenko [1], Knapp and Stein [1], [2]).

The infinitesimal method was pioneered by Harish-Chandra [1]. His work suggests that it might be worthwhile studying the class of $\mathfrak{G}$-modules V having the following properties: (i) V is finitely generated, (ii) each element of V is contained in a $\mathfrak{k}$-stable finite-dimensional subspace of V and the isotypical subspaces of V considered as a $\mathfrak{k}$-module are all finite dimensional. The structure of such $\mathfrak{G}$-modules is not completely determined except in some special cases (cf. Gel'fand and Ponomarev [1]).

Fix a θ-stable CSG $A = A_I A_R$. Then $\hat{A} \simeq \hat{A}_I \times \hat{A}_R$. For $\xi \in \hat{A}$ we write ξ_I, ξ_R for the projections of ξ in $\hat{A}_I$, $\hat{A}_R$. If $\xi_I \in \hat{A}_I'$, then we can associate, with ξ_I, $\sigma(\xi_I) \in \mathscr{E}_2(M)$ and construct the representation $\pi_{\sigma(\xi_I),\xi_R}$. Let Θ^ξ be the character of $\pi_{\sigma(\xi_I),\xi_R}$. We then define, for any $f \in \mathscr{C}(G)$,

$$\hat{f}_A(\xi) = (-1)^{q_M}\varepsilon(\xi_I)\Theta^\xi(f)$$

where q_M, $\varepsilon(\xi_I)$ have the same relation to M as q, $\varepsilon(\hat{b})$ have to G in our statement of the theorem on the discrete series. It can then be proved that $\hat{f}_A$ extends from $\hat{A}_I' \times \hat{A}_R$ to $\hat{A}$ as an element of the Schwartz space $\mathscr{C}(\hat{A})$ of $\hat{A}$, that the extension is unique, and that the extension (denoted by $\hat{f}_A$ again) has the following symmetry property for a suitable $\pm$ 1-valued character ε_A of $W(G/A)$:

$$\hat{f}_A(s\cdot\xi) = \varepsilon_A(s)\hat{f}_A(\xi) \qquad (\xi \in \hat{A}, s \in W(G/A)).$$

Let now A_i ($1 \leqq i \leqq r$) be a complete system of θ-stable CSG's of G no two of which are conjugate in G. For $f \in \mathscr{C}(G)$, write $\hat{f}_i = \hat{f}_{A_i}$, $\varepsilon_i = \varepsilon_{A_i}$, and put, following Harish-Chandra [4], $\hat{f} = (\hat{f}_1, \cdots, \hat{f}_r)$. We call $\hat{f}$ the *invariant Fourier transform* of f. The map $f \mapsto \hat{f}$ is continuous from $\mathscr{C}(G)$ into $\mathscr{C}(A_1) \times \cdots \times \mathscr{C}(A_r)$ and gives rise, via duality, to a Fourier transform theory of tempered invariant distributions on G. The computation of the Fourier transform of the Dirac measure at the identity element 1 of G then leads to the Plancherel formula (Harish-Chandra [4]):

THEOREM. *There exist unique continuous functions C_i on $\hat{A}_i$ ($1 \leqq i \leqq r$) with the following properties:*

(i) $C_i(\xi) = 0$ *if* $\xi_I \in \hat{A}_I \setminus \hat{A}_I'$.

(ii) $C_i(s\cdot\xi) = \varepsilon_i(s)C_i(\xi)$ $(s \in W(G/A_i),\ \xi \in \hat{A}_i)$.

(iii) *C_i has at most polynomial growth on $\hat{A}_i$.*

(iv) *There is a constant $c \neq 0$ such that, for all $f \in \mathscr{C}(G)$,*

$$cf(1) = \sum_{1\leq i\leq r} \int_{\hat{A}_i} C_i(\xi)\hat{f}_i(\xi)\, d_i\xi,$$

where $d_i\xi$ are fixed Haar measures on $\hat{A}_i$ for $1 \leqq i \leqq r$.

One would like to determine the Fourier transform of important tempered invariant distributions associated with G. When $\mathrm{rk}(G/K) = 1$, Sally and Warner [1] have done this for the invariant measures on the conjugacy classes of G.

It is natural to attempt to construct a more general transform for functions in $\mathscr{C}(G)$ that uses the matrix coefficients of $\pi_{\sigma(\xi),\xi_R}$ rather than its character. The resulting transform theory would be capable of dealing with arbitrary (not necessarily invariant) tempered distributions on G. For $G = \mathrm{SL}(2, \boldsymbol{R})$ this was done by Ehrenpreis and Mautner [1] and by Ranga Rao and Varadarajan (unpublished); for the general case see Arthur [1].

It is of interest to examine the Fourier transform on the subspace $\mathscr{C}^p(G)$ of all $f \in \mathscr{C}(G)$ with $(1 + \sigma)^m(Ef) \in L^p(G)$ for all $m \geqq 0$, $E \in \mathscr{D}$ $(1 \leqq p < 2)$. Ehrenpreis and Mautner [1] studied the case $G = \mathrm{SL}(2, \boldsymbol{R})$; the general case has not been completely settled. At the level of the discrete spectrum this is substantially the problem of determining the subset $\mathscr{E}_p(G) \subset \mathscr{E}_2(G)$ of classes that have matrix coefficients in $L^p(G)$; although this has not yet been answered, Trombi and Varadarajan [2] have obtained many partial results, and in particular have proved that if $\omega \in \mathscr{E}_2(G)$ corresponds to $\lambda \in \mathfrak{h}_c^*$, $\omega \in \mathscr{E}_p(G)$ provided $\inf_\alpha |\langle\lambda, \alpha\rangle|$ is sufficiently large. Further, for the case of the continuous spectrum, Trombi and Varadarajan [1] have completely determined the spherical transforms of the convolution algebra $\mathscr{I}^p(G)$ of spherical functions in $\mathscr{C}^p(G)$; among other things they have shown that the spherical transform is a topological algebra isomorphism of $\mathscr{I}^p(G)$ with the multiplication algebra of functions defined and holomorphic on a certain tube domain in $\mathfrak{a}_c^*$, that are in addition invariant under the Weyl group of the symmetric space G/K and rapidly decreasing in a suitable sense (here $\mathfrak{a}$ is the Lie algebra of A and $G = KAN$ is an Iwasawa decomposition).

It should also be mentioned that analogous transform theories have been developed for functions and distributions on homogeneous spaces associated with G, especially G/K. For an account of some of these see Helgason [1], Gangolli [1].

Finally it must be noted that the semisimple groups differ from the compact and abelian groups in their having irreducible unitary representations that do not enter the Plancherel formula. From our point of view these are the irreducible unitary representations whose characters are nontempered. We call the corresponding set of equivalence classes the *exceptional series* of G.

The main technique of studying the exceptional series seems to be that of analytic continuation. More precisely, if A is a θ-stable CSG, we consider the representations $\pi_{\sigma,\xi}$ of G induced by the representation $man \mapsto \mu(m)\xi(a)$ of MA_RN, where ξ is now a quasi-character of A_R, i.e., a member of the group $\tilde{A}_R$ of continuous homo-

morphisms of A_R into $C^\times$, the nonzero complexes. According to a theorem of Harish-Chandra [1], all irreducible representations of G (unitary or not) occur as subquotients of the $\pi_{\sigma,\xi}$ ($\sigma \in \mathcal{E}_2(M)$, $\xi \in \tilde{A}_R$) associated with an Iwasawa type CSG. Consequently it is natural to look for the exceptional series from this large series of representations. Although the theory is incomplete, substantial results have already been obtained; see Kostant [1], Harish-Chandra [5] and Knapp and Stein [1], [2].

References

J. Arthur

[1] *Harmonic analysis on the Schwartz space of a reductive Lie group.* I, II (preprint).

F. Bruhat

[1] *Sur les représentations induites des groupes de Lie*, Bull. Soc. Math. France **84** (1956), 97–205. MR **18**, 907.

L. Ehrenpreis and F. I. Mautner

[1] *Some properties of the Fourier transform on semi-simple Lie groups.* I, II, III, Ann. of Math. (2) **61** (1955), 406–439; Trans. Amer. Math. Soc. **84** (1957), 1–55; ibid. **90** (1959), 431–484. MR **16**, 1017; **18**, 745; **21** #1541.

R. A. Gangolli

[1] *Spherical functions on semisimple Lie groups*, Short courses presented at Washington University, Marcel Dekker, New York, 1972.

I. M. Gel'fand and M. A. Naĭmark

[1] *Unitary representations of the classical groups*, Trudy Mat. Inst. Steklov. **36** (1950); German transl., *Unitäre Darstellungen der klassischen Gruppen*, Akademie-Verlag, Berlin, 1957. MR **13**, 722; **19**, 13.

I. M. Gel'fand and V. A. Ponomarev

[1] *The category of Harish-Chandra modules over the Lie algebra of the Lorentz group*, Dokl. Akad. Nauk SSSR **176** (1967), 243–246 = Soviet Math. Dokl. **8** (1967), 1065–1068. MR **36** #6552.

Harish-Chandra

[1] *Representations of semisimple Lie groups.* II, Trans. Amer. Math. Soc. **76** (1954), 26–65. MR **15**, 398.

[2] *Invariant eigendistributions on a semisimple Lie group*, Trans. Amer. Math. Soc. **119** (1965), 457–508. MR **31** #4862d.

[3] *Discrete series for semisimple Lie groups.* I. *Construction of invariant eigendistributions*; II. *Explicit determination of the characters*, Acta Math. **113** (1965), 241–318; ibid. **116** (1966), 1–111. MR **36** #2744; #2745.

[4] *Harmonic analysis on semisimple Lie groups*, Bull. Amer. Math. Soc. **76** (1970), 529–551. MR **41** #1933.

[5] *On the theory of the Eisenstein integral*, Conference on Harmonic Analysis (College Park, Maryland, 1971), Lecture Notes in Math., vol. 266, Springer-Verlag, New York, 1972.

S. Helgason

[1] *A duality for symmetric spaces with applications to group representations*, Advances in Math. **5** (1970), 1–154. MR **41** #8587.

A. W. Knapp and E. M. Stein

[1] *Intertwining operators for semisimple Lie groups*, Ann. of Math. (2) **93** (1971), 489–578.

[2] *Irreducibility theorems for the principal series*, Conference on Harmonic Analysis (College Park, Maryland, 1971), Lecture Notes in Math., vol. 266, Springer-Verlag, New York, 1972.

B. Kostant
[1] *On the existence and irreducibility of certain series of representations*, Bull. Amer. Math. Soc. **75** (1969), 627–642. MR **39** #7031.
R. P. Langlands
[1] *On the classification of irreducible representations of real algebraic groups* (preprint).
M. S. Narasimhan and K. Okamoto
[1] *An analogue of the Borel-Weil-Bott theorem for hermitian symmetric pairs of non-compact type*, Ann. of Math. (2) **91** (1970), 486–511. MR **43** #419.
K. R. Parthasarathy, R. Ranga Rao and V. S. Varadarajan
[1] *Representations of complex semi-simple Lie groups and Lie algebras*, Ann. of Math. (2) **85** (1967), 383–429. MR **37** #1526.
R. Parthasarathy
[1] *Dirac operator and the discrete series*, Ann. of Math. (2) **96** (1972), 1–30.
P. J. Sally, Jr. and G. Warner
[1] *The Fourier transform on semisimple Lie groups of real rank one*, Acta Math. **131** (1973), 1–26.
W. Schmid
[1] *On a conjecture of Langlands*, Ann. of Math. (2) **93** (1971), 1–42. MR **44** #4149.
P. C. Trombi
[1] *The tempered spectrum of a real semisimple Lie group* (preprint).
P. C. Trombi and V. S. Varadarajan
[1] *Spherical transforms on semisimple Lie groups*, Ann. of Math. (2) **94** (1971), 246–303. MR **44** #6913.
[2] *Asymptotic behaviour of eigenfunctions on a semisimple Lie group: The discrete spectrum*, Acta Math. **129** (1972), 237–280.
N. R. Wallach
[1] *Cyclic vectors and irreducibility for principal series representations*, Trans. Amer. Math. Soc. **158** (1971), 107–113. MR **43** #7558.
D. P. Želobenko
[1] *The analysis of irreducibility in the class of elementary representations of a complex semisimple Lie group*, Izv. Akad. Nauk SSSR Ser. Mat. **32** (1968), 108–133 = Math. USSR Izv. **2** (1968), 105–128. MR **37** #2906.

UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024, U.S.A.

Труды Международного Конгресса Математиков
Ванкувер, 1974

Комплексный Гармонический Анализ на Полупростых Группах Ли*

Д. П. Желобенко

1. **Введение.** Под комплексным гармоническим анализом на группе Ли G мы понимаем анализ (в духе Пэли-Винера) функций на G, операторные образы Фурье которых естественно продолжаются в комплексную область. К этому классу относится групповая алгебра $\mathscr{D}(G)$ всех финитных обобщенных функций на G, а также быстро убывающие (быстрее любой экспоненты) функции на G. Среди подалгебр $\mathscr{D}(G)$ особую роль в теории представлений играют $C_0^\infty(G)$, $\mathscr{D}_e(G)$, где $\mathscr{D}_e(G)$—подалгебра всех функций с носителем в единичной точке $e \in G$.

Метод, развитый в работах автора [2]—[5] для полупростых групп Ли, позволяет свести исследование операторного интеграла Фурье на группе G к исследованию его специального случая—сферического преобразования Фурье, которое индуцируется в классе двусторонне K-инвариантных функций, где K—максимальная компактная подгруппа в G. (Если G—комплексная группа ранга l, то сферическое преобразование Фурье сводится к обычному интегралу Фурье в l-мерном евклидовом пространстве.) Указанный метод является чисто алгебраическим и основан на рассмотрении специальной категории в классе двусторонних $U-K$-модулей, где положено $U = U(\mathfrak{g}^c)$, $\mathfrak{g}^c$—комплексификация алгебры Ли $\mathfrak{g}$ группы G. Посредством этого подхода удается установить аналоги классических теорем Пэли-Винера для функций класса $C_0^\infty(G)$, $\mathscr{D}_e(G)$, а также для многих других классов функций.

Замечательно, что при доказательстве этих теорем используется лишь незначительная информация о мере Планшереля (характер роста и явное

* Not presented in person.

значение в классе двусторонне K-инвариантных функций). Тем не менее полные результаты получены в настоящее время лишь для случая, когда G комплексна (другие случаи упоминаются ниже). Эти результаты имеют существенное применение к теории представлений (классификация неприводимых G-модулей и неприводимых $\mathfrak{g}$-модулей Хариш-Чандры).

2. Преобразование Фурье. В дальнейшем для простоты изложения мы считаем, что G—полупростая комплексная связная группа Ли, $G = KAN$—её разложение Ивасавы. *Преобразованием Фурье* числовой функции $x(g)$ на G называется операторная функция

$$x(\nu, \sigma) = \int x(g) e_{\nu\sigma}(g)\, dg,$$

где dg—мера Хаара на G, $e_{\nu\sigma}$—элементарное представление группы G, индуцированное в $C^\infty(K)$ характером $\theta_{\nu\sigma}(man) = m^\nu a^{\sigma+\rho}$ борелевской подгруппы $B = MAN$. Более детально, пусть $D_{\nu\sigma}$—подмодуль левого регулярного G-модуля в классе $C^\infty(G)$, выделяемый уравнением

$$f(xy) = f(x)\theta_{\nu\sigma}(y)^{-1}, \qquad x \in G,\ y \in B.$$

Операция сужения на подгруппу K изоморфно отображает $D_{\nu\sigma}$ на подпространство $\mathfrak{D}_\nu$ всех функций класса $C^\infty(K)$, удовлетворяющих уравнению $\varphi(km) = \varphi(k)m^{-\nu}$, $k \in K$, $m \in M$. Действие группы G в $\mathfrak{D}_\nu$ обозначается $e_{\nu\sigma}$. При этом мы полагаем

$$m^\nu = \exp(\nu, \ln m), \qquad a^\lambda = \exp(\lambda, \ln a), \qquad \nu \in \Gamma,\ \lambda \in \mathfrak{h},$$

где $(\cdot, \cdot)$—киллингова форма в комплексной алгебре $\mathfrak{g}$, $\mathfrak{h}$—картановская подалгебра в $\mathfrak{g}$, Γ—решетка весов тора M, вложенная в $\mathfrak{h}$ относительно киллинговой формы, $\lambda = \sigma + \rho$, где ρ—полусумма положительных корней в $\mathfrak{h}^c$. (Представления $e_{\nu\sigma}$ при чисто мнимых σ образуют основную серию Гельфанда-Наймарка.)

Преобразование Фурье естественно продолжается на обобщенные функции. Если $x \in \mathscr{D}(G)$, то $x(\nu, \sigma)$ является целой операторной функцией от σ. Если $x \in \mathscr{D}_e(G)$, то $x(\nu, \sigma)$—операторный полином от σ.

3. Пространство X и категория $X^{\lambda\mu}$. Опишем класс пространств, для которых применимы предлагаемые методы комплексного анализа. Пусть X—линейное пространство обобщенных функций на G, удовлетворяющее условиям:

(1) X является двусторонним K-модулем относительно левых и правых сдвигов на G,

(2) X является U-модулем относительно дифференцирования обобщенных функций,

(3) образы Фурье элементов $x \in X$ являются целыми (операторными) функциями от σ.

Пусть Λ—множество всех старших весов группы K. Для каждого $\lambda \in \Lambda$ пусть e^λ—центральный проектор в $C(K)$ со старшим весом λ: $e^\lambda = n_\lambda \bar{\chi}_\lambda$, где χ_λ—характер, n_λ—размерность неприводимого представления π^λ со старшим весом λ. Положим

$$X^{\lambda\mu} = e^\lambda X e^\mu, \qquad X^\# = \sum_{\lambda,\mu} X^{\lambda\mu}.$$

Элементы $x(\nu, \sigma)$ при $x \in X^{\lambda\mu}$ естественно отождествляются с матричными функциями от σ со значениями в $\mathrm{Hom}(L^{\mu}_{\nu}, L^{\lambda}_{\nu})$, где положено $L^{\lambda}_{\nu} = e^{\lambda}\mathfrak{D}_{\nu}$. Заметим, что, по правилу двойственности Фробениуса, $L^{\lambda}_{\nu} \cong E^{\lambda} \otimes E^{\lambda}_{\nu}$ (антилинейно по второму сомножителю), где E^{λ}—пространство π^{λ}, E^{λ}_{ν}—весовое подпространство веса ν. Соответственно, если $\tilde{X}^{\lambda\mu}$—образ Фурье $X^{\lambda\mu}$, то нетрудно показать, что

$$\tilde{X}^{\lambda\mu} \cong \mathscr{M}_{\lambda\mu} \otimes \bar{X}^{\lambda\mu},$$

где $\mathscr{M}_{\lambda\mu} = \mathrm{Hom}(E^{\mu}, E^{\lambda})$ и элементы $\bar{X}^{\lambda\mu}$ являются матричными функциями от σ со значениями в $\mathrm{Hom}(E^{\mu}_{\nu}, E^{\lambda}_{\nu})$.

Если X—алгебра относительно свертки, то $X^{\lambda\mu}X^{\mu\nu} \subset X^{\lambda\nu}$, т.е. семейство $X^{\lambda\mu}$ образует категорию (с единицами e^{λ}).

Заметим, что алгебра $U = U(\mathfrak{g}^c) \cong \mathscr{D}_e(G)$ не удовлетворяет условию (1). В этом случае положим

$$U^{\lambda\mu} = \{u \in U : \mathscr{I}^{\lambda}u \subset U\mathscr{I}^{\mu}\}, \qquad \lambda, \mu \in \Lambda,$$

где $\mathscr{I}^{\lambda} = \ker \pi^{\lambda}$—двусторонний идеал в $U(\mathfrak{k}^c)$, $\mathfrak{k}$—алгебра Ли группы K. Заметим, что преобразование Фурье $u(\nu, \sigma)$ элемента $u \in U$ определяется чисто алгебраически (дифференциал $e_{\nu\sigma}$).

В дальнейшем (§5) условия (1)—(3) будут дополнены.

4. Соотношения симметрии. Между функциями $x(\nu, \sigma)$ существуют алгебраические соотношения, порожденные сплетающими операторами для представлений $e_{\nu\sigma}$:

$$Cx(\nu, \sigma) = x(\nu', \sigma')C,$$

где C—сплетающий оператор для пары $e_{\nu\sigma}$, $e_{\nu'\sigma'}$. Среди таких соотношений особую роль играют условия симметрии относительно группы Вейля:

I. *условия симметрии* I *рода* индуцируются группой Вейля $W = M'/M$ и определяются интегралами Шифмана $C = A(m, \nu, \sigma)$, $m \in M'$, причем $\nu' = w\nu$, $\sigma' = w\sigma$, $w \in W$, где w—класс, содержащий m. В [2] показано, что $A(m, \nu, \sigma)$—рациональная матричная функция, порожденная (мультипликативно) l образующими $A_i(\nu, \sigma)$ (простые рефлексии в W).

II. *Условия симметрии* II *рода* индуцируются группой $W \times W$ (группа Вейля алгебры $\mathfrak{g}^c$) и определяются инфинитезимальными (правыми) сдвигами в особых точках функций $A_i(\nu, \sigma)$, причем ν', σ' выражаются через ν, σ преобразованиями $W \times W$ по параметрам $\nu \pm \sigma$ [2], [5].

Операторы симметрии сохраняют $X^{\lambda\mu}$, $\lambda, \mu \in \Lambda$, и индуцируют алгебраические (матричные) соотношения для элементов $\tilde{X}^{\lambda\mu}$, $\bar{X}^{\lambda\mu}$.

5. Основные результаты. Напомним, что сферическое преобразование Фурье действует в классе X^{00} двусторонне K-инвариантных функций. При этом, согласно §4, элементы $\tilde{X}^{00} = \bar{X}^{00}$ являются W-симметричными.

Предположим, наряду с (1)—(3), выполнение следующих условий.

(4) $X^{00}\mathscr{H}(G) \subset X$, относительно обычного умножения функций на G, где

$\mathscr{H}(G)$—множество всех матричных элементов голоморфных неприводимых представлений группы G.

(5) Все матричные элементы $x(\nu, \sigma)$, $x \in X^{\#}$, принадлежат линейному пространству $\mathscr{Z}$, замкнутому относительно умножения на рациональные функции (т.е. если $z \in \mathscr{Z}$, r—рациональная функция, rz—целая функция, то $rz \in \mathscr{Z}$).

(6) $\bar{X}^{00} = \mathscr{Z}^W$.

Положим $H_\nu^{\lambda\mu} = \mathrm{Hom}(E_\nu^\mu, E_\nu^\lambda)$. Для каждого $\nu \in \Gamma$ пусть ν^+—доминантный образ вектора ν относительно W, W_ν—стационарная подгруппа точки ν.

Теорема 1. *Пусть $Z^{\lambda\mu}$—множество всех функций класса $\mathscr{Z}$ со значениями в $H_\nu^{\lambda\mu}$, удовлетворяющих условиям симметрии* I *и* II *рода. Тогда*

$$\bar{X}^{\lambda\mu} = Z^{\lambda\mu} \quad \textit{для всех } \lambda, \mu \in \Lambda.$$

Теорема 2. *Пусть $Z^{\lambda\mu}(\varepsilon)$, $\varepsilon \in \Lambda$—множество всех элементов $Z^{\lambda\mu}$, равных нулю при $\nu^+ \geqq \varepsilon$. Тогда*

$$\sum_{\delta \leqslant \varepsilon} \bar{U}^{\lambda\delta} \bar{X}^{\delta\mu} = Z^{\lambda\mu}(\varepsilon), \qquad \varepsilon \in \Lambda.$$

Следствие 3. *Пусть $N^\lambda = \sum_{\delta<\lambda} U^{\lambda\delta} X^{\lambda\delta}$ в пространстве $X^\lambda = X^{\lambda\lambda}$. Тогда*

$$X^\lambda / N^\lambda \simeq \mathscr{M}_{\lambda\lambda} \otimes \mathscr{Z}_{W_\lambda}, \qquad \lambda \in \Lambda.$$

Теорема 1 дает полное описание пространства, двойственного к $X^{\#}$ относительно преобразования Фурье (аналог теоремы Пэли-Винера). Заметим, что Теорема 1—частный случай Теоремы 2. Следствие 3 имеет приложения в теории представлений (см. ниже).

Если X—непрерывный K-модуль, то $X^{\#}$ всюду плотно в X, что позволяет продолжить полученные результаты на все пространство X.

Пример. Пространство, двойственное к $X = C_0^\infty(G)$, состоит из всех целых функций $x(\sigma) = \bigoplus_\nu x(\nu, \sigma)$, где $x(\nu, \sigma)$; $\mathfrak{D}_\nu \to \mathfrak{D}_\nu$, удовлетворяющих условиям симметрии I и II рода и оценкам вида

$$(*) \qquad |\sigma|^n \left\| \beta^{n_1} x(\sigma) \beta^{n_2} \right\| \leqq C_{nn_1n_2} e^{\mathrm{Re}\langle r, \sigma\rangle}, \qquad r \in \mathfrak{h}_+, n, n_1, n_2 = 0, 1, 2, \cdots,$$

где положено $\|x\|$—норма оператора x в $L^2(K)$, β—оператор Лапласа-Бельтрами на K, $\mathfrak{h}_+$—замыкание доминантной камеры Вейля в $\mathfrak{h}$. При этом (*) равносильно

$$\mathrm{supp}\, x \subset \{g \in G: \ln \|g\|_i \leqq r_i,\ i = 1, 2, \cdots, l\},$$

где $r_i = (r, \lambda_i)$, λ_i—фундаментальные веса алгебры $\mathfrak{g}$, $\|g\|_i$—норма $\pi^{\lambda_i}(g)$.

6. Схема доказательства. Принципиальным моментом теории является доказательство Теоремы 1 при $X = U = U(\mathfrak{g}^c)$ (операционное исчисление [5]). В общем случае Теорема 2 равносильна системе тождеств

$$\bar{U}_\nu^{\lambda\delta} \bar{X}_\nu^{\delta\delta} \bar{U}_\nu^{\delta\mu} = Z_\nu^{\lambda\mu}(\delta), \qquad \delta = \nu^+,$$

где нижний индекс означает сужение функций $x(\nu, \sigma)$ на фиксированное значение ν, ν^+—доминантный образ вектора ν относительно группы Вейля.

Описание множеств $\bar{U}_\nu^{\lambda\delta}$, $\bar{U}_\nu^{\delta\mu}$ требует глубокого алгебраического анализа [1], [4], [5]. Описание множеств $\bar{X}_\nu^{\delta\delta}$, $\delta = \nu^+$, сводится к сферическому преобразованию Фурье [5]. (Замечательно, что только последняя задача зависит от выбора пространства X.)

Замечание. Первоначальная форма операционного исчисления [4] была основана на расширении алгебры U с помощью проекторов e^λ. Дальнейшие упрощения получены в [1], [5]. Некоторые технические фрагменты операционного исчисления (сферический случай) содержатся в [7].

7. Неприводимые модули Хариш-Чандры. Всякий $\mathfrak{g}$-модуль, градуированный конечномерными $\mathfrak{k}$-подмодулями, называется *модулем Хариш-Чандры*. Примером является модуль $L_{\nu\sigma}$ всех K-финитных векторов $e_{\nu\sigma}$, неприводимый при почти всех σ [3]. Положим

$$V_{\nu\sigma} = L^0_{\nu\sigma}/L^+_{\nu\sigma},$$

где $L^0_{\nu\sigma} = Ue_0$, e_0—ненулевой вектор, принадлежащий K-подмодулю π^{λ_0}, $\lambda_0 = \nu^+$—минимальный из старших весов группы K в $e_{\nu\sigma}$, $L^+_{\nu\sigma}$—максимальный подмодуль в $L^0_{\nu\sigma}$, не содержащий π^{λ_0}.

Теорема 4. *Всякий неприводимый модуль Хариш-Чандры эквивалентен одному из тодулей* $V_{\nu\sigma}$.

При этом $V_{\nu\sigma} \simeq V_{\nu'\sigma'}$ только при $\nu' = w\nu$, $\sigma' = w\sigma$. На каждой W-орбите существует точка ν, σ, для которой $V_{\nu\sigma} = L^0_{\nu\sigma}$. Теорема 4 вытекает из Следствия 3 при $X = U$ ([1] и диссертация автора 1971 г.).

8. Неприводимые G-модули. Условимся рассматривать неприводимые G-модули в секвенциально полных локально выпуклых пространствах. Пусть $\mathfrak{M}_{\nu\sigma}$ —множество всех таких модулей, содержащих $V_{\nu\sigma}$ как всюду плотную часть своего дифференциала.

Теорема 5. *Всякий вполне неприводимый G-модуль эквивалентен одному из модулей класса* $\mathfrak{M}_{\nu\sigma}$.

Теорема 6 вытекает из Следствия 3 при $X = C_0^\infty(G)$ [6]. Условие полной неприводимости можно заменить аналогом леммы Шура в пространстве Гординга топологически неприводимого представления [5].

Следствие 6. *Всякий вполне неприводимый модуль E содержит уплотнение E_0, дуальный модуль к которому содержит старшие векторы (собственные векторы борелевской подгруппы $B = MAN$).*

В частности, для унитарных представлений в качестве E_0 можно выбрать подпространство Гординга в E. Сопряженные старшие векторы можно использовать [5] для классификации вполне неприводимых G-модулей.

9. Вещественные группы. Из вещественных групп наиболее подробно исследованы группы ранга 1 ($SO(n, 1)$, $SU(n, 1)$) и группы с единственным классом сопряженности картановских подгрупп (кватернионные группы

$SL(n, \boldsymbol{H})$). Для общей полупростой группы Ли известны операторы симметрии I рода (интегралы Шифмана), частично исследована их регуляризация (Кнапп и Штейн). Сравнительно детально исследовано сферическое преобразование Фурье (Хариш-Чандра, Хелгасон, Тромби и Варадараджан), с аналогами теорем Пэли-Винера для двусторонне K-инвариантных функций.

Некоторые фрагменты операционного исчисления содержатся в известной работе Б. Костанта (по представлениям класса 0). Аналоги теорем Пэли-Винера получены также для простейших примеров нильпотентных и разрешимых групп Ли (Кумахара и Окамото, Андо), а также для группы движений вещественного евклидова пространства.

Литература

1. M. Duflo, *Représentations irréductibles des groupes semi-simple complexes*, Séminaire J. Dixmier, Paris, 1973.

2. Д.П. Желобенко, *Симметрия в классе элементарных представлений полупростой комплексной группы Ли*, Функциональный анализ и его приложения, **1**, вып. 2 (1967), 15–387.

3. _____, *Анализ неприводимости в классе элементарных представлений полупростой комплексной группы Ли*; Известия АН СССР, сер. матем. **32** (1968), 108–133.

4. _____, *Операционное исчисление на полупростой комплексной группе Ли*, Известия АН СССР, сер. матем. **33** (1969), 931–973.

5. _____, *Гармонический анализ на полупростых комплексных группах Ли*, "Наука", Москва, 1974.

6. Д.П. Желобенко и М.А. Наймарк, *Описание вполне неприводимых представлений полупростой комплексной группы Ли*, Известия АН СССР, сер. матем. **34** (1970), 57–82.

7. R.R. Parthasarathy, R. Ranga Rao and V.S. Varadarajan, *Representations of complex semi-simple Lie groups and Lie algebras*, Ann. of Math. (2) **85** (1967), 383–729.

Университет Дружбы Народов им. Патриса Лумумбы
Москва, СССР

Section 11

Probability and Mathematical Statistics, Potential, Measure and Integration

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Solution to the Buffon-Sylvester Problem and Stereology

R. V. Ambartzumian

The last decade has witnessed a rapid increase of research activity in the subject of geometrical probability. Among other things this has resulted in the rise of two completely novel branches, *stochastic geometry* and *stereology*.

Briefly, stereology is interested in the possibility of making inferences about size and shapes of bodies (say in R^3) based on the investigation of random patterns on lines or planes which intersect the body at random.

In the randomised version of the same problem the bodies may compose a statistical ensemble. In this case the problem is to infer the properties of the ensemble, making use of an appropriate mixture of the samples of intersections, drawn from the individual bodies. Such a situation arises for example when we have non-overlapping random bodies placed in some random way in R^3 (that is, we have a spatial stochastic process of bodies) and the problem is to infer about the ensemble they compose, based on statistical properties of disjoint patterns, appearing on planar or linear sections through the whole space.

If the mutual position of the bodies is also of interest, then it is convenient to think about the union of the bodies as a single random set.

In such cases the assumption of invariance of the spatial process with respect to the group of all Euclidean motions of R^3 (the E.m.-invariance) is usually imposed.

At this step we come quite close to the topic of the stochastic geometry, which devotes itself, as we see it, to investigation of the properties of random sets of different kinds, which satisfy the general condition of invariance of their distribution with respect to groups of transformations of the basic space.

In this connection I would like to mention the recent book *Stochastic geometry*, edited by D.G. Kendall and E.F. Harding.

A glance through the field shows clearly enough that substantial progress here

takes place mostly in those cases in which rather straightforward application of concepts and methods of conventional stochastic process theory is preceded by some specific work of pure geometrical or combinatorial nature. The work of K. Krickeberg and of the late Rollo Davidson in [1] is a bright illustration for this.

In the same manner a number of new tools, now in the process of development and presumably useful in both stereology and stochastic geometry, have as their starting point the combinatorial solution of the Buffon-Sylvester problem in integral geometry. It is the aim of the present report to give a brief account of this development, as well as to describe the B.-S. problem and its solution in more detail.

The classical Buffon's problem of the needle (1776) may be formulated as follows.

Let $\mathfrak{m}$ be the set of all sensed (directed) lines on R^2, $m(\ )$ be the E.m.-invariant measure on $\mathfrak{m}$. Given a needle $\Delta \subset R^2$ (needle = a finite line segment), put $B = \{g \in \mathfrak{m}; g \cap \Delta \neq \emptyset\}$. Find $m(B)$. The answer is of course $m(B) = 4\rho$, ρ is the length of Δ. In the year 1890 J. J. Sylvester considered the following modification of Buffon's problem.

Let $\Delta_1, \cdots, \Delta_N$ be N needles fixed on the plane in general position. Find $m(\bigcap_1^N B_i)$ and $m(\bigcup_1^N B_i)$, $B_i = \{g \in \mathfrak{m}; g \cap \Delta_i \neq \emptyset\}$.

According to J. J. Sylvester [2], "the measures in question become linear Diophantine functions of the sides of the complete $2N$-gonal figure, of which the n pairs of extremities of the needles are the angles".

The problem of finding some general algorithm for calculation of those Diophantine coefficients has been posed explicitly by J. J. Sylvester in [2], and this is what we call the Buffon-Sylvester problem.

The surprisingly simple solution, which has been obtained in [3], [4], is conveniently written when the B.-S. problem is generalised in a way in which all pairs of points appear symmetrically.

Suppose a finite set of points $\{P_i\}_1^n$ is fixed on the plane. Denote by $g_{ij} \in \mathfrak{m}$ a line through P_iP_j (g_{ij} and g_{ji} have opposite directions), by ρ_{ij} the distance between P_i and P_j, by g^+ (g^-) the right (left) half-plane, bounded by g.

Consider the minimal algebra $\mathfrak{a}$ of the subsets of $\mathfrak{m}$, which contains all A_i, $i = 1, \cdots, n$, $A_i = \{g \in \mathfrak{m}; P_i \in g^+\}$.

THEOREM 1. *If in the set $\{P_i\}_1^n$ no three points lie on a line, then for every $A \in \mathfrak{a}$, $m(A) < \infty$,*

$$
\begin{aligned}
m(A) &= \sum_{(i,j)} \rho_{ij} c_{ij}(A), \\
c_{ij}(A) &= I_A(i^+, j^-) + I_A(i^-, j^+) - I_A(i^+, j^+) - I_A(i^-, j^-),
\end{aligned}
\tag{1}
$$

where $I_A(i^x, j^y)$ are defined as conditional limits of the indicator function $I_A(g)$ of the set A, that is, $I_A(i^x, j^y) = \lim I_A(g)$, $g \to g_{ij}$, $P_i \in g^x$, $P_j \in g^y$.

Let us see how (1) is generalised to the case of planes in R^3.

Let $\mathfrak{M}$ be the set of oriented planes in R^3, M be the E.m.-invariant measure on $\mathfrak{M}$, $\{P_i\}_1^n \subset R^3$.

The position of $G \in \mathfrak{M}$, satisfying the condition $P_i \in G$, $P_j \in G$, is specified by

$\phi \in (0, 2\pi)$, the angle of rotation around the P_iP_j axis. Denote such a plane by $G_{i,j,\phi}$. Let G^+ and G^- be the two half-spaces bounded by $G \in \mathfrak{M}$. Again $\mathfrak{A}$ is the minimal algebra, this time of subsets of $\mathfrak{M}$, which contains the sets $A_i = \{G \in \mathfrak{M}: P_i \in G^+\}$.

THEOREM 2. *If in the set* $\{P_i\}_1^n$ *no three points lie on a line, then for every* $A \in \mathfrak{A}$, $M(A) < \infty$,

$$
\begin{aligned}
M(A) &= \frac{1}{2}\sum_{i<j} \rho_{ij}\, c_{ij}(A), \\
c_{ij}(A) &= \int_0^{2\pi} [I_A(i^+, j^-, \phi) + I_A(i^-, j^+, \phi) - I_A(i^+, j^+, \phi) - I_A(i^-, j^-, \phi)]\, d\phi, \\
I_A(i^x, j^y, \phi) &= \lim I_A(G), \qquad G \to G_{i,j,\phi},\ P_i \in G^x,\ P_j \in G^y.
\end{aligned} \tag{2}
$$

The above limits are well defined for almost all $\phi \in (0, 2\pi)$.

Of direct stereological interest is the following corollary of Theorem 1, which is obtained by a rather long process of integration of (1), written for

$$A = \bigcap_1^n B_i, \qquad B_i = \{g \in \mathfrak{m}; g \cap \Delta_i \neq \emptyset\},$$

$\{\Delta_i\}_1^n$ being independent, uniformly distributed random chords of a planar convex domain.

The result refers to convex polyhedrons in R^3 and happens to be

$$\frac{1}{2(n+1)} \sum_k Bl_{n+1}(f_k) = \int x^n\, dL - \frac{n}{2} \int x^n tg\varepsilon_1 tg\varepsilon_2 \cos\delta\, dL. \tag{3}$$

Here f_k is the set of faces of the polyhedron X, $Bl_n(\ )$ are functionals, defined for convex planar domains, which have been introduced by W. Blaschke under the name of "Integrale der Sehnenpotenzen". On the right side of (3), x is the length of $L \cap X$, L is a line in R^3, dL is the element of invariant measure of lines in R^3, ε_1 and ε_2 are the angles between L and the normals $\bar{n}_1$ and $\bar{n}_2$ to ∂X at the points of intersection of L with X, δ is the angle between the planes $L\bar{n}_1$ and $L\bar{n}_2$.

From the point of view of applications, the most desirable would have been the result of the form

$$\sum_k Bl_n(f_k) = Ez(x) \tag{4}$$

where E is the expectation with respect to random L with density proportional to dL, with some function z. Meanwhile (3) is tantamount to $\sum_k Bl_n(f_k) = Ez(x, \Omega_1, \Omega_2)$, Ω_1 and Ω_2 are the directions of $\bar{n}_1$ and $\bar{n}_2$.

We encounter an analogous situation when obtaining expressions for the moments

$$m_n = \sum_k (\pi - \alpha_k) \left|e_k\right|^{n+1} \left(\sum(\pi - \alpha_k) \left|e_k\right|\right)^{-1}$$

($\{e_k\}$ is the set of edges of X, α_k is the angle at e_k, $|\cdot|$ stands for length) by proper integration of (2).

Here an equation of the form

$$m_n = EZ(p; \phi_1, \cdots, \phi_k) \tag{5}$$

is obtained, Z standing for a functional, depending not solely on the polygon $p = X \cap G$, but also on the angles $\phi_1, \cdots, \phi_k$ of intersection of G with the faces of the polyhedron X. In (5), E stands for expectation with respect to random G (with density proportional to dM). Of course, (5) is a weaker result than the equation

$$m_n = EZ(p) \tag{6}$$

about which the applied mathematician would dream. Unfortunately, neither (4) nor (6) is at present known to exist.

Let us look now how (3) applies in the stereology of E.m.-invariant random (loose) packings of nonoverlapping convex polyhedrons in R^3. Assume that the polyhedrons are coloured white, and the complement—black!

First, fix an individual realisation of the packing, and write down (3) for each X_i in the realisation, supplying the corresponding quantities by index i. Sum up over the set $\{i;\ X_i \subset S(R, 0)\}$, $S(R, 0)$ is the sphere of radius R centered at 0. On the right side we obtain

$$\int_{L \cap S(R) \neq \emptyset} \left(\sum_{i \in J(L,R)} x_i^n \left[1 - \frac{n}{2} tg\varepsilon_1^{(i)} tg\varepsilon_2^{(i)} \cos \delta_i \right] \right) dL$$

$$J(L, R) = \{i;\ L \cap X_i \neq \emptyset,\ X_i \subset S(R, 0)\}.$$

Formal application of the law of large numbers yields

$$\sum_{i \in J} x_1^n tg\varepsilon_1^{(i)} tg\varepsilon_2^{(i)} \cos \delta_i = \mathscr{E} x^n \left[1 - \frac{n}{2} tg\varepsilon_1 tg\varepsilon^2 \cos \delta \right] \cdot \operatorname{card} J.$$

On the line L we have a stochastic process $\{I_i, \Omega_1^{(i)}, \Omega_2^{(i)}\}$ (that is the process of white intervals "marked" by orientations at their ends) with distribution not dependent on L. $\mathscr{E}$ stands for expectation with respect to an arbitrary white interval in this process.

It is well known that $\int \operatorname{card} J(L, R)\ dL = \pi \cdot \lambda(R)$, $\lambda(R)$ is the total area of the faces inside $S(R, 0)$. Therefore

$$\mathscr{E} x^n \left[1 - \frac{n}{2} tg\varepsilon_1 tg\varepsilon_2 \cos \sigma \right] = \lim [\pi \lambda(R)]^{-1} \sum Bl_n(f_k), \qquad f_k \in S(R, 0).$$

The limit on the right-hand side is assumed to exist and not to depend on the realisation with prob. 1. It clearly has the meaning $\mathscr{A}^{-1} \mathscr{E} Bl_n(f)$, with f for "arbitrary face" in the spatial process, $\mathscr{A}$ for mean value of the area of such a face. Finally

$$\mathscr{E} Bl_n(f) = \mathscr{A} \mathscr{E} x^n\ [1 - (n/2) tg\varepsilon_1 tg\varepsilon_2 \cos \delta]. \tag{7}$$

Turning to consider the theoretical models of such random packings, the natural first step would be to look for the cases when the term $\mathscr{E} x^n tg\varepsilon_1 tg\varepsilon_2 \cos \delta$ vanishes. This happens, for example, when the components of (X, Ω_1, Ω_2) are independent. The corresponding class of random packings will be called IOSC (independent orientations stereological class). An example of such random packing is obtained when the Poisson field of planar layers is considered. Assume that the union of the

layers is coloured black, and its complement—white. The white random set is of desired type if the layers have random, independent and identically distributed thickness. There are also some other examples.

It is natural to ask how the equation (7) generalises for random sets (white sets in random black-white colouring) of more general nature, when only the assumptions of invariance and piecewise-smoothness of the boundaries is retained.

A rather special process of integration of (1) again leads to a result. In the planar case, this has the form of some relations between distributions associated with marked white interval processes on the arbitrary transection and on some properly defined tangent lines. The terms which involve orientations at the ends of white intervals again disappear for random sets from the IOSC (defined analogously for general random sets).

The following observation gives an idea how rich is the IOSC of these more general random sets (planar case).

Let $\{D_i\}$ be a sequence of random, identically distributed independent bounded convex domains with piecewise-smooth boundaries, $\{t_i\}$ be random set of vectors, with endpoints composing an independent Poisson field in R^2 with constant intensity. The random set $\{P \in R^2;\ P$ is covered by less than k domains from the set $\{t_i D_i\}\}$ belongs to IOSC for every $k = 1, 2, 3, \cdots$ (tD denotes the domain D shifted by vector t).

The following proposition serves to build new examples of random sets from IOSC and increases the interest for the class.

PROPOSITION. *The IOSC is closed with respect to operations of taking set-theoretical intersections and unions (the latter under the additional assumption that the sequence of white marked intervals on the transection is a sequence of independent triples).*

An outstanding unsolved problem in the present theory is to understand the structure of the IOSC, either in special or in general cases. Such an understanding would permit one to judge whether in a given practical situation one can avoid measuring orientations.

References

1. *Stochastic geometry*, Edited by E. F. Harding and D. G. Kendall, Wiley, New York.

2. J. J. Sylvester, *On a funicular solution of Buffon's "Problem of the needle" in its most general form*, Acta. Math. **14** (1891), 185–205.

3. R. V. Ambartzumian, *The solution to the Buffon-Sylvester problem in R^3*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **27** (1973), 53–74.

4. ———, *Combinatorial solution of the Buffon-Sylvester problem*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **23** (1974), 25–31.

ARMENIAN ACADEMY OF SCIENCES
EREVAN, U.S.S.R.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Gaussian Process and How to Approach It

R. M. Dudley

1. de Moivre

Dear listeners, please smile indulgently
At this brief introductory phase
Of somewhat metered, sing-song words;
Not poetry, perhaps, yet not quite prose, I hope.
To speak of "Gaussian" measures, now
Is first of all, to credit C. F. Gauss
For objects he was not the first to find;
Some three quarters of a century before
Carl Friedrich took them up
We find that Abraham de Moivre
Already had described them;
And so to him should go, I think, the fame
Of finding out the laws we call by Gauss's name,
Those bell-shaped curves of density
Whose formulas we now write with ease
In terms of e to minus half x squared
But which de Moivre more laboriously did call
The number which answers to the hyperbolic
Logarithm minus half x times x.
The central limit theorem, too, is found
In Abe de Moivre's book, *Doctrine of Chances*, and if
It's only for binomial distributions, well
Now only after Fourier and then by Paul Lévy
Is rendered easy such a proof;

Without their tools perhaps, dear listener,
You'd demonstrate it as a tour de force;
I've tried without success.
And yet to say "de Moivrian" rather twists the tongue
And it's too late to change the name
So we'll have to find some other way
Of remembering the founder of this line of work;
Let's dedicate now to him, de Moivre
A few moments of our kindest thoughts.

2. Finite-dimensional Gaussian measures. On the real line, the Gaussian or normal probability distribution $N(m, \sigma^2)$ has density $\sigma^{-1}(2\pi)^{-1/2}\exp[-(x-m)^2/2\sigma^2]$ with respect to Lebesgue measure, where $\sigma > 0$ and $m \in R$; $N(m, 0)$ is the unit mass at m. A probability measure on a Euclidean space is called Gaussian if every linear form has a Gaussian law.

Here are two unsolved problems over fifty years old about Gaussian measures and convergence to them.

Problem 1. This was posed by Cantelli in 1917. If X and Y are independent random variables with distribution $N(0,1)$, $f \geqq 0$, f is measurable and $X + f(X)Y$ has a Gaussian distribution $N(0, \sigma^2)$, is f necessarily a constant almost everywhere? Some work on this problem is surveyed by Tricomi [**1967**]. The problem seems to be a mere curiosity, but that will perhaps be unclear until it is solved.

Problem 2. This was implicitly posed by Karl Pearson in 1900: the question of the accuracy of the χ^2 approximation to the X^2 statistic. Given $p_j > 0$, $j = 1, \cdots, m$, $\sum p_j = 1$, $X_1, \cdots, X_n$ independent, $\Pr(X_i = j) = p_j$, $n_j =$ the number of values of $i \leqq n$ such that $X_i = j$,

$$X^2 = \sum_{j=1}^{m} (n_j - np_j)^2/np_j, \qquad \chi^2_{m-1} = \sum_{j=1}^{m-1} G_j^2,$$

where G_j are independent $N(0,1)$. Let Z_s be such that $\Pr(\chi^2_{m-1} \geqq Z_s) = s$. It is known from the multidimensional central limit theorem that $\Pr(X^2 \geqq Z_s) \to s$ as $n \to \infty$ for fixed p_j.

Question. For conventional statistical significance levels including $s = .05$, $.01$ and $.001$, how large must $\min_j np_j$ be (for fixed m) to make the approximation valid within, say, a factor of 5/4?

I have nothing new to say about these finite-dimensional problems; most of the recent progress has been in infinite dimensions.

3. Sample continuity of processes. To define a Gaussian stochastic process, we have a probability space $(\Omega, \mathcal{F}, P)$, a set T, and a function $X(t, \omega)$ on $T \times \Omega$ such that for any $t_1, \cdots, t_n$ in T, $X(t_1, \cdot), \cdots, X(t_n, \cdot)$ have a Gaussian joint distribution on R^n. For simplicity, from now on we only consider Gaussian laws with mean 0. Assume T is equipped with a compact metrizable topology.

Question. Under what conditions is $X(\cdot, \cdot)$ sample continuous, i.e., when is there a process Y such that for each t, $X(t, \omega) = Y(t, \omega)$ almost surely, and such that, for

each ω, $t \to Y(t, \omega)$ is continuous? For this to be true, $t \to X(t, \omega)$ must at least be continuous in probability and hence (using Gaussian) in quadratic mean.

Let $d(s,t) = [E(X(s, \cdot) - X(t, \cdot))^2]^{1/2}$, the intrinsic Hilbert space metric. Looking at T with this metric is essentially equivalent to replacing T by the set $K = \{X(t,\cdot): t \in T\}$ in the Hilbert space $H = L^2(\Omega, P)$ and considering "the" Gaussian process L on H with mean 0 and covariance equal to the inner product. Since Gaussian probability laws are uniquely determined by their means and covariances through L, we can reduce the study of all Gaussian processes to that of L.

A useful sufficient condition for sample continuity can be given in terms of metric entropy as follows. For any $\varepsilon > 0$ let

$$N(\varepsilon) = N(T, d, \varepsilon) = \inf\Big\{n: \exists\, A_1, \cdots, A_n: T \subset \bigcup_{j=1}^{n} A_j, \sup_{s,t \in A_j} d(s,t) \leqq 2\varepsilon,\ j = 1, \cdots, n\Big\}.$$

THEOREM (DUDLEY [**1967**], [**1973**]). *X is sample continuous if*

$$\int_0^1 [\log N(x)]^{1/2}\, dx < \infty.$$

(Note that the square root of the log is the inverse of the exponential of the square, which relates to the Gaussian density.)

THEOREM (FERNIQUE [**1974a,b**]). *If X is a stationary Gaussian process (on R, with joint probability laws invariant under translations of the axis) then it is sample continuous if and only if $\int_0^1 [\log N(x)]^{1/2}\, dx < \infty$, where N is defined for $T = [0,1]$.*

For nonstationary Gaussian processes on [0,1], however, the metric entropy integral condition is no longer necessary for sample continuity. A necessary and sufficient condition was stated by Sudakov [**1971**]; it is not as easy to apply as the metric entropy condition. For proofs of some of Fernique's and Sudakov's results, I have found Chevet [**1974**] a useful reference.

4. The central limit theorem in Banach spaces. Given a probability measure μ on a Banach space, let $X_1, \cdots, X_n, \cdots$ be independent random variables in the Banach space X with distribution μ and $S_n = X_1 + \cdots + X_n$. We say the *central limit theorem* holds for μ if and only if there is a Gaussian probability measure ν on X (i.e., a probability measure such that each continuous linear form has a Gaussian distribution) such that the distribution of $S_n/n^{1/2}$ converges vaguely to ν as $n \to \infty$, in the sense that for every bounded continuous real-valued function f on X, $Ef(S_n/n^{1/2}) \to \int f d\nu$.

We call X a *CLT space* if $\int x\, d\mu = 0$ and $\int \|x\|^2\, d\mu < \infty$ imply the central limit theorem for μ.

Fortet and Mourier [**1955**] proved that L^p is a CLT space for $2 \leqq p < \infty$. There are counterexamples for $p < 2$ and for $p = \infty$. (For $p < 2$, one set of examples uses the existence of stable laws of index < 2.)[1] So for general Banach spaces we need a stronger condition, such as the following:

[1]NOTE ADDED IN PROOF. G. Pisier has proved that X is a CLT space iff it is of *type* 2, i.e. whenever $\sum \|x_n\|^2 < \infty$ and G_n are independent $N(0, 1)$, $\sum G_n x_n$ converges a.s.

THEOREM (JAIN AND MARCUS [1974]). *Given a bounded linear transformation T of a Banach space X into a Banach space Y, with adjoint T^* from Y^* into X^*, let μ be a probability measure on Y^* with $\int z\, d\mu(z) = 0$ and $\int \|z\|^2\, d\mu(z) < \infty$. Let B be the unit ball in X and assume*

$$\int_0^1 [\log N(T(B), \|\cdot\|_Y, \varepsilon)]^{1/2}\, d\varepsilon < \infty.$$

Then the central limit theorem holds in X^* for $T^*\mu$.

References

S. Chevet [1974], *Mesures cylindriques, espaces de Wiener et fonctions aléatoires gaussiennes*, Exposé 8 (Clermont-Ferrand).

R. Dudley [1967], *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Functional Analysis **1**, 290–330. MR **36** #3405.

———, [1973], *Sample functions of the Gaussian process*, Ann. Probability **1**, 66–103.

X. Fernique [1974a], *Des résultats nouveaux sur les processus gaussiens*, C. R. Acad. Sci. Paris Sér. A-B **278**, A363–A365.

———, [1974b], *Minorations des fonctions aléatoires gaussiennes*, Ann. Inst. Fourier, **24** fasc. 2, 61–66.

R. Fortet and E. Mourier [1955], *Les fonctions aléatoires comme éléments aléatoires dans les espaces de Banach*, Studia Math. **15**, 62–79, MR **19**, 1202.

N. Jain and M. Marcus [1974], *Central limit theorems for $C(S)$-valued random variables*, Ann. Probability (to appear).

V. N. Sudakov [1971], *Gaussian random processes and measures of solid angles in Hilbert space*, Dokl. Akad. Nauk SSSR **197**, 43–45 = Soviet Math. Dokl. **12**, 412–415. MR **44** #6027.

F. G. Tricomi [1967], *Einige ungelöste Probleme der klassischen Analysis*, Abh. Math. Sem. Univ. Hamburg **31**, 25–32. MR **36** #1.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139, U.S.A

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Semigroups of Invariant Operators

J. Faraut

I. Introduction. Let G be a locally compact group and K a compact subgroup of G, and let X be the homogeneous space $X = G/K$. We denote by $C_0(X)$ the space of continuous complex functions on X vanishing at infinity. With the norm $\|f\| = \mathrm{Sup}_{x\in X} |f(x)|$ the space $C_0(X)$ is a Banach space.

To any element g of G we associate the transformation τ_g of $C_0(X)$ defined by $\tau_g f(x) = f(g^{-1}x)$.

An operator A on the space $C_0(X)$ is said to be invariant if it commutes with the transformations τ_g.

Let us consider a strongly continuous semigroup of contractions of the space $C_0(X)$, i.e., a family $\{P_t\}_{t\geqq 0}$ of operators of $C_0(X)$ such that

$$P_0 = I, \qquad P_t P_s = P_{t+s}, \tag{1}$$

$$\|P_t\| \leqq 1, \tag{2}$$

$$\forall f \in C_0(X), \qquad \lim_{t\to 0} \|P_t f - f\| = 0 \tag{3}$$

(we shall say a contraction semigroup).

Let us assume that each operator P_t is invariant. Let (D_A, A) be the infinitesimal generator of the semigroup, i.e., the operator defined by

$$D_A = \left\{ f \in C_0(X) \;\middle|\; \lim_{t\to 0} \frac{1}{t}(P_t f - f) \text{ exists} \right\},$$

$$Af = \lim_{t\to 0} \frac{1}{t}(P_t f - f).$$

The infinitesimal generator (D_A, A) has the following properties:

(a) The operator (D_A, A) is closed, and its domain D_A is dense.

(b) The operator (D_A, A) is dissipative, i.e., for any function f of the domain D_A, $f(x) = \|f\| \Rightarrow \mathrm{Re}\, f(x) \leqq 0$.

(c) The operator (D_A, A) commutes with the transformations τ_g, i.e., for any element g of G, $\tau_g D_A = D_A$, $\tau_g A = A\tau_g$. We say also that the operator (D_A, A) is invariant.

At present one can prove that an operator satisfying (a), (b), and (c) is the infinitesimal generator of a contraction semigroup in three cases:

(1) The group G is compact.

(2) The convolution algebra $M^\natural(G)$ of bounded measures on G bi-invariant by K is commutative. This hypothesis is satisfied if G is abelian; also if (G, K) is a Riemannian symmetric pair, then the space X is a Riemannian symmetric space.

(3) The group G is a Lie group and the domain D_A contains the C^∞ functions on X with compact support.

For $\boldsymbol{R}^n$ the result is due to the author [2] (1969), for a Riemannian symmetric space to the author and Harzallah [3] (1971), for a pair (G, K) with the algebra $M^\natural(G)$ commutative to Hirsch and Roth [4] (1973) and for a Lie group to Duflo [1] (1974). The case of positive contractions was studied by Hunt [5] (1956).

II. Properties of dissipative operators.

PROPOSITION II.1. *The infinitesimal generator (D_A, A) of a contraction semigroup $\{P_t\}_{t\geqq 0}$ is dissipative.*

Let f be a function of D_A and x_0 a point of X such that $f(x_0) = \|f\|$; because P_t is a contraction we have

$$\forall\, x, \qquad \operatorname{Re} P_t f(x) \leqq |P_t f(x)| \leqq \|f\| = f(x_0)$$

and then

$$\operatorname{Re} Af(x_0) = \lim_{t\to 0} \frac{1}{t} [\operatorname{Re} P_t f(x_0) - f(x_0)] \leqq 0.$$

PROPOSITION II.2. *If (D_A, A) is a dissipative operator then $\forall\, \lambda > 0$, $\forall\, f \in D_A$, $\|\lambda f - Af\| \geqq \lambda \|f\|$.*

Let f belong to D_A, and let x_0 be a point of X such that $|f(x_0)| = \|f\|$. We can suppose $f(x_0) \geqq 0$ (if necessary one can multiply f by a complex constant) and then $\operatorname{Re} Af(x_0) \leqq 0$; it follows that

$$\|\lambda f - Af\| \geqq \lambda f(x_0) - \operatorname{Re} Af(x_0) \geqq \lambda f(x_0) = \lambda \|f\|.$$

REMARK 1. From Proposition II.2 it follows that if (D_A, A) is a closed dissipative operator, $\forall\, \lambda > 0$, $(\lambda I - A)D_A$ is closed.

REMARK 2. One can prove (see [6]) that a dissipative operator with dense domain is closable and that its closure is also a dissipative operator.

From the Hille-Yosida theorem one can deduce the following (similar to a theorem of Lumer and Phillips [6]).

THEOREM II.3. *An operator (D_A, A) on $C_0(X)$ generates a contraction semigroup if and only if:*

(a) *The operator (D_A, A) is closed and its domain D_A is dense.*

(b) *The operator* (D_A, A) *is dissipative and* $\forall\ \lambda > 0$, $(\lambda I - A)D_A = C_0(X)$.

We give another form of this theorem.

THEOREM II.4. *Let* (D_A, A) *be an operator on* $C_0(X)$ *such that:*
(a) *Its domain* D_A *is dense.*
(b) *The operator* (D_A, A) *is dissipative.*
(c) $\forall\ \lambda > 0$, $(\lambda I - A)D_A$ *is dense.*
Then the operator (D_A, A) *is closable and its closure generates a contraction semigroup.*

III. The group G is compact. Let G be a compact group, K a compact subgroup, and $X = G/K$. Let $\{T_i\}$ be the set of (classes of) irreducible representations of G which occur in the decomposition of $L^2(X)$. Let $\chi_i = \text{Trace}\ (T_i)$ be the character and d_i the dimension of T_i. For a function f of $L^2(X)$ we put $\pi_i f = d_i \chi_i * f$ (we consider a function on X as a function on G right invariant by K). The map π_i is an orthogonal projection. We define $H_i = \pi_i L^2(X)$. H_i is a finite dimensional subspace of $C(X)$. We have the following:

PROPOSITION III.1. *Let* (D_A, A) *be an operator on* $C(X)$, *closed, with dense domain and invariant. Then the subspaces* H_i *are contained in* D_A, *and* A *maps* H_i *into* H_i.

And we deduce

COROLLARY III.2. *Let* (D_A, A) *be an operator on* $C(X)$, *closed, with dense domain and invariant. If the operator* (D_A, A) *is injective then its range is dense.*

The restriction A_i to H_i is an injective endomorphism of the finite dimensional space H_i; it follows that A_i is onto. Then the range of (D_A, A) contains the subspaces H_i, and by the Peter-Weyl theorem its range is dense.

THEOREM III.3. *Let* (D_A, A) *be an operator of* $C(X)$, *dissipative, with dense domain and invariant. Then* (D_A, A) *is closable and its closure* $(D_{\bar{A}}, \bar{A})$ *generates a contraction semigroup of* $C(X)$.

From Proposition II.2 and Remark 2 we know that for $\lambda > 0$ the operator $\lambda I - \bar{A}$ is injective, and from Corollary III.2 it follows that its range is dense. The result follows from Theorem II.3 or II.4.

IV. The algebra $M^1(G)$ is commutative. We assume now that the convolution algebra $M^1(G)$ of bounded measures on G bi-invariant by K is commutative.

PROPOSITION IV.1. *Let* (D_A, A) *be an operator of* $C_0(X)$, *closed, with dense domain and invariant. If* (D_A, A) *is injective then its range is dense.*

We use the Hahn-Banach theorem. Let N be the space of bounded measures on X orthogonal to the range of (D_A, A). We shall prove that $N = \{0\}$. Let μ be a measure of N invariant by K; we have

$$\forall f \in D_A,\ \forall g \in G, \qquad \int A\tau_g f\, d\mu = \int \tau_g Af\, d\mu = 0,$$

that is, $Af * \check{\mu} = 0$ where the measure $\check{\mu}$ is defined by

$$\int f \, d\check{\mu} = \int f(g^{-1}) \, d\mu(g)$$

and where we consider functions and measures on X as functions and measures on G right invariant by K.

From the commutativity of $M^{\natural}(G)$ we deduce, if f is invariant by K, $\check{\mu} * Af = 0$ and $A(\check{\mu} * f) = 0$ because (D_A, A) is closed and invariant. Then for any f of D_A invariant by K, $\int f \, d\mu = 0$, and, as μ is invariant by K, $\forall f \in D_A$, $\int f \, d\mu = 0$.

The domain D_A is dense so that $\mu = 0$. From the invariance of the space N it follows that $N = \{0\}$.

We deduce as before a theorem similar to Theorem III.3.

V. The group G is a Lie group. We assume now that G is a Lie group so that X is a C^∞-differentiable manifold and we use the distribution theory. We denote by $\mathscr{D}(X)$ the space of C^∞ functions on X with compact support.

We say that a distribution T on X is dissipative if

(a) $\forall f \in \mathscr{D}(X), f(0) = \|f\| \Rightarrow \operatorname{Re} \int f \, dT \leqq 0$ (0 denotes the coset eK),

(b) T is invariant by K.

Every dissipative and invariant operator on $C_0(X)$ with domain $\mathscr{D}(X)$ has the form

$$Af(g0) = \int f(gx) \, dT(x)$$

where T is a dissipative distribution on X. We have in particular $Af(0) = \int f \, dT$.

If we consider functions and distributions on X as functions and distributions on G right invariant by K we have $Af = f * \check{T}$.

Outside of a neighbourhood of 0 a dissipative distribution T is a bounded measure, so that it is a Radon measure outside of $\{0\}$, and can be written $T = S + \sigma$ where S is a distribution with compact support and σ a bounded measure.

THEOREM V.1. *Let T be a dissipative distribution on X, and let (D_A, A) be the operator on $C_0(X)$ defined by*

$$D_A = \mathscr{D}(X), \qquad Af(g0) = \int f(gx) \, dT(x).$$

Then the closure of (D_A, A) generates a contraction semigroup.

We have to prove that for all $\lambda > 0$ the space $(\lambda I - A)\mathscr{D}(X)$ is dense. We shall use the Hahn-Banach theorem: let ν be a bounded measure on X such that

$$\forall f \in \mathscr{D}(X), \quad \int (\lambda f - Af) \, d\nu = 0,$$

that is, $\lambda \nu - \nu * T = 0$ and by regularization, $\forall \alpha \in \mathscr{D}(G)$, $\lambda \, \alpha * \nu - \alpha * \nu * T = 0$.

The distribution $\check{T}$ is, as T, dissipative; then by Proposition II.2, $\forall \alpha \in \mathscr{D}(G)$, $\alpha * \nu = 0$. It follows that $\nu = 0$.

Conversely one can prove that the domain of the infinitesimal generator of any invariant contraction semigroup contains $\mathscr{D}(X)$ so that there is a bijection between the dissipative distributions and the invariant contraction semigroups.

It is possible to give for the dissipative distributions an integro-differential representation analogue to the Levy-Kinchine formula.

References

1. M. Duflo, *Semi-groupes de mesures complexes sur un groupe de Lie*, Paris VII, 1974 (preprint).

2. J. Faraut, *Semi-groups de mesures complexes et calcul symbolique sur les générateurs infinitésimaux de semi-groupe d'opérateurs*, Ann. Inst. Fourier (Grenoble) **20** (1970), 235–301.

3. J. Faraut and K. Harzallah, *Semi-groupes d'opérateurs invariants et opérateurs dissipatifs invariants*, Ann. Inst. Fourier (Grenoble) **22** (1972), 147–164.

4. F. Hirsch and J. P. Roth, *Opérateurs dissipatifs et codissipatifs invariants sur un espace homogène, Théorie du potentiel et analyse harmonique*, Lecture Notes in Math., vol. 404, Springer-Verlag, New York, 1974.

5. G. A. Hunt, *Semi-groups of measures on Lie groups*, Trans. Amer. Math. Soc. **81** (1956), 264–293. MR **18**, 54.

6. G. Lumer and R. S. Phillips *Dissipative operators in a Banach space*, Pacific J. Math. **11** (1961), 679–698. MR **24** #A2248.

UNIVERSITÉ LOUIS PASTEUR
STRASBOURG, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Some Mathematical Problems Arising in Robust Statistics

Peter J. Huber

1. Robustness and continuous statistical functionals. Let $x_1, x_2, \cdots$ be observations with values in some complete separable metric space $\mathcal{X}$. Many test statistics and estimates depend on the sample $(x_1, \cdots, x_n)$ only through the empirical distribution

$$\mu_n = \frac{1}{n} \sum \delta_{x_i}, \tag{1}$$

where δ_x denotes the pointmass 1 at x, and thus they can be written as $T_n(x_1, \cdots, x_n) = T(\mu_n)$. For the sake of simplicity we shall assume that the functional T is defined everywhere on the set $\mathcal{M}_1$ of all probability measures on $\mathcal{X}$.

The basic stability or robustness requirement is (Hampel [**1968**], [**1971**]): A small change in μ_n should induce only a small change in $T(\mu_n)$. Here "small" can mean

small changes in many or all of the x_i (rounding and grouping errors);

large changes in a few of the x_i (occasional gross errors).

This amounts to requiring that T should be continuous for the weak topology, i.e., for the weakest topology on $\mathcal{M}_1$ such that the functionals

$$\mu \to \int \psi \, d\mu \tag{2}$$

are continuous for all bounded, continuous functions ψ on $\mathcal{X}$. Conversely, all weakly continuous, linear functionals on $\mathcal{M}_1$ then are of this form (2) with a bounded, continuous ψ.

The standard classical statistics, like the sample mean, $T(\mu_n) = \int x \, d\mu_n$ do not satisfy the robustness requirement. But from the practical point of view—especially if the data are to be processed with little, if any, human intervention—it is crucially important as a safeguard against catastrophes caused by occasional gross errors.

2. Metrics on $\mathcal{M}_1$. If we are to treat quantitative questions, we must metrize the

weak topology. While the topology is forced upon us by the outside requirement of robustness, the metric can be chosen according to convenience.

On the real line, $\mathscr{X} = R$, we can use the *Lévy metric*:

$$d_L(\mu, \nu) \leqq \varepsilon \text{ iff, for all } x,$$
$$\mu\{(-\infty, x - \varepsilon)\} - \varepsilon \leqq \nu\{(-\infty, x)\} \leqq \mu\{(-\infty, x + \varepsilon)\} + \varepsilon.$$

For general $\mathscr{X}$, there is the *Prohorov metric*:

$d_P(\mu, \nu) \leqq \varepsilon$ iff, for all measurable sets A, $\mu(A) \leqq \nu(A^\varepsilon) + \varepsilon$.

Here, A^ε denotes the closed ε-neighborhood of A.

This has a particularly nice interpretation due to Strassen [**1965**]: $d_P(\mu, \nu) \leqq \varepsilon$ iff there is a measure λ on $\mathscr{X} \times \mathscr{X}$ with the marginals μ, ν and such that $\lambda\{(x, y) \mid d(x, y) \leqq \varepsilon\} \geqq 1 - \varepsilon$.

For certain purposes, the *bounded Lipschitz metric* d_{BL} may be more convenient. Assume that the given metric d on $\mathscr{X}$ is bounded by 1 (if necessary, replace it by $d/(1 + d)$). Then $d_{BL}(\mu, \nu) \leqq \varepsilon$ iff for all functions ϕ satisfying $|\phi(x) - \phi(y)| \leqq d(x, y)$ we have $\left|\int \phi\, d\mu - \int \phi\, d\nu\right| \leqq \varepsilon$.

Also this has a nice interpretation (proved by Kantorovič and Rubinšteĭn [**1958**] for compact $\mathscr{X}$, but the result generalizes): $d_{BL}(\mu, \nu) \leqq \varepsilon$ iff there is a measure λ on $\mathscr{X} \times \mathscr{X}$ with the marginals μ, ν such that $\int d(x, y)\, d\lambda \leqq \varepsilon$.

It is straightforward to show that $d_P^2 \leqq d_{BL} \leqq 2d_P$.

In the following, d_* will denote any metric (or pseudo-metric) compatible with the weak topology and with the usual Euclidean metric on lines: For any μ_0, $\mu_1 \in \mathscr{M}_1$,

$$d_*(\mu_t, \mu^s) = O(|t - s|) \tag{3}$$

where

$$\mu_t = (1 - t)\mu_0 + t\mu_1, \qquad 0 \leqq t \leqq 1. \tag{4}$$

3. Differentiable functionals. Differentiable statistical functionals were introduced by von Mises [**1937**]; their importance for the heuristics of robustness was first recognized by Hampel [**1968**].

Assume that T can be linearized at a fixed μ in the sense that there is a linear functional L such that

$$T(\nu) - T(\mu) = L(\nu - \mu) + o(d_*(\mu, \nu)). \tag{5}$$

If T is continuous at $\nu = \mu$, then also L is; hence it must have the form (2). Without loss of generality we standardize $\phi = \phi_\mu$ such that $\int \phi_\mu\, d\mu = 0$, and then (5) can be rewritten as

$$T(\nu) - T(\mu) = \int \phi_\mu\, d\mu + o(d_*(\mu, \nu)). \tag{6}$$

Fréchet differentiability (6) is often difficult to establish, but ϕ_μ is easily calculable as a Volterra derivative: Put $\mu_0 = \mu$, $\mu_1 = \delta_x$ in (4); then

$$\phi_\mu(x) = \lim_{t \to 0} \frac{T(\mu_t) - T(\mu)}{t}. \tag{7}$$

In view of this, $\psi_\mu(x)$ can be interpreted as the differential influence of one observation with value x toward the value of the estimate (Hampel [**1968**]).

Assume now that the x_i are independent with common distribution μ. The typical asymptotical normality proofs use linear approximations. For instance, one may be able to show that

$$n^{1/2}(T(\mu_n) - T(\mu)) = \frac{1}{n^{1/2}} \sum \psi_\mu(x_i) + o_p(1). \tag{8}$$

Here the right-hand side is asymptotically normal in view of the central limit theorem. If

$$d_*(\mu, \mu_n) = O_p(n^{-1/2}) \tag{9}$$

then (6) implies (8), but in general neither of the two relations implies the other.

Since (9) would allow very streamlined asymptotic normality proofs, this raises a puzzling question: For which d_* and μ do we have (9)?

Assume first $\mathscr{X} = R$. Then it is well known that (9) holds for d_L and arbitrary μ. If $\mu = U(0,1)$ is the uniform distribution on the unit interval, it also holds for d_P and d_{BL} (Dudley [**1969**]). It is false if μ has sufficiently long tails (rational tails suffice, $\mu\{|x| > t\} \sim t^{-k}$ for some k).

If μ is the uniform distribution on a k-dimensional cube, then (9) fails for d_P and d_{BL}, if $k \geqq 3$, and the question seems to be open for $k = 2$.

Apart from that, we seem to lack the machinery for proving or disproving Fréchet differentiability. For instance, linear combinations of order statistics correspond to functionals of the form

$$T(\mu) = \int F^{-1}(t) M(dt), \tag{10}$$

where $F(x) = \mu\{(-\infty, x)\}$ and M is some signed measure on (0, 1). The gap between the necessary and the sufficient conditions for differentiability of (10) is uncomfortably wide.

4. Capacities and exact finite sample results. Classical statistical decision theory is concerned with parametric families $(P_\theta)_{\theta \in \Theta}$ of probability measures and tries to find "optimal" (admissible, minimax, etc.) decision procedures, assuming that the true underlying distribution belongs to the family (P_θ). But, to be honest, we can at best hope that the true probability distribution P lies in the set $\mathscr{P}_\theta = \{P \in \mathscr{M}_1 \mid d_*(P_\theta, P) \leqq \varepsilon\}$ for some θ. Of course, then there is no longer a single "true" value of the parameter θ, but concepts like interval estimates, confidence intervals, etc., remain meaningful.

It turns out that the approach works if (and essentially only if) the sets $\mathscr{P}_\theta$ can be described as $\mathscr{P}_\theta = \{P \in \mathscr{M}_1 \mid \forall\, A,\ P(A) \leqq v_\theta(A)\}$ where v_θ is an alternating capacity of order 2.

A real valued function v on the set of all subsets of $\mathscr{X}$ is called a *capacity* if

$$A \subset B \Rightarrow v(A) \leqq v(B),$$
$$A_n \uparrow A \Rightarrow v(A_n) \uparrow v(A),$$
$$A_n \downarrow A,\ A_n \text{ closed} \Rightarrow v(A_n) \downarrow v(A),$$

and it is called alternating of order 2 (or 2-*alternating*) if $v(A \cup B) + v(A \cap B) \leqq v(A) + v(B)$.

All our capacities will furthermore be normalized such that $v(\emptyset) = 0$, $v(\mathcal{X}) = 1$.

The requirement that v is 2-alternating is crucial; it is equivalent to the following property: For every upper semicontinuous function $h \geqq 0$ on $\mathcal{X}$ one has

$$\sup_{P \leqq v} \int h \, dP = \sup \int_0^\infty P\{h > t\} \, dt = \int_0^\infty v\{h > t\} \, dt.$$

In other words, not only the upper bounds for probabilities, but also those for expectations, can be expressed in terms of v.

Interestingly, Prohorov neighborhoods can be expressed in terms of a 2-alternating capacity (the set function $A \to P(A^\varepsilon) + \varepsilon$ is 2-alternating).

There are at least two explicitly solvable special cases fitting into this general framework:

(i) $(P_\theta)_{\theta \in \{\theta_1, \theta_2\}}$ has only two members (Strassen [**1964**], Huber [**1965**], Huber and Strassen [**1973**]). This generalizes the Neyman-Pearson lemma to capacities. At the same time, this gives a natural generalization of the Radon-Nikodym derivative to capacities.

(ii) $(P_\theta)_{\theta \in R}$ is a location family, where P_θ has density $f(x - \theta)$ with respect to Lebesgue measure, such that $-\log f$ is convex (Huber [**1968**]).

The obvious and intriguing question now is: How much of classical probability and statistics generalizes to capacities?

References

R. M. Dudley **(1969)**, *The speed of mean Glivenko-Cantelli convergence*, Ann. Math. Statist. **40**, 40–50. MR **38** #5270.

F. R. Hampel **(1968)**, *Contributions to the theory of robust estimation*, Unpublished Ph.D. Thesis, University of California, Berkeley.

——— **(1971)**, *A general qualitative definition of robustness*, Ann. Math. Statist. **42**, 1887–1896. MR **46** #1013.

P. J. Huber **(1965)**, *A robust version of the probability ratio test*, Ann. Math. Statist. **36**, 1753–1758. MR **32** #3209.

——— **(1968)**, *Robust confidence limits*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **10**, 269–278. MR **39** #3661.

P. J. Huber and V. Strassen **(1973)**, *Minimax tests and the Neyman-Pearson lemma for capacities*, Ann. Statist. **1**, 251–263, ibid. **2**, 223–224.

L. Kantorovič and G. Š. Rubinštein **(1958)**, *On a space of completely additive functions*, Vestnik, Leningrad. Univ. **13**, no. 7, 52–59. (Russian) MR **21** #808.

R. von Mises **(1937)**, *Sur les fonctions statistiques*, Conférence de la Réunion Internationale des Math., Gauthier-Villars, Paris, pp. 1–8; Also in: *Selecta R. von Mises*. Vol. II, pp. 388–394.

V. Strassen **(1964)**, *Messfehler und Information*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **2**, 273–305. MR **29** #2120.

——— **(1965)**, *The existence of probability measures with given marginals*, Ann. Math. Statist. **36**, 423–439. MR **31** #1693.

SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH, SWITZERLAND

Actes du Congrès International des Mathematiciens
Vancouver, 1974

Théorie du Potentiel Récurrent (Résultats Récents)

J. Neveu

1. Soit P un noyau markovien défini sur l'espace mesurable $(E, \mathscr{A})$ que nous supposerons séparable. L'ensemble des fonctions mesurables $h: E \to (0, 1)$ étant désigné par H, nous poserons $U_h = \sum_N (PM_{1-h})^n P$ pour toute fonction $h \in H$ en convenant de noter M_k le noyau de multiplication par la fonction k $(k \in H)$; ces noyaux positifs U_h $(h \in H)$ vérifient les relations

$$U_h = \sum_N (U_k M_{k-h})^n U_k \quad \text{lorsque } h \leqq k \text{ dans } H, \tag{1}$$

et donc aussi les équations résolvantes $U_h = U_k + U_h M_{k-h} U_k = U_k + U_k M_{k-h} U_h$ $(h \leqq k)$.

Les noyaux U_h ont plusieurs interprétations probabilistes intéressantes. D'abord si $(X_n, n \in N)$ est une chaîne de Markov de probabilité de transition P et si $\nu_A = \min (n: n \geqq 1, X_n \in A)$, on a $U_{1_A} f = E \cdot (\sum_1^{\nu_A} f(X_n))$ et en particulier $U_{1_A} 1_A = P \cdot (\nu_A < \infty)$. Deuxièmement la résolvante d'un processus markovien de sauts $(X_t, 0 \leqq t < \zeta)$ de durées de séjour exponentielles de paramètres $q(x)$ $(x \in E)$ et de lois de sauts données par le noyau P vaut $\lambda V_\lambda = M_{h_\lambda} + M_{1-h_\lambda} U_{h_\lambda} M_{h_\lambda}$ ou $h_\lambda = \lambda/\lambda + q$. Enfin si $(X_t, 0 \leqq t < \zeta)$ est un processus de Hunt sur un espace l.c.d. (= localement compact à base dénombrable) et si $(U_\lambda, \lambda > 0)$ désigne sa résolvante, alors en prenant $P = U_1$, on trouve que $U_h f = E \cdot (\int_0^\infty \exp(-H_t) f(X_t)\, dt)$ si $H_t = \int_0^t h(X_s)\, ds$; dans ce cas d'ailleurs les noyaux U_h peuvent être définis plus généralement pour toute fonction bornée et même pour toute fonction mesurable $h: E \to R_+$ telle que $H_t < \infty$ p.s. $(t \in R_+)$ tout en vérifiant encore les relations (1) ci-dessus [**2**].

(Considérons les noyaux positifs Q tels que $P \geqq PQ$ (ou que $P \geqq QP$) parmi lesquels figurent les M_h $(h \in H)$; les noyaux positifs $U_Q = \sum_N (P(I - Q))^n P$ qui leur sont associés vérifient des relations analogues à (1) et notamment la relation

$\sum_N (U_Q Q)^n U_Q = U_0$ qui entraîne que $U_0 Q$ et $Q U_0$ sont des noyaux potentiels. Cette remarque devrait être intéressante.)

2. Etant donné une mesure positive m sur $(E, \mathscr{A})$, le noyau markovien P est dit m-irréductible si l'absolue continuité $m \ll \sum_N^* P^n(x, \cdot)$ a lieu pour tout $x \in E$. D'après un résultat de Jain et Jamison ([**6**], voir aussi [**10**]) un tel noyau P ne peut être que transitoire ou récurrent; de manière plus précise ou bien il existe une fonction mesurable strictement positive h telle que $U_0 h$ soit bornée, ou bien après s'être restreint à un sous-ensemble de E P-invariant de mesure m pleine convenable on a $U_h h = 1$ pour tout $h \in H$ tel que $m(h) \neq 0$ tandis que $U_0 f \equiv +\infty$ dès que $m(f) \neq 0$. On montre d'autre part que P est m-irréductible dès que $\sum_N P^n(\cdot, A) > 0$ *presque partout* pour tout A non négligeable, c'est à dire dès que l'opérateur défini par P sur $L^\infty(m)$ est irréductible, à moins que les mesures m et $\sum_N P^n(x, \cdot)$ ne soient étrangères pour tout $x \in E$ [**9**]. Dans le cas de la promenade aléatoire à gauche de loi λ définie sur un groupe l.c.d. G par $P_\lambda f = \int_G f(\cdot, x)\, d\lambda(x)$, ce noyau P_λ est irréductible par rapport à une mesure de Haar à droite μ de G si et seulement si G est le plus petit groupe fermé portant λ et s'il existe un entier $n \geqq 1$ tel que λ^{n*} et μ ne soient pas étrangères; une telle probabilité λ sur G est dite *étalée*.

3. Dès 1956 Harris introduisit la condition de récurrence: "$U_h h = 1$ pour $h \in H$ tel que $m(h) \neq 0$" (sous la forme équivalente: $U_A 1_A = 1$ si $m(A) \neq 0$) pour en déduire l'existence et l'unicité à un facteur constant près d'une mesure positive σ-finie et P-invariante μ sur E telle en outre que $m \ll \mu$. Nous avons démontré que cette condition entraîne aussi l'existence d'une fonction strictement positive h_1 dans H telle que $U_{h_1} \geqq 1 \otimes \mu$, ce qui permet de développer une théorie du potentiel (du type logarithmique) très complète pour le noyau P à partir du noyau positif

$$W_{h_1} = \sum_N (V_{h_1} M_{h_1})^n V_{h_1} \qquad (V_{h_1} = U_{h_1} - 1 \otimes \mu \geqq 0)$$

(voir [**10**]). Remarquons que le résultat de Jain et Jamison, celui de Harris et le notre affirmant l'existence d'une fonction h_1 telle que $U_{h_1} \geqq 1 \otimes \mu$ se déduisent assez simplement du lemme suivant en se servant des équations résolvantes.

LEMME [**10**]. *Si p et q sont deux fonctions mesurables strictement positives sur $(E, \mathscr{A})^{2\otimes}$ et si m est une mesure positive non nulle, la fonction $r(x, z) = \int_E p(x, y) q(y, z)\, dm(y)$ peut être minorée par le produit $a(x)b(z)$ de deux fonctions mesurables strictement positives sur $(E, \mathscr{A})$.*

Si E est un espace l.c.d., il n'est pas difficile de montrer qu'un noyau markovien P est récurrent au sens de Harris dès qu'il est fortement fellérien (Pf est continue si f est borélienne bornée) et que $U_0(\cdot, 0) \equiv +\infty$ pour tout ouvert non vide 0. Semblablement la promenade aléatoire P_λ sur le groupe l.c.d. G est récurrente au sens de Harris dès que λ est étalée et que $\sum_N \lambda^{n*}(0) \equiv \infty$ (0: ouvert non vide).

4. Pour développer la théorie du potentiel on introduit le cône convexe S des fonctions f dites spéciales, c'est à dire des fonctions positives telles que $U_h(f)$ soit bornée pour tout $h \in H$ ($\mu(h) \neq 0$) ou ce qui est équivalent telles que $W_{h_1}(f)$ soit

bornée (our un h_1 fixé); ces fonctions sont nécessairement μ-intégrables. On introduit aussi le vectoriel Σ des charges spéciales, c'est à dire des fonctions de S-S de μ-intégrale nulle. S et Σ ont la propriété importante d'être stable par P et plus généralement par les $M_h U_h$. Un des résultats principaux de la théorie est alors que pour tout $f \in \Sigma$, la fonction $W_{h_1} f$ est à une constante additive près l'unique solution bornée de l'équation de Poisson $(1 - P)u = Pf$. Les fonctions de Σ satisfont aussi à un principe semicomplet du maximum relativement à W_{h_1} (pour tout ceci, voir [10]).

Métivier ([7] a démontré pour les fonctions spéciales le théorème ergodique quotient suivant: Quelles que soient les probabilités λ_1, λ_2 sur E et quelles que soient $f_1, f_2 \in S$ $(\mu(f_2) \neq 0)$,

$$\lambda_1\Big(\sum_1^n P^m f_1\Big) \Big/ \lambda_2\Big(\sum_1^n P^m f_2\Big) \to \mu(f_1)/\mu(f_2) \quad \text{lorsque } n \uparrow \infty.$$

(Pour des fonctions de $L^1(\mu)$, le résultat analogue est faux comme l'a montré Krengel.) En appliquant ce théorème aux promenades aléatoires gauche et droite associée à une probabilité λ définie sur un groupe G l.c.d., on en déduit que G ne peut porter de promenade récurrente au sens de Harris que si G est unimodulaire.

Au départ je me suis intéressé aux fonctions spéciales parceque les fonctions bornées spéciales f peuvent être caractérisées comme étant les fonctions positives bornées telles que $U_{cf} \geqq 1 \otimes \nu$ pour une mesure positive non nulle ν et une constante $c > 0$. Il s'en suit [4] que lorsque le noyau markovien P admet un noyau dual $\hat{P}$ par rapport à μ, c'est à dire s'il existe $\hat{P}$ est un noyau markovien tel que $\int Pfg \, d\mu = \int f\hat{P}g \, d\mu$, les fonctions bornées positives f telles que $U_{cf} \geqq 1 \otimes \mu$ pour une constante $c > 0$ sont exactement les fonctions bornées spéciales à la fois pour $\hat{P}$ et pour P (à une modification de $\hat{P}$ près). On notera que les fonctions spéciales pour P, resp. pour $\hat{P}$, ne sont pas nécessairement les mêmes.

Lorsque E est un espace l.c.d. et lorsque P est féllérien, il existe une fonction spéciale continue strictement positive de sorte que toute fonction continue positive à support compact est nécessairement spéciale. En général les constantes positives ne sont pas spéciales; pour que 1 soit spéciale il est en fait nécessaire et suffisant d'après Brunel et Revuz [4] que P soit quasi-compact.

5. L'opérateur W_h peut être déterminé explicitement dans un certain nombre de cas. D'abord si E est fini et si P est une matrice markovienne récurrente irréductible, le noyau W_h est donné explicitement par

$$W_h = (I - P + Ph \otimes h\mu)^{-1}P - (\mu(h))^{-1}(1 \otimes \mu);$$

ce noyau existe ici pour tout $h \geqq 0$ non nul et d'après nos résultats il est positif si h est assez petit.

Pour un processus de Hunt récurrent au sens de Harris les relations entre les opérateurs U_h et W_h et l'opérateur infinitésimal A du processus ont été étudiées dans [1] où l'on trouvera comme application la détermination de l'opérateur W_h pour les processus de diffusion récurrents sur R. D'autre part dans le cas d'un processus à accroissements independants stationaires sur R de fonction de Lévy ϕ,

en supposant la fonction $1/\psi$ intégrable à l'infini et le processus récurrent au sens de Harris, on trouve que le noyau W_h possède la densité

$$-l(y-x)+\left(\int (l(z-x)+l(y-z))h(z)\,dz \Big/ \int h(z)\,dz\right)+c_h$$

où c_h est la constante $\iint l(y-x)h(y)h(x)\,dx\,dy/(\int h(z)\,dz)^2$ et où l est la fonction continue donnée par la formule

$$l(x)=\int_R (e^{ixt}-1-ixt\,1_{|t|\leq 1})\frac{dt}{\psi(t)}+ax \qquad (a \text{ constante arbitraire}).$$

6. Si G est un groupe opérant sur E en laissant P et μ invariants et si on pose $T_g f = f(g^{-1}\cdot)$ sur E, il est facile de voir que $T_g W_{h_1} T_g^{-1} = W_{T_g h_1}$ et que le cône S des fonctions spéciales est invariant par les T_g ($g \in G$). En se plaçant alors dans le cas où P admet un noyau dual $\hat{P}$ par rapport à μ pour simplifier, la question se pose de savoir si par une modification inessentielle de W_{h_1} en $W = W_{h_1} - (a \otimes \mu + 1 \otimes \hat{a}\mu)$ où a et $\hat{a}$ désignent deux fonctions positives, il est possible de construire un opérateur potentiel W invariant par le groupe $(T_g,\ g \in G)$. Cette question revient à trouver des fonctions a, $\hat{a}$ telles que

$$T_g a - a + \frac{1}{\mu(h_1)} W_{h_1}(T_g h_1) = c_g, \quad T_g \hat{a} - \hat{a} + \frac{1}{\mu(h_1)} \hat{W}_{h_1}(T_g h_1) = \hat{c}_g \qquad (g \in G)$$

pour des constantes $\hat{c}_g$, c_g dont la somme doit en outre être égale pour tout g fixé à la fonction constante $(W_{h_1}(T_g h_1) + W_{T_g h_1}(h_1))/\mu(h_1)$.

La question précédente admet une réponse immediate s'il existe une fonction spéciale et cospéciale invariante par les T_g comme c'est le cas pour tout espace homogène *compact* $E = G/K$ et tout noyau P de convolution par une probabilité K-invariante sur E, récurrent au sens de Harris.

D'autre part lorsque $G = E$ est un groupe l.c.d. unimodulaire, Brunel et Revuz [3] ont pu construire un opérateur potentiel invariant W en définissant les fonctions a et $\hat{a}$ par les formules

$$\lim_{n\to\infty} W_{h_1} P_n(x,\cdot) = a(x)\mu, \qquad \lim_{n\to\infty} \hat{W}_{h_1} \hat{P}_n(x,\cdot) = \hat{a}(x)\mu$$

(limites vagues sur E, uniformes en x sur tout compact), la difficulté étant évidemment de montrer que ces limites existent. Le noyau W qui est un noyau de convolution sur G applique les fonctions spéciales dans les fonctions finies bornées supérieurement de telle manière que $(I-P)Wf = Pf$ pour tout $f \in S$; en outre $\lim_{n\to\infty} P_k f = Wf$ pour toute charge spéciale f à la fois pour P et $\hat{P}$.

Bibliographie

1. F. Bertein et C. Cocozza, *Note*, C. R. Acad. Sci. Paris (à paraître).

2. F. Bronner, *Méthodes probabilistes pour la détermination de résolvantes markoviennes associées à un noyau verifiant le principle complet du maximum*, Ann. Inst. H. Poincaré **11** (1975).

3. A. Brunel et D. Revuz, *Marches récurrentes au sens de Harris sur les groupes localement compacts*, Ann. École. Norm. Sup. **7** (1974).

4. ———, *Quelques applications probabilistes de la quasi-compacité*, Ann Inst. H. Poincaré **10** (1974), 301–338.

5. T. E. Harris, *The existence of stationary measures for certain Markov processes*, Proc. Third Berkeley Sympos. on Math. Statist. and Probability, 1954/55, vol. II, Univ. of California Press, Berkeley, Calif., 1956, pp. 113–124. MR **18**, 941.

6. N. Jain and B. Jamison, *Contribution to Doeblin's theory of Markov processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **8** (1967), 19–40. MR **36** #4643.

7. M. Métivier, *Théorème limite quotient pour les chaînes de Markov récurrentes au sens de Harris*, Ann. Inst. H. Poincaré Sect. B **8** (1972), 93–105. MR **46** #4613.

8. J. Neveu, *Bases mathématiques du calcul des probabilités*, 2ieme ed., Masson, Paris, 1970. MR **42** #6885.

9. ———, *Sur l'irréductibilité des chaînes de Markov*, Ann. Inst. H. Poincaré **8** (1972), 249–254.

10. ———, *Potentiel markovien recurrent des chaînes de Harris*, Ann. Inst. Fourier (Grenoble) **22** (1972), fasc. 2, 85–130.

UNIVERSITÉ DE PARIS
PARIS, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Functional Equations and Characterization of Probability Distributions

C. Radhakrishna Rao

1. Introduction. The importance of characterization of probability distributions in problems of statistical inference has been stressed in a recent book by Kagan, Linnik and Rao (Russian ed. 1971, English ed. 1973) which will be referred to as KLR in the rest of the paper. There are different types of characterization problems of which the following two seem to have received considerable attention.

One is to assume a model underlying a stochastic phenomenon and derive the appropriate distribution of an observable random variable. Classical examples are the derivation of the normal distribution from Hagan's hypothesis on errors of measurement and from Maxwell's hypothesis on velocities of molecules in a gas (see Rao [**4**, pp. 160–161]).

A second type which opened up a rich area of research is what may be described as characterization of probability distributions through properties of sample statistics. More precisely the problem can be stated as follows:

Let (A, B) and (C, D) be two measure spaces and $T : A \to C$ be a measurable mapping of (A, B) into (C, D). Further let p be a probability measure on (A, B) and p_T the probability measure induced by (statistic) T on (C, D). Further let π be a specified property of p_T. The problem is to find the class

$$P = \{p : p_T \text{ has the specified property } \pi\}. \tag{1.1}$$

The mathematical problem is interesting when π is a weak property and P is a small class. A famous example is the Darmois-Skitovic theorem: Let $X_1, \cdots, X_n$ be independent variables, $T_1 = a_1X_1 + \cdots + a_nX_n$ and $T_2 = b_1X_1 + \cdots + b_nX_n$ be linear functions where a_i, b_i are nonzero, and π be the property that T_1 and T_2 are independently distributed. Then each X_i is normally distributed.

In most cases the problem of characterization reduces to finding the solution of a functional equation. Indeed, the study of characterization problems has led to several new functional equations, not all of which have been satisfactorily solved. In my article, I shall confine myself to the special area of characterizing the multivariate normal distribution (m.n.d.) on R^p, the Euclidean space of p dimensions (with possible generalizations to other spaces), describe the nature of functional equations involved and mention some unsolved problems (listed as Problems 1—7).

2. The Cauchy equation and generalization. Two basic results on functional equations which led to the solution of many characterization problems are given in Lemmas 1 and 2 (see KLR, pp. 29–37, 471–476 for more general results).

LEMMA 1 (KLR, P. 29). *Let $f_1, \cdots, f_k$ be continuous complex-valued functions on R^1 such that, for given a_i all different,*

$$f_1(t + a_1u) + \cdots + f_k(t + a_ku) = A(t) + B(u), \qquad \forall\ t, u \in O^1 \tag{2.1}$$

where O^p denotes a region covering the origin in R^p. Then f_i, A, B are all polynomials of degree $\leqq k$ in O^1.

LEMMA 2 (KLR, P. 471). *Let g_i be a continuous complex-valued function on R^{p_i}, and A_i' and B_i' be matrices of rank p_i and orders $p \times p_i$ and $m \times p_i$ respectively, $i = 1, \cdots, k$, such that $\forall\ \boldsymbol{t} \in O^p$ and $\boldsymbol{u} \in O^m$,*

$$g_1(\boldsymbol{A}_1\boldsymbol{t} + \boldsymbol{B}_1\boldsymbol{u}) + \cdots + g_k(\boldsymbol{A}_k\boldsymbol{t} + \boldsymbol{B}_k\boldsymbol{u}) = C(\boldsymbol{t}) + D(\boldsymbol{u}). \tag{2.2}$$

Then $C(\boldsymbol{t})$ and $D(\boldsymbol{u})$ are polynomials of degree $\leqq k$ in O^p and O^m respectively.

Let us consider a simple case of the equation (2.2) with $k = 1$,

$$g(\boldsymbol{t} + \boldsymbol{u}) = g(\boldsymbol{t}) + g(\boldsymbol{u}) \qquad \forall\ \boldsymbol{t}, \boldsymbol{u} \in O^p \tag{2.3}$$

which is the famous Cauchy equation with a linear function as its solution. Let us restrict the validity of (2.3) to only pairs $\boldsymbol{t}, \boldsymbol{u} \in O^p$ such that the inner product

$$(\boldsymbol{t}, \boldsymbol{u}) = 0. \tag{2.4}$$

What is the solution of (2.3) with the restriction (2.4)? The answer is given in Lemma 3, which is proved by using Lemma 1.

LEMMA 3. *If g satisfies* (2.3) *with the restriction* (2.4), *then g is a polynomial of degree $\leqq 2$ in O^p.*

The solution is no longer linear but is still of the polynomial type.

It may be noted that if the restriction on $(\boldsymbol{t}, \boldsymbol{u})$ is of the type

$$(\boldsymbol{t}, \boldsymbol{u}) = [(\boldsymbol{t}, \boldsymbol{t})(\boldsymbol{u}, \boldsymbol{u})]^{1/2} \cos \alpha,$$

where $\cos \alpha \neq 0$, then again the solution is linear.

It will be of interest to consider other restrictions which may lead to different types of solutions. I mention one such possibility which has applications in characterization problems.

Problem 1. Suppose for any given $\boldsymbol{t} \in O^p$, there exists $\boldsymbol{u}\ (\neq 0) \in O^p$ such that $g(\boldsymbol{at} + \boldsymbol{bu}) = g(\boldsymbol{at}) + g(\boldsymbol{bu})\ \forall\, a, b \in O^1$. What is the solution for g?

Note that no relationship between $\boldsymbol{t}$ and $\boldsymbol{u}$ such as (2.4) is specified. A possible solution for g is

$$g(\boldsymbol{t}) = h(t_1, \cdots, t_i) + Q(t_{i+1}, \cdots, t_p) \tag{2.5}$$

where $t_1, \cdots, t_p$ are components of t in some order, h is an arbitrary function and Q is a quadratic function.

As an application of Lemma 3 we have Theorem 1 characterizing a m.n.d., while a stronger result is true for $p = 2$ as in Theorem 2.

THEOREM 1. *Let $\boldsymbol{X}$ be a p-vector r.v. (random variable) such that $\boldsymbol{a}'\boldsymbol{X}$ and $\boldsymbol{b}'\boldsymbol{X}$ are independently distributed for all $\boldsymbol{a}, \boldsymbol{b} \in R^p$ such that $(\boldsymbol{a}, \boldsymbol{b}) = 0$. Then $\boldsymbol{X}$ has m.n.d.*

THEOREM 2. *Let $\boldsymbol{X}$ be a bivariate r.v. (with components $\in R^1$), $\boldsymbol{A}$ and $\boldsymbol{B}$ be two given nonsingular 2×2 matrices such that $\boldsymbol{A}^{-1}\boldsymbol{B}$ or $\boldsymbol{B}^{-1}\boldsymbol{A}$ has no zero element. If the components of $\boldsymbol{BX}$ are independent and so are also the components of $\boldsymbol{AX}$, then $\boldsymbol{X}$ has a bivariate n.d.*

Theorem 2 shows that to assert bivariate normality of $\boldsymbol{X}$, it is only necessary to find just two pairs of linear functions such that the functions in each pair are independently distributed. For general p, Theorem 1 requires independence for a very wide class of pairs of linear functions. We pose the following problems.

Problem 2. Let $\boldsymbol{X}$ be a p-vector r.v. Suppose that for any given $\boldsymbol{a} \in R^p$, there exists $\boldsymbol{b}\ (\neq 0) \in R^p$ such that $\boldsymbol{a}'\boldsymbol{X}$ and $\boldsymbol{b}'\boldsymbol{X}$ are independently distributed. Then what can be said about $\boldsymbol{X}$?

If (2.5) is the only solution, then some of the components of $\boldsymbol{X}$ have an arbitrary distribution and the rest have a m.n.d.

Problem 3. What is the smallest class of pairs of vectors $(\boldsymbol{a}, \boldsymbol{b})$ such that $\boldsymbol{a}'\boldsymbol{X}$ and $\boldsymbol{b}'\boldsymbol{X}$ are independent, which ensures multivariate normality of $\boldsymbol{X}$? (When $p = 2$, just two pairs are sufficient.)

3. Generalization of Darmois-Skitovic theorem. Let $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_k$ be k independent p-vector variables such that

$$\boldsymbol{A}_1\boldsymbol{X}_1 + \cdots + \boldsymbol{A}_k\boldsymbol{X}_k \quad \text{and} \quad \boldsymbol{B}_1\boldsymbol{X}_1 + \cdots + \boldsymbol{B}_k\boldsymbol{X}_k, \tag{3.1}$$

where $\boldsymbol{A}_i, \boldsymbol{B}_i$ are nonsingular matrices, are independently distributed. Then Ghurye and Olkin [1] showed that each $\boldsymbol{X}_i$ has p-variate n.d. This result is obtained by writing down the functional equation satisfied by g_i, the log of characteristic function (c.f.) of $\boldsymbol{X}_i$,

$$g_1(\boldsymbol{A}_1\boldsymbol{t} + \boldsymbol{B}_1\boldsymbol{u}) + \cdots + g_k(\boldsymbol{A}_k\boldsymbol{t} + \boldsymbol{B}_k\boldsymbol{u}) = C(\boldsymbol{t}) + D(\boldsymbol{u}), \qquad \forall\, \boldsymbol{t}, \boldsymbol{u} \in O^p \tag{3.2}$$

where O^p is a suitable neighbourhood of the origin in which the logs of all functions are well defined, and applying Lemma 2, which shows that $C(\boldsymbol{t})$ is a polynomial which being a c.f. must be of degree $\leqq 2$. Then $\sum \boldsymbol{A}_i\boldsymbol{X}_i$ has m.n.d. and hence $\boldsymbol{A}_i\boldsymbol{X}_i$ and $\boldsymbol{X}_i$ have m.n.d. for each i.

The crucial step in the proof is to show that $C(\boldsymbol{t})$ is a polynomial, which very

much depends on the finite number of terms on the left-hand side of the equation (3.2). It is not clear what happens when the linear forms (3.1) contain an infinite number of terms. We may formulate the problem as follows.

Problem 4. Suppose $\{\boldsymbol{X}_i\}$ is an infinite sequence of independent p-vector r.v.'s such that

$$T_1 = \sum_1^\infty \boldsymbol{A}_i\boldsymbol{X}_i \quad \text{and} \quad T_2 = \sum_1^\infty \boldsymbol{B}_i\boldsymbol{X}_i \tag{3.3}$$

are independently distributed. Then what can be said about the distribution of $\boldsymbol{X}_i$?

A solution to this problem depends on the nature of the solution for $C(\boldsymbol{t})$ in (3.2) when $k = \infty$. When $p = 1$, and the $\boldsymbol{A}_i, \boldsymbol{B}_i \in R^1$ satisfy some conditions it is shown that X_i are normally distributed (KLR, pp. 34, 94). The proof does not easily generalize to $p > 1$.

4. Characterization through constancy of regression. Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ be independent and identically distributed p-vector r.v.'s such that the conditional expectation

$$E(\boldsymbol{X}_1 - \boldsymbol{A}\boldsymbol{X}_2 \,|\, \boldsymbol{X}_1 + \boldsymbol{B}'\boldsymbol{X}_2) = \boldsymbol{0} \tag{4.1}$$

for given nonsingular matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. What can be said about the distribution of $\boldsymbol{X}_1$? We may suppose that $\boldsymbol{X}_1$ has first moment.

A complete solution to the problem is available when $p = 1$ (KLR, pp. 158–161). A complete solution when $p = 2$ and a partial solution for $p > 2$ are given by Khatri and Rao [3]. We shall examine the nature of the functional equation for general p.

Let $g(\boldsymbol{t})$ be the log c.f. of $\boldsymbol{X}_1$ and define by $\boldsymbol{G}(\boldsymbol{t}) = \partial g/\partial \boldsymbol{t}$, the vector of partial derivatives of $g(\boldsymbol{t})$ with respect to the elements of $\boldsymbol{t}$. Then it is easy to show that (4.1) implies

$$G(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{G}(\boldsymbol{B}\boldsymbol{t}) \quad \text{or} \quad \boldsymbol{A}^{-1}\boldsymbol{G}(\boldsymbol{t}) = \boldsymbol{G}(\boldsymbol{B}\boldsymbol{t}). \tag{4.2}$$

The problem is to solve (4.2) for $\boldsymbol{G}$ given $\boldsymbol{A}$ and $\boldsymbol{B}$, and eventually to determine g such that $\boldsymbol{G}(\boldsymbol{t}) = \partial g/\partial \boldsymbol{t}$.

It is interesting to note that an equation of the type (4.2) occurs in the study of optimization problems and structural stability studied by Andronov and Pontrjagin (see Robbins [5]). In their problem $\boldsymbol{A}^{-1}(= \boldsymbol{D}$ say) and $\boldsymbol{B}$ stand for C^1 diffeomorphisms from a smooth manifold M onto itself and $\boldsymbol{G}$ is a homeomorphism such that $\boldsymbol{D} \circ \boldsymbol{G} = \boldsymbol{G} \circ \boldsymbol{B}$ in which case $\boldsymbol{B}$ and $\boldsymbol{D}$ are said to be topologically conjugate. Theorem 3 considers the special case of (4.2) when $\boldsymbol{A} = \boldsymbol{B}^{-1}$.

THEOREM 3. *Let*

$$\boldsymbol{B} = \delta_1\boldsymbol{Q}_1\boldsymbol{P}_1' + \cdots + \delta_r\boldsymbol{Q}_r\boldsymbol{P}_r' \tag{4.3}$$

be the singular value decomposition of $\boldsymbol{B}$, *where* $\boldsymbol{Q}_i$ *and* $\boldsymbol{P}_i$ *are matrices of order* $p \times m_i$ *with orthonormal vectors corresponding to multiplicity* m_i *of the root* δ_i. *If* $\boldsymbol{A} = \boldsymbol{B}^{-1}$ *in* (4.2) *then* $g(\boldsymbol{t})$ *is of the form*

$$g(\boldsymbol{P}\boldsymbol{t}) = h_1(\boldsymbol{t}_1) + \cdots + h_r(\boldsymbol{t}_r), \tag{4.4}$$

$$g(\boldsymbol{Q}\boldsymbol{t}) = \delta_1^2 h_1(\delta_1^{-1}\boldsymbol{t}_1) + \cdots + \delta_r^2 h_r(\delta_r^{-1}\boldsymbol{t}_r), \tag{4.5}$$

where $\boldsymbol{t}_i$ is a subvector of $\boldsymbol{t}$ of order m_i, and h_i are suitable functions. Then h_i in (4.4), (4.5) *satisfy the equation*

$$\sum h_i(\boldsymbol{P}_i'\boldsymbol{Q}_1\boldsymbol{t}_1 + \cdots + \boldsymbol{P}_i'\boldsymbol{Q}_r\boldsymbol{t}_r) = \sum \delta_i^2 h_i(\delta_i^{-1}\boldsymbol{t}_i). \tag{4.6}$$

Thus, the solution of the characterization problem (4.1) even in the special case $\boldsymbol{A} = \boldsymbol{B}^{-1}$ depends on the solution of the functional equation (4.6) which is of the form discussed in KLR (p. 476) but not solved in generality. Solutions have been found for $p = 1, 2$ and for general p when the matrices $\boldsymbol{P}_i'\boldsymbol{Q}_j$ satisfy the conditions given in KLR (p. 476, Theorem A.5.3), leading to multivariate normality of $\boldsymbol{X}_i$ in (4.1). A solution to the equation (4.6) in the general case is of interest.

Problem 5. Let h_i be a continuous complex-valued function on R^{m_i}, $\boldsymbol{P}_i$ and $\boldsymbol{Q}_i$ be partitions with m_i columns of orthogonal matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$, and δ_i be positive numbers for $i = 1, \cdots, r$. What are the solutions for h_i of the functional equation

$$\sum h_i(\boldsymbol{P}_i'\boldsymbol{Q}_1\boldsymbol{t}_1 + \cdots + \boldsymbol{P}_i'\boldsymbol{Q}_r\boldsymbol{t}_r) = \sum \delta_i^2 h_i(\delta_i^{-1}\boldsymbol{t}_i)? \tag{4.7}$$

Now we state a general problem:

Problem 6. What is the solution for $\boldsymbol{G}$ or g ($\boldsymbol{G} = \partial g/\partial \boldsymbol{t}$) of the equation

$$\boldsymbol{A}_1\boldsymbol{G}(\boldsymbol{B}_1\boldsymbol{t}) + \cdots + \boldsymbol{A}_k\boldsymbol{G}(\boldsymbol{B}_k\boldsymbol{t}) = \boldsymbol{0} \tag{4.8}$$

where $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ are given matrices and k may be infinite?

The solution seems to be difficult even for $k = 2$. A solution to the general equation (4.8) would enable us to characterize the probability distribution of $\boldsymbol{X}_i$ by the condition (where $\boldsymbol{X}_i$ are identically distributed)

$$E(\sum \boldsymbol{A}_i\boldsymbol{X}_i \,|\, \sum \boldsymbol{B}_i\boldsymbol{X}_i) = \boldsymbol{0}. \tag{4.9}$$

5. Generalization to other spaces. Throughout this paper, the variables like $\boldsymbol{t}$, $\boldsymbol{u}$ are considered to belong to R^p. All the problems could be generalized to variables belonging to other spaces or topological groups. For instance the equation (3.2), in terms of f_i, the c.f. of $\boldsymbol{X}_i$, is of the form

$$\prod_1^k f_i(\boldsymbol{A}_i\boldsymbol{t} + \boldsymbol{B}_i\boldsymbol{u}) = h(\boldsymbol{t})\cdot m(\boldsymbol{u}). \tag{5.1}$$

One may generalize the problem as follows:

Problem 7. Let $\boldsymbol{t}, \boldsymbol{u} \in X$, a Hausdorff topological group, f_i be conditionally positive definite functions and A_i, B_i be continuous automorphisms of X. Then what are the solutions for f_i of

$$\prod_1^k f_i(\boldsymbol{A}_i\boldsymbol{t}\boldsymbol{B}_i\boldsymbol{u}) = h(\boldsymbol{t})\cdot m(\boldsymbol{u})? \tag{5.2}$$

This problem has been considered by Schmidt [6] and solved when $k = 2$. A solution in the general case (including $k = \infty$) would be of interest.

Similar generalizations can be made of Lemma 3 where $\boldsymbol{t}$, $\boldsymbol{u}$ can belong to a space furnished with an inner product, Problem 1 to a general space, Theorem 2 to a r.v. with components defined on more general spaces instead of R^1 and so on.

References

1. S. G. Ghurye and I. Olkin (1962), *A characterization of the multivariate normal distribution*, Ann. Math. Statist. **33**, 533–541. MR **25** #657.

2. A. M. Kagan, Ju. V. Linnik and C. R. Rao (1973), *Characterization problems in mathematical statistics*, "Nauka", Moscow, 1967; English transl., Appl. Probability and Statist., vol. 1, Wiley, New York, 1973; French transl., Dunod, Paris, 1974. MR **47** #4362.

3. C. G. Khatri and C. R. Rao (1974), *Characterization of multivariate normal distribution by constancy of regression of one linear function on another* (in preparation).

4. C. R. Rao (1973), *Linear statistical inference and its applications*, Wiley, New York.

5. J. W. Robbins (1971), *A structural stability theorem*, Ann. of Math. (2) **94**, 447–493. MR **44** #4783.

6. K. Schmidt (1973), *A class of infinitely divisible positive definite functions on topological groups*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **25**, 97–102.

INDIAN STATISTICAL INSTITUTE
NEW DELHI, INDIA

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Random Time Evolution of Infinite Particle Systems

Frank Spitzer

This is an outline of progress in this field during the last four-year period, which was largely the result of major developments in equilibrium statistical mechanics during the preceding four-year period. A detailed version of this outline will appear in [27].

1. In 1968 ([2], [3]), Dobrušin introduced the notions of an infinite Gibbs state (IGS) and of a Markov random field (MRF), and showed that a MRF is a natural generalization of a stationary Markov process. Somewhat later [(1], [23], [22], [6]), it was realized that every MRF is an IGS with nearest neighbor potential, and vice versa.

2. In 1969, Lanford and Ruelle [15] independently defined IGS and proved a variational characterization analogous to the classical one for finite Gibbs states.

3. The year 1970 saw the first time evolutions which have a given MRF or IGS as equilibrium state. These were on one hand birth and death, or spin flip, evolutions [4], already studied in one dimension by Glauber [5]. On the other hand, time evolutions which preserve the number of particles were proposed in [24].

4. Rigorous existence proofs of such time evolutions as Markovian Feller semigroups T_t, $t \geqq 0$, acting on the continuous function space $C(\Omega)$, where $\Omega = \{0,1\}^{\mathbf{Z}^\nu}$ is the configuration space, were given during 1971 and 1972 by Dobrušin [4], Harris [7], Holley [9], Liggett [16]. Such a semigroup is ergodic if there exists a probability measure ν on Ω such that $\mu T_t \Rightarrow \nu$ for all μ. Then ν is the unique equilibrium state.

5. Dobrušin [4] showed that a birth and death evolution with strictly positive

rates is ergodic if the interaction (dependence of rates on neighboring sites) is weak. The case of zero birth rate when all neighboring sites are vacant was studied by Harris [8], who proved ergodicity when the death rates are sufficiently small, nonergodicity in the contrary case. Holley [10], [11], [12], [13] obtained deep and detailed results when the birth rates are such that the equilibrium states are the equilibrium states for the Ising model. Thus he shows that T_t is nonergodic exactly when the corresponding Ising model exhibits phase transition. The reason one obtains a complete theory in this case was clarified [25], [21] by showing that T_t acts time reversibly in an equilibrium state exactly when the equilibrium state is a MRF (equilibrum state for the Ising model). A surprising irreversible case (the voter model) was recently (1974) completely analyzed by Holley and Liggett [14].

6. The jump processes with constant speed and exclusion have been completely analyzed [17], [26], [18] in the case when the transition function $P(x,y)$ is symmetric. When P is recurrent or a random walk transition function, then the only equilibrium states are the exchangeable measures (convex combinations of Bernoulli product measures). When P is unsymmetric, recent work of Liggett [19], [20] suggests interesting conjectures.

Bibliography

1. M. B. Averintsev (1970), *On a method of describing discrete parameter random fields*, Problemy Peredači Informacii **6**, 100–109. (Russian)

2. R. L. Dobrušin (1968), *Description of a random field by means of conditional probabilities and conditions for its regularity*, Teor. Verojatnost. i Primenen. **13**, 201–229 = Theor. Probability Appl. **13**, 197–224. MR **37** #6989.

3. ———(1968), *Gibbsian random fields for lattice systems with pairwise interactions*, Funkcional. Anal. i Priložen. **2**, no. 4, 31–43 = Functional Anal. Appl. **2**, 292–301. MR **40** #3862.

4. ———(1971), *Markov processes with a large number of locally interacting components*, Problemy Peredači Informacii **7**, 70–87. (Russian)

5. R. Glauber (1963), *The statistics of the stochastic Ising model*, J. Mathematical Phys. **4**, 294–307.

6. R. G. Grimmett (1973), *A theorem about random fields*, Bull. London Math. Soc. **5**, 81–84.

7. T. E. Harris (1972), *Nearest-neighbor Markov interaction processes on multidimensional lattices*, Advances in Math. **9**, 66–89. MR **46** #6512.

8. ——— (1974), *Contact interactions on a lattice* (to appear).

9. R. Holley (1972), *Markovian interaction processes with finite range interactions*, Ann. Math. Statist. **43**, 1961–1967.

10. ——— (1971), *Free energy in a Markovian model of a lattice spin system*, Comm. Math. Phys. **23**, 87–99. MR **45** #1535.

11. ——— (1972), *An ergodic theorem for interacting systems with attractive interactions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **24**, 325–334.

12. ——— (1974), *Some remarks on the* FKG *inequality*, Comm. Math. Phys. **36**, 227–231.

13. ——— (1974), *Recent results on the stochastic Ising model*, Rocky Mountain J. Math. **4**, 479–496.

14. R. Holley and T. Liggett (1974), *Ergodic theorems for weakly interacting infinite systems and the voter model*, Ann. Probability (to appear).

15. O. E. Lanford, and D. Ruelle (1969), *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. **13**, 194–215. MR **41** #1343.

16. T. M. Liggett (1972), *Existence theorems for infinite particle systems*, Trans. Amer. Math. Soc. **165**, 471–481. MR **46** #8328.

17. ——— (1973), *A characterization of the invariant measures for an infinite particle system with interactions*, Trans. Amer. Math. Soc. **179**, 433–453.

18. ——— (1974), *A characterization of the invariant measures for an infinite particle system with interactions*. II, Trans. Amer. Math. Soc. **198**, 207–213.

19. ——— (1974), *Convergence to total occupancy in an infinite particle system with interactions* (to appear).

20. ——— (1974), *Ergodic theorems for the asymmetric simple exclusion process*, Trans. Amer. Math. Soc. (to appear).

21. K. Logan (1974), *Time reversible evolutions in statistical mechanics*, Illinois J. Math. (to appear).

22. C. Preston (1974), *Gibbs states on countable sets*, Cambridge Univ. Press, New York.

23. F. Spitzer (1971), *Markov random fields and Gibbs ensembles*, Amer. Math. Monthly **78**, 142–154. MR **43** #2773.

24. ——— (1970), *Interaction of Markov processes*, Advances in Math. **5**, 246–290. MR **42** #3856.

25. ——— (1974), *École d'été à St. Flour*, Lecture Notes in Math., vol. 390, Springer-Verlag, New York.

26. ——— (1974), *Recurrent random walk of an infinite particle system*, Trans. Amer. Math. Soc. **198**, 191–199.

27. ——— (1975), *Random time evolution of infinite particle systems*, Advances in Math.

CORNELL UNIVERSITY
ITHACA, NEW YORK 14850, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Limit Theorems for Dependent Random Variables Under Various Regularity Conditions

V. Statulevičius

1. Introduction. Let X_t, $t = 0, 1, 2, \cdots$, be a sequence of random variables defined on the probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$ with values in R^k, $k \geqq 1$. In limit theorems of probability theory asymptotic properties of the distribution $P_n(A) = \boldsymbol{P}\{Z_n \in A\}$ of the normalized sum

$$Z_n = B_n^{-1}(S_n - A_n), \qquad S_n = \sum_{j=1}^{n} X_j$$

are investigated most often as $n \to \infty$, where A_n and B_n^{-1} are a nonrandom vector and a matrix, respectively.

The following are implied here:

(a) the determination of approximating distributions $G(A)$ for the distribution $P_n(A)$;

(b) the investigation of the accuracy of the approximation $P_n(A) - G(A)$;

(c) the improvement of this accuracy by adding terms of the asymptotic expansion to $G(A)$;

(d) the investigation of probabilities with large deviations when A is removed together with n and $P_n(A) \to 0$ when $n \to \infty$.

The following classes of sets A are usually considered: a class B of Borel sets, and a class E of convex measurable sets. The distance $\rho(P_n, G)$ between P_n and G is also investigated in various metrics.

Similar problems arise in investigating a still more complicated distribution of a random process $Z_n(t)$, $0 \leqq t \leqq 1$, formed by partial sums $S_0 = 0$, $S_1, \cdots, S_n$ in the configuration of a random polygonal line with vertices at the points $(t_k, B_n^{-1}(S_k - A_k))$, $k = 0, 1, \cdots, n$, $t_0 = 0 \leqq t_1 \leqq \cdots \leqq t_n = 1$, or the distribution of functionals of $Z_n(t)$, for example, $\max_{0<k\leqq n} S_k$ and so on.

It stands to reason that we can speak of general theories in these problems only when the variables $X_1, \cdots, X_n$ are stochastically independent or, in some sense, weakly dependent on one another.

At the present time, in the case of independent summands particularly great attention is paid to the questions (b)—(d). Significant results have been obtained here, especially in the case of finite-dimensional X_j. Due to the absence of good analytical methods it is much more difficult (excluding (a)) to discover those peculiarities which arise in the case of dependent summands, though limit theorems for the distribution of sums of dependent random variables become more and more topical problems in the statistics of random processes, statistical physics, additive number theory and so on.

We present some results for problems of the types (b)—(d) under various conditions of weak dependence.

2. Conditions of weak dependence. Let $\mathscr{F}_s^t = \sigma\{X_u,\ s \leqq u \leqq t\}$ denote the σ-algebra of events generated by the random variables X_u, $0 \leqq s \leqq u \leqq t$. $L(\mathscr{F}_s^t)$ stands for the totality of all $\mathscr{F}_s^t$-measurable random variables with finite variance.

The following conditions of weak dependence are usually considered:

(I) strong mixing (SM):

$$\sup_t \sup_{A \in \mathscr{F}_0^t;\ B \in \mathscr{F}_{t+s}^\infty} \left|P(AB) - P(A)P(B)\right| = \alpha(s) \to 0 \qquad (s \to \infty),$$

(II) full regularity (FR):

$$\sup_t E\left\{ \operatorname*{var}_{B \in \mathscr{F}_{t+s}^\infty} \left|P(B|\mathscr{F}_0^t) - P(B)\right| \right\} = \beta(s) \to 0 \qquad (s \to \infty),$$

(III) uniformly strong mixing (USM):

$$\sup_t \sup_{A \in \mathscr{F}_0^t;\ P(A)>0;\ B \in \mathscr{F}_{t+s}^\infty} \frac{\left|P(AB) - P(A)P(B)\right|}{P(A)} = \varphi(s) \to 0 \qquad (s \to \infty),$$

(IV) Markov type regularity (RMT):

$$\left|P(AB|C) - P(A|C)P(B|C)\right| \leqq \begin{cases} \gamma_1(s) & \text{(RMT I)} \\ P(A|C)\gamma_2(s) & \text{(RMT II)} \end{cases}$$

for all $A \in \mathscr{F}_0^t$, $C \in \mathscr{F}_{t+1}^{t+s-1}$, $B \in \mathscr{F}_{t+s}^\infty$, where $\gamma_i(s) \to 0$ $(s \to \infty)$, $i = 1, 2$;

(V) regularity of correlation functions:

Let $\boldsymbol{E}|X_t|^k < \infty$ for all t and $S_X^{(k)}(t_1, \cdots, t_k) = \Gamma\{X_{t_1}, \cdots, X_{t_k}\}$ be a correlation function of a random process X_t of the kth order, i.e., a simple semi-invariant of a random vector $(X_{t_1}, \cdots, X_{t_k})$. The regularity condition lies in that $S_X^{(k)}(t_1, \cdots, t_k)$ should be sufficiently small when $\max(t_j - t_i) \to \infty$, for instance, in the sense of the existence of integrals $\int \cdots \int S_X^{(k)}(t_1, \cdots, t_k)\, d\Psi(t_1, \cdots, t_k)$ for all $k \geqq 1$.

The first condition was introduced by M. Rosenblatt (1955, [1]), the second by A. N. Kolmogorov, the third by I. A. Ibragimov (1959, [2]), the fourth by B. Rjauba and V. Statulevičius (1962, [3]) and the fifth by V. P. Leonov and A. N. Širjaev (1959, [4]).

Analogues of conditions (I)—(III) in terms of conditional expectations were

considered in the classical work by S. N. Bernstein [5].

In some works (see, for example, [6], [7]) the asymptotic behavior of P_{Z_n} is investigated under elegant conditions on $\boldsymbol{P}\{X_j \in A \,|\, S_{j-1}\}$, $j = 1, 2, \cdots, n$. These conditions, however, are hardly verifiable and they do not separate the properties of dependence from the individual properties of summands, though such a separation would be quite desirable.

Instead of condition (III) one may introduce condition (III'): $\rho(s) \to 0$ $(s \to \infty)$, where $\rho(s)$ is a maximal coefficient of the correlation between the past $\mathscr{F}_0^t$ and the future $\mathscr{F}_{s+t}^{\infty}$ of the process X_t:

$$\rho(s) = \sup \frac{\boldsymbol{E}(X - \boldsymbol{E}X)(Y - \boldsymbol{E}Y)}{\sigma X \sigma Y}$$

in which the supremum is taken over all random variables $X \in L(\mathscr{F}_0^t)$ and $Y \in L(\mathscr{F}_{t+s}^{\infty})$, and t, while σ here denotes standard deviation. The maximal correlation coefficient was considered in [8]—[10]. We always have $\rho(s) \leqq 2\varphi^{1/2}(s)$. Similarly it is possible to define the maximal kth order correlation coefficient

$$\rho_k(s) = \sup \frac{\left|\hat{\boldsymbol{E}}\xi_1\xi_2 \cdots \xi_k\right|}{\prod_{j=1}^{k} \boldsymbol{E}^{1/k}\left|(\xi_j - \boldsymbol{E}\xi_j)^k\right|},$$

where the supremum is taken over all $\xi_1 \in L(\mathscr{F}_0^t)$, $\xi_j \in L(\mathscr{F}_{u_j}^{v_j})$, $j = 2, \cdots, k-1$, $\xi_k \in \mathscr{F}_{t+s}^{\infty}$ with $\boldsymbol{E}|\xi_j|^k < \infty$ and over all t and $u_j \leqq v_j$, where $0 \leqq t < u_2 \leqq v_2 < \cdots < u_{k-1} \leqq v_{k-1} < t + s$. The symbol $\hat{\boldsymbol{E}}$ is defined as follows: The sign "$\frown$" over Y denotes the centering $\hat{Y} = Y - EY$; then

$$\hat{\boldsymbol{E}}Y_1 \cdots Y_r = \boldsymbol{E}Y_1\overbrace{Y_2 \cdots Y_{r-1}\hat{Y}_r}.$$

It is evident that $\rho_2(s) = \rho(s)$. Theorems of large deviations for $\boldsymbol{P}\{Z_n > x\}$ under the condition $\rho_k(s) \leqq k!\, L^k \exp\{-\beta_n \cdot s\}$ with all $k \geqq 2$ were proved in [11] and [12].

If the random variables $Y_1, \cdots, Y_r$ are related to a Markov chain $x(t)$, $t = 1, 2, \cdots, r$, $Y_t = f(x(t))$ with transition probability $P_t(x, A)$ and initial distribution $P_1(A)$, then

$$\hat{\boldsymbol{E}}\, Y_1^{\nu_1} \cdots Y_r^{\nu_r} = \int \cdots \int f^{\nu_1}(x_1) P(dx_1) \prod_{j=2}^{r} f^{\nu_j}(x_j)\,(P_j(x_{j-1}, dx_j) - P_j(dx_j))$$

where $P_t(A) = \boldsymbol{P}\{x(t) \in A\}$.

Let us consider the kth order correlation among the indicators I_{A_j} of sets $A_j \in \mathscr{F}_{u_j}^{v_j}$, $j = 2, \cdots, k-1$, $A_1 \in \mathscr{F}_0^t$, $A_k \in \mathscr{F}_{t+s}^{\infty}$. Assume that $\alpha(A_1, \cdots, A_k) = \boldsymbol{E}I_{A_1} \cdots I_{A_k}$. Let us say that the condition (α_k) is satisfied if

$$\sup_{t,\, u_j \leqq v_j,\, A_j} \left|\alpha(A_1, \cdots, A_k)\right| = \alpha_k(s) \to 0 \qquad (s \to \infty)$$

as well as the condition (φ_k) if

$$\sup_{t,\, u_j \leqq v_j,\, A_j:\, P(A_1) > 0} \frac{\left|\alpha(A_1, \cdots, A_k)\right|}{P(A_1)} = \varphi_k(s) \to 0 \qquad (s \to \infty).$$

There exists a positive constant c_k, depending only on k, such that

$$\rho_k(s) \leqq c_k \varphi_k^{1/k}(s).$$

The properties of $\alpha(A_1, \cdots, A_k)$ were studied by Kolmogorov and Žurbenko [13]. The conditions (α_k) and (φ_k) were applied in [35] in investigating the accuracy of the approximation of $P_n(A)$ by a normal distribution.

Note that (α_2) coincides with condition (I) and (φ_2) with (III). When k increases the conditions (α_k) and (φ_k) become more and more restrictive, and transform into RMT I and RMT II, respectively.

We shall not linger on the conditions for convergence of the distributions P_{Z_n} and $P_{Z_n(\cdot)}$ to limit ones. We shall only direct the reader to the papers [1] — [41], where one can also find a rather complete bibliography. We shall only illustrate an interesting result by Ibragimov [14]. Assume that X_t is a strictly stationary sequence with $\boldsymbol{E}X_t \equiv 0$ and $\rho(s) \to 0$ as $s \to \infty$. Then either $\sup_n \sigma^2 S_n < \infty$ or $\sigma^2 S_n = n \cdot h(n)$, where $h(n)$ is a slowly varying Karamata function. If in addition $\boldsymbol{E}|X_t|^{2+\delta} < \infty$ for some $\delta > 0$ and $\sigma^2 S_n \to \infty$ $(n \to \infty)$, then

$$\lim_{n\to\infty} \boldsymbol{P}\{S_n/\sigma S_n < x\} = \Phi(x),$$

$$\Phi(x) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{x} e^{(-y^2/2)}\, dy.$$

3. Large deviations.

THEOREM 1. *If the condition* RMT II *with*

$$\gamma_2(s) \leqq e^{-\gamma_n \cdot s}, \qquad \gamma_n > 0, \tag{1}$$

is satisfied, $X_j \in R$, $\boldsymbol{E}X_j = 0$, *and*

$$\boldsymbol{E}\{|X_j^p| \mid \mathscr{F}_0^{j-1}\} \leqq p! L K^{p-2} \sigma^2 X_j, \qquad j = 1, \cdots, n,$$

with probability 1 *for all* $p \geqq 3$, *then for* $Z_n = S_n/B_n$, $B_n^2 = \sigma^2 S_n$, *the relation of large deviations*

$$\frac{\boldsymbol{P}\{Z_n > x\}}{1 - \Phi(x)} = \exp\{(x^3/\Delta_n)\lambda_n(x/\Delta_n)\}\left(1 + \theta_{H,\delta}\frac{x+1}{\Delta_n}\right) \tag{2}$$

holds when $0 \leqq x \leqq \delta \cdot \Delta_n$, $\delta < \delta_H$, *where* $\theta_{H,\delta}$ *is bounded by a constant depending only on* H *and* δ,

$$\lambda_n(t) = \sum_{k=0}^{\infty} \lambda_{kn} t^k, \quad |\lambda_{kn}| \leqq \frac{1}{(k+3)\delta_H^{k+2}}, \quad k = 0, 1, \cdots.$$

Here

$$\Delta_n = \frac{c_{KL} \cdot \gamma_n \cdot B_n}{\max_{1\leqq k\leqq u} \sigma X_k}, \qquad H = C_{K,L} \frac{\sum_{j=1}^{n} \sigma^2 X_j}{B_n^2}.$$

The positive constants $c_{K,L}$ and $C_{K,L}$ depend only on K and L, and $\delta_H > 0$ is a maximal root of the equation $6H\delta_H/(1 - \delta_H)^3 = 1$ (estimation for $\theta_{\delta,H}$ can be found in [15]).

REMARK 1. If $|X_j| \leqq C^{(n)}, j = 1, \cdots, n$, with probability 1, then in this case

(2) holds when $\Delta_n = (\gamma_n \cdot B_n)/(H_2 \cdot C^{(n)})$, $H = 2H_1$, where $H_1 > 0$ and $H_2 > 0$ are absolute constants included by the estimate for the semi-invariant $\Gamma_k\{S_n\}$ of the kth order of the sum S_n:

$$|\Gamma_k\{S_n\}| \leqq \frac{k! H_1 H_2^{k-2} C^{(n)k-2} B_n^2}{\gamma_n^{k-2}}, \qquad k = 3, 4, \cdots. \tag{3}$$

REMARK 2. If X_t, $t \in (-\infty, \infty)$, is a strictly stationary process, $\boldsymbol{E}|X_t|^k < \infty$ for all $k \geqq 1$ and

$$\int_0^T \cdots \int_0^T |s_X(t_1, \cdots, t_{k-1}, 0)|\, dt_1 \cdots dt_{k-1} \leqq k! H_3 H_4^{k-2}$$

for all $k \geqq 3$, then for $\boldsymbol{P}\{Z_T > x\}$, where $Z_T = (\zeta_T - \boldsymbol{E}\zeta_T)/\sigma\zeta_T$, $\zeta_T = \int_0^T X_t\, dt$, the relation of large deviations (2) is valid when $\Delta_T = cT^{1/2}/H_4$, $H = H_3/c^2$, if $\sigma^2\zeta_T \geqq cT$ (see [11]).

The method of proving such theorems is as follows. The kth order semi-invariant of the sum

$$\Gamma_k\{S_n\} = \sum_{1 \leqq t_1, \cdots, t_k \leqq n} \Gamma\{X_{t_1}, \cdots, X_{t_k}\}$$

may be exactly expressed in terms of $\hat{\boldsymbol{E}} X_{t_{j_1}} \cdots X_{t_{j_m}}$, $m \leqq k$. When $|X_j| \leqq C^{(n)}$, $j = 1, \cdots, n$, and (1) is true the estimation

$$|\hat{\boldsymbol{E}} X_{t_{j_1}} \cdots X_{t_{j_m}}| \leqq C \cdot C^{(n)m-2} \exp\{-\gamma_n \cdot (t_{j_m} - t_{j_1})\} \sigma X_{t_{j_1}} \cdot \sigma X_{t_{j_m}}$$

holds if $t_{j_1} \leqq \cdots \leqq t_{j_m}$, where C is an absolute constant. Hence we get estimates of the type (3) for $\Gamma_k\{S_n\}$. Further, if $|\Gamma_k\{Z\}| \leqq k!\, H/\Delta^{k-2}$ for all $k \geqq 3$, $\boldsymbol{E}Z = 0$, $\sigma^2 Z = 1$, then (2) is valid for $\boldsymbol{P}\{Z > x\}$ with the parameters H and Δ (see [15]). Estimates for $\Gamma\{X_{t_1}, \cdots, X_{t_k}\}$ under the RMT II conditions were obtained by I. G. Žurbenko [27], [28] as well.

4. Rate of convergence and asymptotic expansion.

THEOREM 2. *If* $X_j \in R^k$, $\boldsymbol{E}X_j = 0, j = 1, \cdots, n$, *and* RMT II *with* $\gamma_2(s) \leqq 1/(\gamma_n^a \cdot s^a)$ *is satisfied for all* $1 \leqq s \leqq n$ *and some* $a > 3, \gamma_n > 0$, $\boldsymbol{E}|X_j|^3 < \infty, j = 1, \cdots, n$, *then*

$$\sup_E |\boldsymbol{P}\{S_n \in A\} - \Phi_{s_n}(A)| \leqq C(a, k) \sup_{|u|=1;\, u \in R^k} \frac{\sum_{i=1}^n \operatorname{ess\,sup} \boldsymbol{E}\{|(X_j, u)|^3 | \mathscr{F}_1^{i-1}\}}{\gamma_n^2 (\sigma^2\{S_n, u\})^{3/2}}. \tag{4}$$

Here (x, y) denotes the scalar product, and $C(a, k)$ depends only on k and a.

THEOREM 3. *If* $X_j \in R$, $\boldsymbol{E}X_j = 0$, $\boldsymbol{E}|X_j|^3 < \infty, j = 1, \cdots, n$, *and* RMT II *with* $\gamma_2(s) \leqq e^{-\gamma_n \cdot s}$ *is satisfied, then there exists an absolute constant* C_1 *such that*

$$\sup_x \left| \boldsymbol{P}\left\{\frac{\max_{1 \leqq k \leqq n} S_k}{B_n} < x\right\} - \left(\frac{2}{\pi}\right)^{1/2} \int_0^x e^{-y^2/2}\, dy \right| \leqq C_1 \frac{1}{\gamma_n B_n} \frac{\max_{1 \leqq j \leqq n} \operatorname{ess\,sup} \boldsymbol{E}\{|X_j|^3 | \mathscr{F}_1^{j-1}\}}{\min_{1 \leqq j \leqq n} \operatorname{ess\,inf} \sigma^2\{|X_j| \mathscr{F}_1^{j-1}\}}, \qquad B_n^2 = \sigma^2 S_n.$$

In fact condition RMT II is necessary in order to obtain the theorems of large deviations such as Theorem 1 (there is an example indicating that relation (2) is not

valid without RMT). To obtain (4), however, it is not required. Further, let $X_j \in R$, $j = 1, \cdots, n$, and

$$L_{rn} = \frac{\sum_{j=1}^{n} \operatorname{ess\,sup} \boldsymbol{E}\{|X_j|^r | \mathscr{F}_0^{j-1}\}}{\gamma_n^{r-1} B_n^r}, \qquad B_n^2 = \sigma^2 S_n.$$

THEOREM 4. *If the condition* (φ_3) *with*

$$\varphi_3(s) \leqq \frac{1}{\gamma_n^a \cdot s^a}, \qquad 1 \leqq s \leqq n, \tag{5}$$

is satisfied for some $a > 3$, $\boldsymbol{E}|X_j|^3 < \infty$, $j = 1, \cdots, n$, *then there exists a constant* C_a *depending only on a such that*

$$\sup_x |\boldsymbol{P}\{Z_n < x\} - \Phi(x)| \leqq C_a \cdot L_{3n}.$$

THEOREM 5. *If* $|X_j| \leqq C^{(n)}$, $j = 1, \cdots, n$, *with probability* 1 *and the condition* (α_3) *with*

$$\alpha_3(s) \leqq 1/\gamma_n^a \cdot s^a, \qquad 1 \leqq s \leqq n, \tag{6}$$

is satisfied for some $a > 3$ *then there exists a constant* C_a' *such that*

$$\sup_x |\boldsymbol{P}\{Z_n < x\} - \Phi(x)| \leqq C_a' \frac{C^{(n)}}{\alpha_n \cdot B_n \cdot (1 + |x|^3)}.$$

If we want to get an asymptotic expansion for $\boldsymbol{P}\{Z_n < x\} - \Phi(x)$ with the help of Čebyšev-Hermite polynomials

$$\boldsymbol{P}\{Z_n < x\} = \Phi(x) + \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} \sum_{\nu=1}^{r-3} P_{\nu n}(x) + \theta_r L_{rn}$$

with as simple a structure of the remainder term as that of L_{rn}, the condition (φ_r) will be needed here if $\boldsymbol{E}|X_j|^r < \infty$ and (α_r) when $|X_j| \leqq C^{(n)}$, $j = 1, \cdots, n$ (see [35]). Note that in Theorems 2–5 the conditions for $\gamma_2(s)$, $\varphi_r(s)$, $\alpha_r(s)$ are not best possible; they can be weakened. And in general in asymptotic expansions one can get estimates of the order $(\log n)/(n^{1/2})^{r-2}$ by imposing stronger restrictions on $\varphi_m(s)$, $m < r$, but it will be very difficult to describe the structure of the remainder term exactly. For instance, if $\varphi(s) \equiv \varphi_2(s) \leqq C_3 e^{-b \cdot s}$, $\boldsymbol{E}|X_j|^3 \leqq C_4$, $j = 1, \cdots, n$, then it is possible to obtain the estimate

$$\boldsymbol{P}\{Z_n < x\} = \Phi(x) + O((\log n)/n^{1/2})$$

when $\sigma^2\{S_l - S_k\} \asymp l - k$, $1 \leqq k < l \leqq n$. The rate of convergence under the condition (III) was dealt with in [22], [32], [29].

In the remainder terms of Theorems 2–5 the conditional moments $\boldsymbol{E}\{|X_j|^3 | \mathscr{F}_1^{j-1}\}$ can be replaced by absolute ones $\boldsymbol{E}|X_i|^3$ by multiplying the remainder terms by $\log n$. The method of proof rests on the accurate investigation of logarithmic derivatives

$$\Gamma_k\{S_n, t\} = \frac{d^k}{dt} \log f_{S_n}(t)$$

of the characteristic function of the sum S_n when $|t| \leqq L_{3n}^{-1/(r-2)} B_n^{-1}$.

We shall present here one more theorem which employs the regularity conditions (V).

Let X_t be a strictly stationary random sequence with $\boldsymbol{E}X_t \equiv 0$ and $\boldsymbol{E}|X_t^p| < \infty$ for all $p \geqq 1$. Let $F(\lambda) = \int_0^\lambda f(\lambda)\, d\lambda$ be a spectral function of the process X_t with bounded density $\sup_\lambda f(\lambda) \leqq L$. Let

$$\hat{F}_T(\lambda) = \frac{1}{2\pi T} \int_0^T \left| \sum_{t=1}^{T} e^{-ist} X_t \right|^2 ds$$

be the estimation for $F(\lambda)$.

THEOREM 6. *If*

$$\sum_{k-1} \cdots \sum \left| S_X^{(k)}(t_1, \cdots, t_{k-1}, 0) \right| \leqq k!\, H_5 H_6^{k-2},$$

then

$$\sup_x \left| \boldsymbol{P}\left\{ T^{1/2} \sup_{0 \leqq \lambda \leqq \pi} \left| \hat{F}_T(\lambda) - F(\lambda) \right| \leqq \sigma_T \cdot x \right\} - \boldsymbol{P}\left\{ \sup_{0 \leqq t \leqq 1} |w(t)| \leqq x \right\} \right| \leqq C_{L,H,H_6} \frac{\log(1+T)}{\sigma_T T^{1/2}}$$

where $w(t)$ is a Wiener process, $\sigma_T^2 = 2\pi \int_0^\pi \lambda^2 f(\lambda) d\lambda$.

Estimations of such a type for the Gaussian sequence were obtained by T. Arak [37].

Bibliography

1. M. Rosenblatt, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci. U.S.A. **42** (1956), 43–47. MR **17**, 635.

2. I. A. Ibragimov, *Some limit theorems for stochastic processes stationary in the strict sense*, Dokl. Akad. Nauk SSSR **125** (1959), 711–714; English transl., Selected Transl. in Math. Statist. and Probability, vol. 4, Amer. Math. Soc., Providence, R. I., 1963, pp. 99–103. MR **21** #3923.

3. B. Rjauba, *The central limit theorem for sums of series of weakly dependent random variables*, Litovsk. Mat. Sb. **2** (1962), no. 2, 193–205. (Russian) MR **29** #6526.

4. V. P. Leonov and A. N. Širjaev, *Sur le calcul des semi-invariants*, Teor.-Verojatnost. i Primenen. **4** (1959), 342–355 = Theor. Probability Appl. **4** (1959), 319–329. MR **22** #5058.

5. S. Bernstein, *Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendents*, Math. Ann. **97** (1927), 1–59.

6. I. I. Gihman, *Limit theorems for sequences of series of random variables*, Teorija Slučaĭnyh Processov **2** (1974), 37–47. (Russian)

7. A. Dvoretzky, *Asymptotic normality for sums of dependent random variables*, Proc. Sixth Berkeley Sympos. on Math. Statist. and Probability, Univ. of California Press, Berkeley, Calif., 1973, pp. 513–535.

8. H. Gebelein, *Das statistische Problem der Korrelation als Variations-und Eigenwertproblem und sein Zusammenhang mit der Ausgleichungsrechnung*, Z. Angew. Math. Mech. **21** (1941), 364–379. MR **4**, 104.

9. O. V. Sarmanov, *The maximum correlation coefficient (symmetric case)*, Dokl. Akad. Nauk SSSR **120** (1959), 715–718; English transl., Selected Transl. in Math. Statist. and Probability, vol. 4, Amer. Math. Soc., Providence, R. I., 1963, pp. 271–275. MR **21** #2334.

10. A. Rényi, *On measures of dependence*, Acta Math. Acad. Sci. Hungar. **10** (1959), 441–451. MR **22** #6005.

11. V. Statulevičius, *Limit theorems for random functions.* I, Litovsk. Mat. Sb. **10** (1970), 583–592. (Russian) MR **43** #2765.

12. ———, *On limit theorems for random processes*, Second Japan-USSR Sympos. on Probability Theory, Kyoto, 1972, pp. 138–152.

13. I. Žurbenko, *On mixing conditions for random processes*, Slučaïnye Processov **3** (1973), 57–62. (Russian)

14. I. Ibragimov, *Remark on the central limit theorem for dependent random variables*, Teor. Verojatnost. i Primenen. (to appear).

15. V. A. Statulevičius, *On large deviations*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **6** (1966), 133–144. MR **36** #4612.

16. I. A. Ibragimov, *Some limit theorems for stationary processes*, Teor. Verojatnost. i Primenen. **7** (1962), 361–392 = Theor. Probability Appl. **7** (1962), 349–381. MR **26** #5634.

17. V. A. Voklonskiĭ and Ju. A. Rozanov, *Some limit theorems for random functions.* I, Teor. Verojatnost. i Primenen. **4** (1959), 186–207 = Theor. Probability Appl. **4** (1959), 178–197. MR **21** #4477; **22** #12586.

18. V. A. Statulevičius, *Limit theorems for sums of random variables that are connected in a Markov chain.* I, II, III, Litovsk. Mat. Sb. **9** (1969), 345–362, 635–672; ibid. **10** (1970), 161–169. (Russian) MR **42** #1188.

19. Ju. A. Davydov, *Mixing conditions for Markov chains*, Teor. Verojatnost. i Primenen. **18** (1973), 321–338. (Russian)

20. M. Rosenblatt, *Central limit theorem for stationary processes*, Proc. Sixth Berkeley Sympos. on Math. Statist. and Probability, Univ. of California Press, Berkeley, Calif., 1973, pp. 551–561.

21. R. Cogburn, *The central limit theorem for Markov processes*, Proc. Sixth Berkeley Sympos. on Math. Statist. and Probability, Univ. of California Press, Berkeley, Calif., 1973, pp. 485–512.

22. Ch. Stein, *A bound for error in the normal approximation to the distribution of a sum of dependent random variables*, Proc. Sixth Berkeley Sympos. on Math. Statist. and Probability, Univ. of California Press, Berkeley, Calif., 1973, pp. 583–602.

23. W. Philipp, *The central limit problem for mixing sequences of random variables*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **12** (1969), 155–171. MR **39** #7660.

24. R. J. Serfling, *Contributions to central limit theory for dependent variables*, Ann. Math. Statist. **39** (1968), 1158–1175. MR **37** #3637.

25. H. Cohn, *On a class of dependent random variables*, Rev. Roumaine Math. Pures Appl. **10** (1965), 1593–1606. MR **34** #5136.

26. Ja. G. Sinaĭ, *On limit theorems for stationary processes*, Teor. Verojatnost. i Primenen. **7** (1962), 213–219. (Russian) MR **24** #A2990.

27. I. G. Žurbenko, *Strong estimates for mixed semi-invariants of random processes*, Sibirsk. Mat. Ž. **13** (1971), 293–308 = Siberian Math. J. **13** (1971), 202–213. MR **45** #6141.

28. ———, *Spectral semi-invariants of stationary processes with sufficiently strong mixing*, Teor. Verojatnost. i Primenen. **17** (1972), 150–152 = Theor. Probability Appl. **17** (1972), 151–152. MR **45** #7806.

29. V. V. Petrov, *On the central limit theorem for m-dependent variables*, Proc. All-Union Conf. Theory Prob. and Math. Statist. (Erevan, 1958), Izdat. Akad. Nauk Armjan. SSR, Erevan, 1960, pp. 38–44; English transl., Selected Transl. Math. Statist. and Probability, vol. 9, Amer. Math. Soc., Providence, R. I., 1971, pp. 83–88. MR **34** #848.

30. P. H. Diananda, *The central limit theorem for m-dependent variables*, Proc. Cambridge Philos. Soc. **51** (1955), 92–95. MR **16**, 724.

31. W. Hoeffding and H. Robbins, *The central limit theorem for dependent random variables*, Duke Math. J. **15** (1948), 773–780. MR **10**, 200.

32. W. Philipp, *The remainder in the central limit theorem for mixing stochastic processes*, Ann. Math. Statist. **40** (1969), 601–609. MR **39** #4912.

33. H. Bergström, *Reduction of the limit problem for sums of random variables under a mixing condition*, Proc. Fourth Conference on Probability Theory, 1971, pp. 107–120.

34. A. Gut, *On the moments of some first passage times for sums of dependent random variables*, Stochastic Processes Appl. **2** (1974), 115–126.

35. V. Statulevičius, *Asymptotic analysis of the distribution of sums of dependent random variables and the regularity conditions*, Abstracts of Comm., Internat. Conference on Probability and Statist., Vilnius, 1973. pp. 249–252.

36. V. P. Leonov, *Some applications of higher semi-invariants to the theory of stationary random processes*, "Nauka", Moscow, 1964. (Russian) MR **30** #5371.

37. T. Arak, *On the rate of convergence of the distribution of the maximum deviation of an estimate of the spectrum of a Gaussian random sequence*, Dokl. Akad. Nauk SSSR **201** (1971), 1019–1021 = Soviet Math. Dokl. **12** (1971), 1747–1750. MR **45** #2787.

38. A. N. Kolmogorov and Ju. A. Rozanov, *On a strong mixing condition for a stationary random Gaussian process*, Teor. Verojatnost. i Primenen. **5** (1960), 222–227 = Theor. Probability Appl. **5** (1960), 204–208. MR **24** #A3009.

39. R. L. Dobrušin, *Central limit theorem for non-stationary Markov chains.* I, II, Teor. Verojatnost. i Primenen. **1** (1956), 72–89, 365–425 = Theor. Probability Appl. **1** (1956), 65–80, 329–383. MR **19**, 184; **20** #3592.

40. B. Rosen, *On the central limit theorem for sums of dependent random variables*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **7** (1967), 48–82. MR **34** #8465.

41. T. Hoglund, *Central limit theorems and statistical inference for finite Markov chains*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **29** (1974), 123–151.

INSTITUTE OF PHYSICS AND MATHEMATICS, VILNIUS STATE UNIVERSITY
VILNIUS, U.S.S.R.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Theory of Harmonic Spaces

Bertram Walsh

Let $(X, \mathscr{H})$ be a harmonic space in the sense of H. Bauer's address. This report is devoted to exposition of some things one can say, and questions one can ask, about X and $\mathscr{H}$ from the standpoints of topology, functional analysis, and sheaf theory. While it will seem that potential theory does not appear explicitly and probabilistic considerations are absent, one should keep in mind W. Hansen's work [13] showing that the operators we shall examine are either semigroup generators or potential kernels.

J.-M. Bony [6], [7] has shown that if X and $\mathscr{H}$ are sufficiently differentiable, then on an open dense set $\mathscr{H}$ is the solution sheaf of a second-order equation $Lu = 0$, where $Lu = f$ is (locally) solvable for a large class of functions f and solutions are superharmonic if $f \geqq 0$. The situation in $(X, \mathscr{H})$ is thus the same as the classical one in open sets of $\boldsymbol{R}^n$ where $L = \Delta$ and u, f are $\mathscr{C}^\infty$ functions, i.e. (in the usual sheaf-theoretic formulation), $0 \to \mathscr{H} \to \mathscr{E} \xrightarrow{\Delta} \mathscr{E} \to 0$ is an exact sequence of sheaves ($\mathscr{E}$ being the sheaf of [germs of] $\mathscr{C}^\infty$ real-valued functions). Since it is a "fine resolution" of $\mathscr{H}$, this sequence makes it possible in the classical situation to determine the sheaf cohomology groups with coefficients in $\mathscr{H}$. Since $\mathscr{E}$ is a sheaf of modules over the $\mathscr{C}^\infty$ functions it is also possible, for each $\mathscr{C}^\infty$ function m, to define an operator $\theta_m: u \to \Delta u + m \cdot u$. It is classical that the sequence $0 \to \mathscr{G}_m \to \mathscr{E} \xrightarrow{\theta_m} \mathscr{E} \to 0$ is also exact and that $\mathscr{G}_m = \operatorname{Ker} \theta_m$ satisfies the Brelot axioms. Harmonic structures "near" the classical one thus exist.

Assuming only the usual axioms for $(X, \mathscr{H})$ (and the weakest convergence axiom), one can construct ([28], [29], [30]) fine sheaves $\mathscr{R}$ and $\mathscr{Q}$ and a homomorphism Δ such that $0 \to \mathscr{H} \to \mathscr{R} \xrightarrow{\Delta} \mathscr{Q} \to 0$ is exact; moreover, $\mathscr{Q}$ is a sheaf of modules over the sheaf $\mathscr{B}$ of germs of bounded Borel functions on X, so it makes sense to define operators $\theta_M: u \to \Delta u + u \cdot M$ over U for any $M \in \Gamma(U, \mathscr{Q})$. Any $\mathscr{G}_M = \operatorname{Ker} \theta_M$

is a sheaf of continuous functions for which $(U, \mathscr{G}_M)$ is a harmonic space [30]. $\mathscr{R}$ is the sheaf of germs of continuous functions that are local differences of continuous $\mathscr{H}$-superharmonic functions; if $1 \in \mathscr{H}$ this is a sheaf of algebras possessing partitions of unity, so in all cases it is locally fine and hence fine. To define $\mathscr{Q}$, for each open $U \subseteq X$ let $\mathfrak{P}_U$ be its cone of continuous $\mathscr{H}$-potentials, and if $V \subseteq U$ let $r_{VU} : \mathfrak{P}_U \to \mathfrak{P}_V$ be defined by $r_{VU}p = q$, where $q \in \mathfrak{P}_V$ is the potential part of the Riesz decomposition $p|V = q + h$ on V. The r_{VU}'s extend linearly to the spaces $\mathfrak{D}_U = \mathfrak{P}_U - \mathfrak{P}_U$, and $\mathscr{Q}$ is the sheaf associated with the presheaf $\{\mathfrak{D}_U, r_{VU}\}$. The operator $\Delta : \mathscr{R} \to \mathscr{Q}$ is defined similarly at the presheaf level: If s is a continuous superharmonic function on an open set U it determines a continuous potential p in any open V with compact closure contained in U, namely the potential part p of the Riesz decomposition $s|V = p + h$ on V; one verifies that the map $s \to p$ determines a linear homomorphism $\Delta : \mathscr{R} \to \mathscr{Q}$ with $\mathscr{H} = \operatorname{Ker} \Delta$. $\mathscr{Q}$ is fine as a consequence of the theory of specific restriction of potentials [15] or specific multiplication [11], [27], according to which there is an action of the bounded Borel functions on open $U \subseteq X$ upon the potentials on U that has all the properties of the classical action given by $f \cdot p = \int G_U(\cdot, y) f(y)\, d\mu(y)$ where $U \subseteq \boldsymbol{R}^n$ is open, G_U is its Green's function for the Laplace equation, and μ is the unique measure for which $p = \int G_U(\cdot, y)\, d\mu(y)$. In any harmonic space $r_{VU}(f \cdot p) = (f|V) \cdot r_{VU}p$ holds, so $\mathscr{Q}$ is a sheaf of modules over $\mathscr{B}$ and is therefore fine.

For each open $U \subseteq X$ the space $\Gamma(U, \mathscr{Q})$ can be made a Fréchet space, and indeed a Banach space if U is compact. $\mathscr{Q}$ is the analogue of the sheaf $\mathscr{L}$ of (germs of) measures defined as follows in the classical setting: $\mu \in \mathscr{M}(U)$ belongs to $\mathscr{L}$ if $\int_K G_U(\cdot, y)\, d|\mu|(y)$ is continuous for every $U \subseteq \boldsymbol{R}^n$ possessing a Green's function and every compact $K \subseteq U$. If $(X, \mathscr{H})$ has local Green's functions [11], [15], [16], then the family $\mathscr{L}$ of measures formally defined in the same way is a sheaf isomorphic to $\mathscr{Q}$ as a sheaf of $\mathscr{B}$-modules, and Δ becomes an operator sending functions to measures. If the adjoint sheaf [15] exists locally, then [29] it too has a resolution $0 \to \mathscr{H}^* \to \mathscr{R}^* \to \mathscr{L}^* \to 0$ and there is a duality relation (I) $\Gamma(U, \mathscr{H}^*) \cong H^1_K(U, \mathscr{H})'$, and a similar one with $\mathscr{H}$ and $\mathscr{H}^*$ interchanged [29]; natural locally convex topologies are present. If X is compact, then the convergence principle implies that $\Gamma(X, \mathscr{H}^*)$ is finite-dimensional and thus so is $H^1(X, \mathscr{H})$; if $1 \in \overline{\mathscr{H}}^*$ the dimension is 0 or 1.

Even without duality the short fine resolution implies $H^q(U, \mathscr{H}) = 0$ for $q \geqq 2$. If X is a compact Brelot harmonic space with $1 \in \overline{\mathscr{H}}$, or if a consistent family of ways of solving the exterior Dirichlet problem for a basis of neighborhoods of ∞ in the one-point compactification $\hat{X}$ of X is given (normal operators [28] generalizing the notions of L. Sario [22] or a full-harmonic structure like those of Z. Kuramochi [18], with corresponding sheaf of "full-harmonic" functions $\mathscr{H}^{\mathfrak{Q}}$ over X), then it is possible to determine [28] the Čech $H^1(X, \mathscr{H})$ (or $H^1(\hat{X}, \mathscr{H}^{\mathfrak{Q}})$, mutatis mutandis) by solving Cousin problems. This is most perspicaciously done using a theorem of H. Schaefer [23] about positive operators on spaces $\mathscr{C}(Y)$.

(II) If $1 \notin \mathscr{H}_X$ then $H^1(X, \mathscr{H}) = 0$, while if $1 \in \mathscr{H}_X$ then $\dim H^1(X, \mathscr{H}) = 1$; the dual of $H^1(X, \mathscr{H})$ has a "positive" generator identifiable with the classical flux.

(III) One can determine $H^1(X, \mathscr{H})$ for general compact X by perturbing the sheaf $\mathscr{H}$ [**30**]; one defines $\theta_M: \mathscr{R} \to \mathscr{Q}$ and $\mathscr{G}_M = \operatorname{Ker} \theta_M$ as above and $\mathscr{G}_M$ is a harmonic structure, though one can only preserve a weak convergence axiom. If M is sufficiently "large" in a natural ordering of $\mathscr{Q}$, then $\mathscr{G}_M$ will possess a strong potential. With $U = X$, it follows from this fact and the R.-M. Hervé extension theorem [**15**] that $\Gamma(X, \mathscr{G}_M) = 0$ and $H^1(X, \mathscr{G}_M) = 0$. $\Gamma(X, \mathscr{R})$ and $\Gamma(X, \mathscr{Q})$ can be topologized as Banach spaces; θ_M is an isomorphism of these spaces and differs from Δ by the operator $u \to u \cdot M$, which has compact square. Fredholm index theory then makes $\dim \Gamma(X, \mathscr{G}_M) - \dim H^1(X, \mathscr{G}_M)$ independent of M and thus zero; in particular, $\dim \Gamma(X, \mathscr{H}) = \dim H^1(X, \mathscr{H})$, the index-zero theorem of [**30**].

The following questions arise with respect to these results (three are keyed by Roman numbers):

(I) Even for the weakest axioms, for any open $V \subseteq U$ there is a natural mapping $H^1_k(V, \mathscr{H}) \to H^1_k(U, \mathscr{H})$ and thus a natural transpose $H^1_k(U, \mathscr{H})' \to H^1_k(V, \mathscr{H})'$; the family of all such spaces and maps forms a presheaf that is a natural adjoint object to $\mathscr{H}$. This is naturally isomorphic to $\mathscr{H}^*$ when the latter exists. Can one realize it as a sheaf of functions, measures, or anything reasonable-looking? This might localize and extend J. C. Taylor's results [**25**].

(II) If $1 \in \mathscr{H}$ and the Brelot axioms hold, there is a natural way to use local cones of potentials to order each $H^1_k(U, \mathscr{H})$. A positive linear functional on $H^1_k(X, \mathscr{H})$ is a flux functional as in [**28**]. Any normal structure $\mathfrak{L}$ with $1 \in \mathscr{H}^{\mathfrak{L}}$ generates a flux functional; in particular, if $(X, \mathscr{H})$ has no global potential the least Dirichlet solutions of [**28**] can be used to show that such a functional exists (see also V. Anandam [**1**], [**2**]). Does such a functional exist in general? If it does, it can be used to construct normal structures.

(III) A number of interesting unanswered questions concern perturbed harmonic structures. In [**30**] it is shown that the weakest "Bauer" convergence axiom [**11**] is inherited from $\mathscr{H}$ by any $\mathscr{G}_M$; it can be shown that nuclearity of section spaces is inherited if $M \geqq 0$. Are side conditions on M needed in order that the "Doob" or "Brelot" convergence axioms be preserved? For a "completeness" question, it is easy to see that if $M \geqq 0$ the sheaf $\mathscr{H}$ dominates $\mathscr{G}_M$ in P. Loeb's ordering [**19**]. Is it true that if $\mathscr{H} \geqq \mathscr{K}$ in that ordering, then there exists $0 \leqq M \in \Gamma(X, \mathscr{Q})$ with $\mathscr{K} = \mathscr{G}_M$?

For global problems, suppose X compact. If $M \in \Gamma(X, \mathscr{Q})$ is so positive that $\mathscr{G}_{\lambda M}$ has a potential for large $\lambda \geqq 0$, then λ's for which $\mathscr{G}_{\lambda M} \neq \{0\}$ can be interpreted as eigenvalues. Can one arrange that they have multiplicity 1, or is that true for "generic" M (cf. [**26**])? Is the map $f \to f \cdot M$ from $\Gamma(X, \mathscr{R})$ to $\Gamma(X, \mathscr{Q})$ nuclear, or does it have a nuclear power, for suitable M, or generic M, or all M?

(IV) One would like information about the kind of topological spaces that can support harmonic structures. Strong hypotheses on $\mathscr{H}$ can restrict X severely [**17**]. If nonpolar points are admitted, X can be very unpleasant [**10**], [**11**]: It is not hard to see that there are infinite-dimensional X's that support Brelot harmonic structures. These spaces are not homogeneous, and an excellent test question is whether $\boldsymbol{T}^N$ supports a group-invariant harmonic structure; perturbation theory and

harmonic analysis can be brought to bear here. A negative answer would suggest that "sufficiently homogeneous" harmonic spaces may have to be locally Euclidean.

A drawback of knowing the spaces $H^q(X, \mathscr{H})$ for all q is that one knows they contain no information about X. The classical case shows that an axiomatic theory of biharmonic spaces [**24**] would contain no more. Suppose X compact; replacing $\mathscr{H}$ by a suitable $\mathscr{G}_M$ if necessary, one might as well assume $1 \in \mathscr{H}$. One can then formally construct the quotient sheaf $\mathscr{H}/\boldsymbol{R}$; classically, this is the sheaf of germs of harmonic differentials. The cohomology exact sequence associated with $0 \to \boldsymbol{R} \to \mathscr{H} \to \mathscr{H}/\boldsymbol{R} \to 0$ can be used to show that if X has a basis of open U's with $H^1(U, \boldsymbol{R}) = \{0\}$, then the section spaces of $\mathscr{H}/\boldsymbol{R}$ have Fréchet topologies such that restriction to relatively compact subsets is a compact operator. That is a strong condition—e.g., it implies that $H^1(X, \boldsymbol{R})$ and $H^1(X, \mathscr{H}/\boldsymbol{R})$ have finite dimensions differing at most by 1—and if it could be established under hypotheses on $\mathscr{H}$ rather than X it would severely limit the choice of X—e.g., exclude $\boldsymbol{T}^N$. In the best of all possible worlds one would like to represent $\mathscr{H}/\boldsymbol{R}$ as a sheaf of sections of a vector bundle—is this possible over a locally compact group using the distribution theory of J. Riss [**21**] or something similar?

If X is compact and $1 \in \mathscr{H}$ there is a natural map $H^1(X, \boldsymbol{R}) \to H^1(X, \mathscr{H})$ that is either onto or zero; examination of simple cases suggests that these cases correspond to Brownian motion with or without diffusion respectively. Is this, or a variant of it, the way to distinguish these cases for general $(X, \mathscr{H})$?

References

1. I. V. Anandam, *Harmonic spaces with positive potentials and nonconstant harmonic functions*, Rend. Circ. Mat. Palermo **21** (1972), fasc. 1–2, 149–167.

2. ———, *Espaces harmoniques sans potential positif*, Ann. Inst. Fourier (Grenoble) **22** (1972), fasc. 4, 97–160.

3. H. Bauer, *Harmonische Räume und ihrer Potentialtheorie*, Lecture Notes in Math., no. 22, Springer-Verlag, Berlin and New York, 1966. MR **35** #1801.

4. C. Berg and G. Forst, *Non-symmetric translation-invariant Dirichlet forms*, Invent. Math. **21** (1973), fasc. 3, 199–212.

5. J. Bliedtner, *Harmonische Gruppen und Huntsche Faltungskerne*, Seminar über Potentialtheorie, Springer, Berlin, 1968, pp. 69–102. MR **41** #500.

6. J.-M. Bony, *Détermination des axiomatiques de théorie du potentiel dont les fonctions harmoniques sont différentiables*, Ann. Inst. Fourier (Grenoble) **17** (1967), fasc. 1, 353–382.

7. ———, *Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés*, Ann. Inst. Fourier (Grenoble) **19** (1969), fasc. 1, 277–304. MR **41** #7486.

8. M. Brelot, *Lectures on potential theory*, Lectures on Math., 19, Tata Institute of Fundamental Research, Bombay, 1960. MR **22** #9749.

9. ———, *Axiomatique des fonctions harmoniques*, 2ième. éd., Séminaire de Mathématiques Supérieures, no. 14, (Été, 1965), Les Presses de l'Université de Montréal, Qué., 1969. MR **40** #393.

10. C. Constantinescu and A. Cornea, *Examples in the theory of harmonic spaces*, Seminar über Potentialtheorie, Springer, Berlin, 1968, pp. 161–171. MR **38** #6090.

11. ———, *Potential theory on harmonic spaces*, Die Grundlehren der math. Wissenschaften, Band 158, Springer-Verlag, Berlin and New York, 1972.

12. G. Forst, *Symmetric harmonic groups and translation invariant Dirichlet spaces*, Invent. Math. **18** (1972), 143–182. MR **48** #551.

13. W. Hansen, *Cohomology in harmonic spaces*, Lecture Notes in Math., vol. 226, Springer-Verlag, New York and Berlin, 1971, pp. 63–101.

14. ———, *Perturbation of harmonic spaces and construction of semigroups*, Invent. Math. **19** (1973), 149–164.

15. R.-M. Hervé, *Recherches axiomatiques sur la théorie des fonctions surharmoniques et du potentiel*, Ann. Inst. Fourier (Grenoble) **12** (1962), 415–571. MR **25** #3186.

16. K. Janssen, *On the existence of a Green function for harmonic spaces*, Math. Ann. **208** (1974), 295–303.

17. J. Král and J. Lukeš, *Indefinite harmonic continuation*, Časopis Pěst. Mat. **98** (1973), 87–94, 100. MR **47** #8894.

18. Z. Kuramochi, *Evans-Selberg's theorem on abstract Riemann surfaces with positive boundaries*, I, Proc. Japan. Acad. **32** (1956), 228–233. MR **18**, 290.

19. P. A. Loeb, *An axiomatic treatment of pairs of elliptic differential equations*, Ann. Inst. Fourier (Grenoble) **16** (1966), fasc. 2, 167–208. MR **37** #3039.

20. L. Pontrjagin, *Continuous groups*, GITTL, Moscow, 1938; English transl., Princeton Math. Ser., vol. 2, Princeton Univ. Press, Princeton, N. J., 1939. MR **1**, 44.

21. J. Riss, *Éléments de calcul, différentiel et théorie des distributions sur les groupes abeliens localement compacts*, Acta Math. **89** (1953), 45–105. MR **14**, 949.

22. L. Sario, *A linear operator method on arbitrary Riemann surfaces*, Trans. Amer. Math. Soc. **72** (1952), 281–295. MR **13**, 735.

23. H. Schaefer, *Some spectral properties of positive linear operators*, Pacific J. Math. **10** (1960), 1009–1019. MR **22** #5893.

24. E. P. Smyrnélis, *Axiomatique d'un problème de Dirichlet dans les espaces biharmoniques*, C. R. Acad. Sci. Paris Sér. A-B **274** (1972), A1897–A1900. MR **46** #5660.

25. J. C. Taylor, *Duality and the Martin compactification*, Ann. Inst. Fourier (Grenoble) **22** (1972), 95–130.

26. K. Uhlenbeck, *Eigenfunctions of Laplace operators*, Bull. Amer. Math. Soc. **78** (1972), 1073–1076. MR **47** #7771.

27. G. F. Vincent-Smith, *A construction of the kernel in axiomatic potential theory*, J. London Math. Soc. (2) **6** (1972), part 1, 24–28. MR **47** #5284.

28. B. Walsh, *Flux in axiomatic potential theory*. I: *Cohomology*, Invent. Math. **8** (1969), 175–221. MR **42** #532.

29. ———, *Flux in axiomatic potential theory*. II: *Duality*, Ann. Inst. Fourier (Grenoble) **19** (1969), fasc. 2, 371–417. MR **42** #2023.

30. ———, *Perturbation of harmonic structures and an index-zero theorem*, Ann. Inst. Fourier (Grenoble) **20** (1970), fasc. 1, 317–359. MR **43** #554.

RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Stochastic Integrals in the Plane

John Walsh

The work described here was done in collaboration with R. Cairoli; it will be published in full elsewhere. In that article, however, we treat integration with respect to arbitrary square-integrable martingales, which requires the construction of some expensive machinery. In the present paper we will try to sketch some parts of the subject in their simplest terms.

I. White noise and line integrals. Let us begin with white noise in $\boldsymbol{R}^2_+$ (the positive quadrant of the plane). This is a finitely additive set function W defined on the Borel subsets of $\boldsymbol{R}^2_+$ such that

(i) $W(A)$ is a $N(0, |A|)$ random variable;

(ii) if $A \cap B = \varnothing$, then $W(A)$ and $W(B)$ are independent.

It is natural to consider stochastic integration with respect to W—this has in fact been done by numerous authors—but let us proceed slowly and first define a stochastic process W_{st} whose parameter set is $\boldsymbol{R}^2_+$. Let R_{st} denote the rectangle $[0, s] \times [0, t]$, and define $W_{st} = W(R_{st})$. (We use W to denote both the measure and the process; this will not cause confusion.)

W_{st}, $s, t \geqq 0$, is called the *two-parameter Wiener process*, or the *Brownian sheet* (to visualize its sample paths, picture a wrinkled bed sheet). It is a continuous mean zero Gaussian process; its covariance function is easily calculated from (i) and (ii). If t is fixed, the process $s \to W_{st}$ is a Brownian motion, as can be seen from its covariance function. Since the theory of stochastic integration with respect to Brownian motion is well known, we can have stochastic line integrals with respect to W along the lines $t =$ constant. By symmetry, we can also integrate along the lines $s =$ constant. Putting these two together, we can integrate along polygonal curves in $\boldsymbol{R}^2_+$ consisting of a finite number of horizontal and vertical segments. We call such curves *staircases*.

The above remarks tacitly assume we are integrating in the direction of increasing s and t, but we can define the integral in the direction of decreasing s and t simply by changing sign. We denote the line integral of ϕ over an oriented staircase Γ by $\int_\Gamma \phi\, \partial W$.

It is convenient to introduce an analogue of the integral of a differential form. This is totally trivial in the present setting but serves to simplify our notation. If Γ is a staircase, $\int_\Gamma \phi \partial_1 W$ is the line integral of ϕ over the *horizontal* segments of Γ, and $\int_\Gamma \phi\, \partial_2 W$ is the integral over the *vertical* segments. Then we have

$$\int_\Gamma \phi\, \partial W = \int_\Gamma \phi\, \partial_1 W + \int_\Gamma \phi\, \partial_2 W.$$

All three of the above line integrals can be extended to sufficiently regular curves by a limiting argment, but staircases will suffice here.

If $f(x, s, t)$ is twice continuously differentiable in its arguments, we can write Ito's formula in differential notation:

$$\begin{aligned} \partial_1 f(W_{st}, s, t) &= \frac{\partial f}{\partial x}(W_{st}, s, t)\, \partial_1 W \\ &\quad + \left(\frac{t}{2}\frac{\partial^2 f}{\partial x^2}(W_{st}, s, t) + \frac{\partial f}{\partial s}(W_{st}, s, t)\right) ds \end{aligned} \tag{1}$$

with the symmetric equation holding for $\partial_2 f$.

II. Martingales. Stochastic integration is inextricably mixed with martingale theory, and, before we go further, we should look at martingales in our context. We are working in the plane, so the processes we consider will have $\boldsymbol{R}_+^2$ as a parameter set. We give $\boldsymbol{R}_+^2$ the usual partial order: $(s, t) \prec (u, v) \Leftrightarrow s \leqq u$ and $t \leqq v$.

Define σ-fields $\mathscr{F}_z$, $z \in \boldsymbol{R}_+^2$, by $\mathscr{F}_z = \sigma\{W_\zeta, \zeta \prec z\}$. A stochastic process $\{M_z, z \in \boldsymbol{R}_+^2\}$ is a *martingale* relative to the fields $(\mathscr{F}_z)$ if

(i) $E\{|M_z|\} < \infty$, all z;

(ii) M_z is $\mathscr{F}_z$-measurable;

(iii) $z \prec z' \Rightarrow E\{M_{z'} \mid \mathscr{F}_z\} = M_z$.

Thus we are talking about martingales with a partially ordered two-dimensional parameter set. Cairoli [1] has proved versions of both the maximal inequality and the martingale convergence theorem in this setting, but on balance it seems that relatively little is known about such martingales, and almost nothing about the corresponding sub- and supermartingales. This is one of the principal difficulties of stochastic integration in higher dimensions. Indeed, extensions of the classical theory can be quite delicate, as is indicated by the following two facts:

1°. If $\{M_z, z \in \boldsymbol{R}_+^2\}$ is a separable martingale relative to $(\mathscr{F}_z)$ which is bounded in $L \log L$, its sample functions are a.s. continuous.

2°. There exists a separable, uniformly integrable—and hence L^1-bounded—martingale which is everywhere discontinuous with probability one.

REMARK. The continuity of $L \log L$-bounded martingales is a property of the particular fields $(\mathscr{F}_z)$. For more general fields it is not even known whether or not all bounded martingales have a right continuous version.

III. Surface integrals. Stochastic integration with respect to W can be defined,

following Ito, exactly as in the classical case. We will outline this briefly.

If $A \subset \boldsymbol{R}^2_+$ is a closed rectangle with lower left-hand corner z_0, define ϕ by

$$\phi_z = \phi_0 I_A(z), \qquad z \in \boldsymbol{R}^2_+, \tag{2}$$

where ϕ_0 is $\mathscr{F}_{z_0}$-measurable and square-integrable. Then let

$$(\phi \cdot W)_z = \int_{R_z} \phi \, dW =_{\text{def}} \phi_0 W(R_z \cap A).$$

This defines a stochastic process $\phi \cdot W$ which is

(i) a.s. continuous in z,

(ii) a martingale, and

(iii) $E\{(\phi \cdot W)_z^2\} = \int_{R_z} E\{\phi_\zeta^2\} \, d\zeta$.

If ϕ is simple, i.e., a finite sum of processes of the form (2), its integral is defined by linearity. In general, if ϕ satisfies

(a) ϕ_z is $\mathscr{F}_z$-measurable,

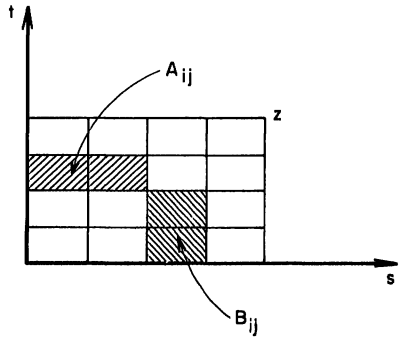
(b) $(z, \omega) \to \phi_z(\omega)$ is $\mathscr{B} \times \mathscr{F}$-measurable, and

(c) $\int_{R_z} E\{\phi_\zeta^2\} \, d\zeta < \infty$ for all $z \in \boldsymbol{R}^2_+$,

then we can find a sequence of simple ϕ_n for which $\int_{R_z} E\{(\phi_n - \phi)^2\} \, d\zeta \to 0$ for each z. The integrals $(\phi_n \cdot W)_z$ then converge in L^2, a subsequence converges a.s., and even uniformly for z in compacts thanks to Cairoli's maximal inequality. We define $\phi \cdot W = \lim \phi_n \cdot W$. Then $\phi \cdot W$ also satisfies (i), (ii), and (iii).

REMARK. In contrast to the classical case, there seem to be genuine difficulties in extending the integral very far beyond the square-integrable case.

IV. The measure J. It turns out that the integral we have just defined is not sufficient in this theory, and a second type of integral is needed. Wong and Zakai [3] have introduced a type of multiple integral which is exactly what is needed for a complete discussion, but we will be able to get by here with something simpler. The basic difficulty is that W, whatever else it may be, is not a product measure. We want to introduce a measure J which will in some sense satisfy $dJ \approx \partial_1 W \partial_2 W$, where $\partial_1 W$ and $\partial_2 W$ are respectively the "horizontal" and "vertical" increments of W_{st}. To be specific, let us fix a point $z \in \boldsymbol{R}^2_+$ and divide R_z into 2^{2n} congruent rectangles as in the diagram below.



Let A_{ij} and B_{ij} be the cross-hatched rectangles above. Note that if z_{ij} is the lower left-hand corner of the ijth rectangle, then $W(A_{ij}) = W_{z_{i,j+1}} - W_{z_{ij}}$. We define J_z^n by

$$J_z^n = \sum_{i,j=1}^{2^n} W(A_{ij})W(B_{ij}),$$

and define J_z by $J_z = \lim_{n\to\infty} J_z^n$.

One can show this limit exists and defines a continuous martingale which is orthogonal to W, that is $E\{J_z W_{z'}\} = 0$ for all z, z'. We get the desired measure, again denoted by J, by setting $J(R_z) = J_z$, and extending it to all rectangles in the obvious way. Integrals with respect to J are defined exactly as in §III.

V. Green's formula. We need to introduce the notion of a stochastic partial derivative. We define this globally: A process $\{\Phi_{st}, s, t \geqq 0\}$ has a *stochastic partial* (ϕ, ψ) *with respect to* (W, t) if there exist adapted measurable processes ϕ and ψ such that, for each s, $t \geqq 0$,

$$\Phi_{st} = \Phi_{0s} + \int_0^t \phi_{sv}\, \partial_2 W_{sv} + \int_0^t \psi_{sv}\, dv. \tag{3}$$

We will often simply speak of a stochastic partial with respect to t. We remark that if Φ is a martingale, then ψ vanishes a.e.

THEOREM 1 (GREEN'S FORMULA FOR RECTANGLES). *Let Φ be an adapted, measurable, L^2-bounded process which has stochastic partials (ϕ, ψ) with respect to (W, t). Suppose further that $\psi \equiv 0$. Let A be a rectangle whose boundary ∂A is oriented clockwise. Then*

$$\int_{\partial A} \Phi\, \partial_1 W = \int_A \Phi\, dW + \int_A \phi\, dJ. \tag{4}$$

Similarly, if Φ has stochastic partials (ϕ, ψ) with respect to (W, s) and if ψ vanishes, then

$$\int_{\partial A} \Phi\, \partial_2 W = -\int_A \Phi\, dW - \int_A \phi\, dJ. \tag{5}$$

REMARKS. 1°. The hypothesis $\psi \equiv 0$ serves purely to simplify (4) and (5). If ψ did not vanish, we would have to add the iterated integral $\int (\int \psi_{uv}\, \partial_1 W_{uv})\, dv$ to the right-hand side of (4). However, in our applications, Φ will be a martingale, so ψ will vanish.

2°. (4) and (5) are true for regions A with sufficiently regular boundaries.

For a quick application of the theorem, take $\Phi_{st} = W_{st}$, $\phi \equiv 1$, and apply (4) to $A = R_{st}$. Using Ito's formula (1) on the line integral, we get an expression for J:

$$J_{st} = \frac{1}{2}(W_{st}^2 - st) - \int_{R_{st}} W\, dW.$$

VI. Holomorphic processes. A process $\Phi = \{\Phi_z, z \in \boldsymbol{R}_+^2\}$ is said to be *holomorphic in* $\boldsymbol{R}_+^2$, or, more simply, holomorphic, if there exists an adapted measurable process $\phi = \{\phi_z, z \in \boldsymbol{R}_+^2\}$ such that $E\{\phi_z^2\}$ is bounded for z in compacts and such that, for all $z \in \boldsymbol{R}_+^2$ and any staircase $\Gamma \subset \boldsymbol{R}_+^2$ with initial point 0 and final point z,

$$\Phi_z = \Phi_0 + \int_\Gamma \phi\, \partial W \tag{6}$$

where ϕ_0 is constant. We call ϕ the *derivative* of Φ.

It is easily seen that the line integral of ϕ around any closed staircase vanishes, and we could treat this subject from the point of view of path independent integrals. But the structure of holomorphic processes bears a striking resemblance in some respects to that of classical holomorphic functions of a complex variable, and it seems worthwhile to bring this out here. We should emphasize, though, that our processes are real-, not complex-valued.

A holomorphic process is necessarily a martingale, being defined as a stochastic integral. The class of holomorphic processes is nontrivial, since W is holomorphic (its derivative is identically one). One might ask if W^2, W^3, etc., were holomorphic. They are not, but we should remember that the stochastic analogue of z^n is not W^n, but $H_n(W_{st}, st)$, where $H_n(x, t)$ is the nth Hermite polynomial.

Here, H_n is defined by

$$H_n(x, t) = \frac{(-t)^n}{n!} e^{x^2/2t} \frac{\partial^n}{\partial x^n} e^{-x^2/2t}. \tag{7}$$

Then $H_0 \equiv 1$, $H_1(x, t) = x$, $H_2(x, t) = \frac{1}{2}(x^2 - t)$ and $H_3(x, t) = (x^3 - 3xt)/3!$. For each fixed t, the set $\{H_n(x, t)\}_{n=0}^{\infty}$ is a complete set of orthogonal polynomials relative to the weight function $e^{-x^2/2t}\,dx$. In particular, since W_{st} is $N(0, st)$, we have

$$\begin{aligned} E\{H_n(W_{st}, st)H_m(W_{st}, st)\} &= 0 \qquad \text{if } m \neq n, \\ &= (st)^n/n! \quad \text{if } m = n. \end{aligned} \tag{8}$$

Let us apply Ito's formula (1) to the process $H_n(W_{st}, st)$ along the line $t =$ constant. If we use the facts (derivable from (7)) that

$$\frac{\partial}{\partial x} H_n = H_{n-1} \quad \text{and} \quad \frac{1}{2}\frac{\partial^2}{\partial x^2} H_n + \frac{\partial}{\partial t} H_n = 0, \tag{9}$$

we see that

$$H_n(W_{st}, st) = \int_0^s H_{n-1}(W_{ut}, ut)\, \partial_1 W_{ut};$$

by symmetry, this is also

$$= \int_0^t H_{n-1}(W_{sv}, sv)\, \partial_2 W_{sv}.$$

We conclude from this that $\{H_n(W_{st}, st),\ s, t \geqq 0\}$ is a holomorphic process, with derivative $H_{n-1}(W_{st}, st)$. As sums of these processes are also holomorphic, the class of holomorphic processes is evidently relatively large. Now let us turn to the general case.

THEOREM 2. *If Φ is holomorphic and Γ is a closed staircase, then $\int_\Gamma \Phi\, \partial W = 0$.*

This is almost immediate. It reduces to the case where Γ is the boundary A of a rectangle, where it follows directly from (4) and (5) and the observations that Φ, being holomorphic, has a stochastic partial ϕ with respect to both s and t, and that

$$\int_{\partial A} \Phi\, \partial W = \int_{\partial A} \Phi\, \partial_1 W + \int_{\partial A} \Phi\, \partial_2 W.$$

It follows that we can define a holomorphic process Ψ with derivative Φ by $\Psi_z = \int_0^z \Phi\, \partial W$. Thus the integral of a holomorphic process is holomorphic. What about the derivative? This is a more delicate question, but it still has a positive answer.

THEOREM 3. *Let Φ be holomorphic. Then Φ admits a derivative ϕ which is also holomorphic.*

It follows from this that Φ has holomorphic derivatives of all orders. We denote the nth derivative of Φ by $\Phi^{(n)}$. We then have the analogue of Taylor's theorem, which gives the basic structure of holomorphic processes.

THEOREM 4. *If Φ is holomorphic, then*

$$\Phi_{st} = \sum_{n=0}^{\infty} \Phi_0^{(n)}\, H_n(W_{st}, st),$$

where the above series converges in L^2 for all $s, t \geqq 0$.

REMARKS. 1°. $\mathscr{F}_0$ is trivial, so the coefficients $\Phi_0^{(n)}$ are constant.

2°. Theorem 4 implies Theorem 3—indeed, the derivative of Φ is evidently

$$\Phi'_{st} = \sum_{n=1}^{\infty} \Phi_0^{(n)}\, H_{n-1}(W_{st}, st).$$

However, by far the most difficult part of the proof of Theorem 4 lies in establishing Theorem 3.

The most striking aspect of Theorem 4 is that we knew a priori only that Φ_{st} was $\mathscr{F}_{st}$-measurable. This is much weaker than what turns out to be true, namely that Φ_{st} is actually a function of W_{st}.

Let us close this article with one further result which indicates that the existence of stochastic partials is more demanding than one might think. In proving Theorem 3, it turns out that one constructs a holomorphic version of the derivative using only the fact that Φ is a martingale and has stochastic partials with respect to both s and t, but *without using the fact that they are equal.* Thus, applying Theorem 1 to this derivative, we get:

THEOREM 5. *Suppose $\{M_z,\ z \in \mathbf{R}_+^2\}$ is a square-integrable martingale which has stochastic partials with respect to both s and t. Then M is a holomorphic process.*

References

1. R. Cairoli, *Une inégalité pour martingales à indices multiples et des applications*, Séminaire de Probabilités. IV (Univ. Strasbourg, 1968/69), Lecture Notes in Math., vol. 124, Springer-Verlag, Berlin, 1970, pp. 1–27. MR 42 #5312.

2. R. Cairoli and J. B. Walsh, *Stochastic integrals in the plane*, Acta Math. (to appear).

3. E. Wong and M. Zakai, *Martingales and stochastic integrals for processes with a multi-dimensional parameter*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **29** (1974), 109–122.

UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BRITISH COLUMBIA, CANADA V6T 1W5

Section 12

Complex Analysis

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Theory of Factorization and Boundary Properties of Functions Meromorphic in the Disk

M. M. Džrbašjan

1. The now classical formula of Jensen and Nevanlinna and the most important notation of the characteristic function which is deduced from it constitute the basis of the modern theory of meromorphic functions.

The problem of a complete description of the class of functions $F(z)$ meromorphic in the disk $|z| < 1$ for which the termwise passage to the limit in the Jensen-Nevanlinna formula is possible, yielding a representation for the function $\log F(z)$ in the entire open disk $|z| < 1$, was first posed and solved by R. Nevanlinna in the mid-twenties [1].

Like the investigations in the value distribution theory of meromorphic functions, the solution of this problem is also based on the determination of the characteristic function $T(r; F)$.

Defining the class N as the set of functions meromorphic in $|z| < 1$ for which $\sup_{0<r<1} T(r; F) < +\infty$, R. Nevanlinna showed that the mentioned termwise passage to the limit is possible only for the functions of the class N. Thus he proved the possibility of complete parametric representation and factorization of the class N.

2. In connection with this basic statement of R. Nevanlinna and its method of proof, the following problem arises naturally.

Do there exist other, more general, formulae of the Jensen-Nevanlinna type, permitting the establishment of parametric representation and factorization for more classes larger than N and for more restricted classes of functions meromorphic in the disk?

In the early papers [2], [3] of the author apparently the first attempt in the direc-

tion of solving a problem of this kind was undertaken, although the result obtained there was far from complete.

Considerably later [4] the author returned to the problem of factorization of meromorphic functions in a general form, constructing a general theory of parametric representation of N_α classes $(-1 < \alpha < +\infty)$ of functions meromorphic in the disk $|z| < 1$ as complete as the now classical theory of the class N. These results were presented in detail in the monograph [5] by the author.

3. The definitive and complete theory of factorization of functions meromorphic in the disk $|z| < 1$, including the theory of the N_α classes as a special case, was constructed in the recent investigations by the author [6], [7].

A cursory survey of these results follows.

Denote by Ω the set functions $\omega(r)$, $\omega(r) \in L[0, 1)$, subject to the following conditions:

(1) $\omega(r) > 0$, $\omega(0) = 1$, continuous on $[0, 1]$;

(2) the integral $k_\omega = \int_0^1 (1 - \omega(x))/x\, dx$ converges absolutely.

In the class of functions $\varphi(r)$, possessing a bounded and piecewise continuous derivative on $[0, 1)$, the operator $L^{(\omega)}\{\varphi(r)\}$ will be defined as follows:

$$L^{(\omega)}\{\varphi(r)\} = \varphi(0) + r\int_0^1 \varphi'(r\tau)\,\omega(\tau)\,d\tau, \quad r \in [0,1], \tag{1}$$

$$L^{(\omega)}\{\varphi(r)\} \equiv \varphi(r), \text{ if } \omega(r) \equiv 1, \ r \in [0,1). \tag{2}$$

The operator $L^{(\omega)}\{\varphi(r)\}$ is an essential generalization of the Riemann-Liouville operator $D^{-\alpha}$ $(-1 < \alpha < +\infty)$ so far as in the special case $\omega(x) = (1 - x)^\alpha$ $(-1 < \alpha < +\infty)$ the identity

$$L^{(\omega)}\{\varphi(r)\} \equiv \Gamma(1 + \alpha)\, r^{-\alpha} D^{-\alpha} \varphi(r)$$

holds.

The functions

$$C(z; \omega) = \sum_{k=0}^{\infty} \frac{z^k}{\Delta(k)}, \quad S(z; \omega) = 1 + 2\sum_{k=1}^{\infty} \frac{z^k}{\Delta(k)}, \tag{3}$$

where $\Delta(0) = 1$, $\Delta(\lambda) = \lambda \int_0^1 \omega(r)\, r^{\lambda-1}\, dr$, $\lambda \in (0, +\infty)$, and also the following three functions, which play an important role in the entire theory,

$$V_\omega(re^{i\varphi}; \zeta) = L^{(\omega)}\left\{\log\left|1 - \frac{re^{i\varphi}}{\zeta}\right|\right\},$$

$$W_\omega(re^{i\varphi}; \zeta) = \frac{1}{2\pi}\int_0^{2\pi} S(e^{i\vartheta}z; \omega)\, V_\omega(e^{i\vartheta}; \zeta)\, d\zeta, \tag{4}$$

$$A_\omega(z; \zeta) = \left(1 - \frac{z}{\zeta}\right)\exp(-W_\omega(z; \zeta)) \qquad (0 < |\zeta| \leqq 1)$$

are associated with the operator $L^{(\omega)}$.

THEOREM 1. *For an arbitrary function $F(z)$ meromorphic in the disk $|z| < 1$ having the zeros $\{a_\mu\}$ and poles $\{b_\nu\}$, and for arbitrary $\omega(r) \in \Omega$, ρ $(0 < \rho < 1)$ the formula*

$$\log F(z) = i \operatorname{Arg} c_\lambda + \lambda k_\omega + \lambda \log (z/\rho)$$

$$+ \sum_{0<|a_\mu|\leqq\rho} \log A_\omega\left(\frac{z}{\rho}; \frac{a_\mu}{\rho}\right) - \sum_{0<|b_\nu|\leqq\rho} \log A_\omega\left(\frac{z}{\rho}; \frac{b_\nu}{\rho}\right)$$

$$+ \frac{1}{2\pi} \int_0^{2\pi} S\left(e^{i\vartheta} \frac{z}{\rho}; \omega\right) L^{(\omega)} \{\log |F(\rho e^{i\vartheta})|\} \, d\vartheta \qquad (|z| < \rho) \tag{5}$$

holds, where λ *and* c_λ *are defined from the decomposition* $F(z) = c_\lambda z^\lambda + c_{\lambda+1} z^{\lambda+1} + \cdots$ $(c_\lambda \neq 0)$ *in a vicinity of the point* $z = 0$.

Note that in the special case $\omega(r) \equiv 1$, when the formula (5) coincides with that of Jensen-Nevanlinna, here, as in the theory of R. Nevanlinna, the general formula (5) leads naturally to the definition of the functions

$$m_\omega(r; F) \equiv \frac{1}{2\pi} \int_0^{2\pi} L_+^{(\omega)} \{\log |F(re^{i\varphi})|\} \, d\varphi \tag{6}$$

and

$$N_\omega(r; F) \equiv \int_0^1 \frac{n(t; \infty) - n(0; \infty)}{t} \, \omega\left(\frac{t}{r}\right) dt + n(0; \infty) \{\log r - k_\omega\}, \tag{7}$$

where $L_+^{(\omega)} = \max \{L^{(\omega)}, 0\}$ and $n(t, \infty)$ is the number of poles of the function $F(z)$ in the disk $|z| \leqq t$, counted in accordance to their multiplicity.

By means of these functions the function

$$T_\omega(r; F) \equiv m_\omega(r; F) + N_\omega(r; F) \tag{8}$$

is finally defined, which we shall call the ω-characteristic. It coincides with the function $T(r; F)$ of R. Nevanlinna when $\omega(r) \equiv 1$.

Finally, to every function $\omega(r) \in \Omega$ an $N\{\omega\}$ class is associated, as the set of functions $F(z)$ meromorphic in the disk $|z| < 1$ subject to the condition $\sup_{0<r<1} T_\omega(r; F) < +\infty$.

By this the $N\{\omega\}$ class coincides with the class N of R. Nevanlinna in the special case when $\omega(r) \equiv 1$. The comparison between the $N\{\omega\}$ classes, when $\omega(r) \not\equiv 1$, and the class N is given in the following theorem.

THEOREM 2. *Let* $\omega(r) \in \Omega$. *Then*

1. *If* $\omega(r)$ *is nondecreasing on* [0, 1), *then* $N\{\omega\} \subset N$.
2. *If* $\omega(r)$ *is nonincreasing on* [0, 1), *then* $N \subset N\{\omega\}$.

4. Passing to the fundamental theorems of the theory of factorization of the $N\{\omega\}$ classes, consider first the theorem about the parametric representation of the $N\{\omega\}$ classes.

THEOREM 3. *The* $N\{\omega\}$ *class coincides with the set of functions representable in the disk* $|z| < 1$ *in the form*

$$F(z) = e^{i\gamma + \lambda k_\omega} z^\lambda \frac{B_\omega(z; a_\mu)}{B_\omega(z; b_\nu)} \exp\left\{\frac{1}{2\pi} \int_0^{2\pi} S(e^{-i\vartheta} z; \omega) \, d\psi(\vartheta)\right\}, \tag{9}$$

where

$$B_\omega(z; a\mu) = \prod_{\mu=1}^{\infty} A_\omega(z; a_\mu), \; B_\omega(z; b_\nu) = \prod_{\nu=1}^{\infty} A_\omega(z; b_\nu) \tag{10}$$

are converging products from the $N\{\omega\}$ class, $\psi(\vartheta)$ is a real function on $[0, 2\pi]$ with $\bigvee_0^{2\pi} \psi(\vartheta) < +\infty$, $\lambda \leqslant 0$ is an arbitrary integer, γ is an arbitrary real number.

In the case when the function $\omega(r) \in \Omega$ is nondecreasing on [0,1] and thus the inclusion $N\{\omega\} \subset N$ is true, there holds

THEOREM 4. 1. *If the sequence $\{z_k\}_1^\infty$ satisfies the condition*

$$\sum_{k=1}^{\infty} \int_{|k|}^{1} \omega(x)\, dx < +\infty, \tag{11}$$

then the representation

$$B_\omega(z; z_k) = B(z; z_k) \exp\left\{\frac{1}{2\pi} \int_0^{2\pi} S(e^{-i\vartheta} z; \omega)\, d\mu(\vartheta)\right\} \tag{12}$$

holds, where $\mu(\vartheta)$ is a certain nonincreasing and bounded function on $[0, 2\pi]$.

2. *The class $N\{\omega\} \subset N$ coincides with the set of functions representable in the disk $|z| < 1$ in the form*

$$F(z) = e^{i\gamma + \lambda k_\omega} z^\lambda \frac{B(z; a_\mu)}{B(z; b_\nu)} \exp\left\{\frac{1}{2\pi} \int_0^{2\pi} S(e^{-i\vartheta} z; \omega)\, d\psi(\vartheta)\right\} \tag{13}$$

where $B(z; a_\mu)$ and $B(z; b_\nu)$ are the Blaschke functions with zeros satisfying the conditions (11).

3. *Any function $F(z) \in N\{\omega\} \subset N$ is representable in the form $F(z) = f_1(z)/f_2(z)$ ($|z| < 1$) where $f_k(z) \in N\{\omega\}$, $|f_k(z)| \leqq 1$ ($|z| < 1$) are analytical in the disk $|z| < 1$.*

THEOREM 5. *For any function $F(z)$ meromorphic in the disk $|z| < 1$ there exists a function $\omega_F(r) \in \Omega$ such that $F(z) \in N\{\omega_F\}$.*

5. As is well known, the class N possesses important boundary properties.

For any function $F(z) \in N$ the limit

$$F(e^{i\vartheta}) = \lim_{r \to 1-0} F(re^{i\vartheta}) \tag{14}$$

exists for almost all $\vartheta \in [0, 2\pi]$; by this, if $F(z) \not\equiv 0$ then,

$$\int_0^{2\pi} \left|\log\left|F(e^{i\vartheta})\right|\right| d\vartheta < +\infty. \tag{15}$$

The boundary properties of the $N\{\omega\}$ classes have been investigated with respect to Theorem 2.

The case $N\{\omega\} \supset N$. Here the following theorems are established.

THEOREM 6. *Let $F(z) \in N\{\omega\}$, where $\omega(r)$ is nonincreasing and satisfies the condition* Lip 1 *on every interval $[0, \Delta]$ ($0 < \Delta < 1$).*[1] *Then*:

1. *The limit*

$$\lim_{r \to 1-0} \operatorname{Re} L^{(\omega)} \{\log F(re^{i\vartheta})\} = \psi'(\vartheta) \in L(0, 2\pi) \tag{16}$$

exists almost everywhere on $[0, 2\pi]$.

[1]Below the class of such functions will be denoted by $\bar{\Omega}$.

2. *We have as well*

$$\lim_{r\to 1-0} \operatorname{Re} L^{(\omega)}\{\log B_\omega(re^{i\varphi})\} = 0 \tag{17}$$

almost everywhere on $[0, 2\pi]$.

Recently [8], the existence of these limits was established generally, and, in particular, for the case when the real part sign is dropped.

The following theorem of uniqueness is an enlargement of the now classical theorem by Szegö for the classes $N\{\omega\} \supset N$.

THEOREM 7. *Let* $f(z) \in N\{\omega\}$ *be analytical in the disk* $|z| < 1$. *Then:*

1. *If* $\omega(r) \in \tilde{\Omega}$ *is nonincreasing on* $[0, 1)$ *then its boundary values*

$$L^{(\omega)}\{\log|f(e^{i\vartheta})|\} = \lim_{r\to 1-0} L^{(\omega)}\{\log|f(re^{i\vartheta})|\}$$

are such that

$$\int_0^{2\pi} L^{(\omega)}\{\log|f(e^{i\vartheta})|\}\, d\vartheta > -\infty. \tag{18}$$

2. *There is no function* $f(z) \not\equiv 0$ *analytical in the disk* $|z| < 1$ *for which*

$$\int_0^{2\pi} L^{(\omega)}\{\log|f(re^{i\vartheta})|\}\, d\vartheta = -\infty, \tag{19}$$

or

$$\sum_{(\mu)} \int_{|a_\mu|}^{1} \omega(x)\, dx = +\infty \qquad \{f(a_\mu) = 0\}. \tag{20}$$

The case $N \supset N\{\omega\}$. In connection with the known boundary properties of classes $N\{\omega\} \subset N$ noted above, the following questions arise naturally.

Does the "thin" exceptional set $E \subset [0, 2\pi]$ of linear measure zero for the classes $N\{\omega\} \subset N$, where the limit $F(e^{i\vartheta})$ of a function $F(z) \in N\{\omega\}$, perhaps not exist?

Can there be stated something else about the boundary values of a function $F(z) \in N\{\omega\} \subset N$, except the boundedness of the integral (15)?

The author and V.S. Zakharian succeeded in obtaining the positive solutions of both these problems in the papers [9], [10] in terms of the ω-capacity notation introduced by us. The ω-capacity is associated with the integral

$$\bigcup_{\omega}(z; \mu) = \int_0^{2\pi} |C(ze^{-i\vartheta}; \omega)|\, d\mu, \qquad \mu(E) = 1, \tag{21}$$

and coincides with the $(1 + \alpha)$-capacity in the Frostman sense when $\omega(x) = (1 - x)^\alpha$ $(-1 < \alpha < 0)$. The solution of the problem posed above about the boundary properties of the functions of the $N\{\omega\} \subset N$ classes is contained in the following theorems.

THEOREM 8. *For any function* $F(z) \in N\{\omega\} \subset N$ *the bounded limit* $F(e^{i\vartheta}) = \lim_{r\to 1-0} F(re^{i\vartheta})$ *exists for all* $\vartheta \in [0, 2\pi]$ *except, perhaps, an exceptional set* $E \subset [0, 2\pi]$ *having zero* ω*-capacity.*

THEOREM 9. *Let $F(z) \in N\{\omega\} \subset N$ and $E \subset [0, 2\pi]$ be any set with positive ω-capacity. Let, further, $\mu \prec E$, $\mu(E) = 1$ be any measure with the property $\bigcup_\omega(\mu) = \sup_{|z|\leq 1} \bigcup_\omega(z; \mu) < +\infty$. Then the boundary values $F(e^{i\vartheta})$ of the function $F(z)$ satisfy the condition*

$$\int_E \left|\log\left| F(e^{i\vartheta})\right|\right| d\mu(\vartheta) = \int_0^{2\pi} \left|\log\left| F(e^{i\vartheta})\right|\right| d\mu(\vartheta) < +\infty.$$

As is well known, if $f(z)$ is an analytical function from the class N, then $f(z) \equiv 0$ ($|z| < 1$), if $f(e^{i\vartheta}) = 0$, $\vartheta \in E$, meas $E > 0$.

THEOREM 10. *Let $f(z) \in N\{\omega\}$ be analytical in the disk $|z| < 1$ and $f(e^{i\vartheta}) = 0$, $\vartheta \in E$, where* meas $E = 0$ *but E has positive ω-capacity. Then $f(z) \equiv 0$ ($|z| < 1$).*

References

1. R. Nevanlinna, *Eindeutige analitische Funktionen*, Die Grundlehren der math. Wissenschaften, Band 46, Springer-Verlag, Berlin, 1936; English transl., Die Grundlehren der math. Wissenschaften, Band 162, Springer-Verlag, Berlin and New York, 1970. MR **43** #5003.

2. M. M. Džrbašjan, *On the canonical representation of functions meromorphic in the unit disk*, Dokl. Akad. Nauk Armjan. SSR **3** (1945), no. 1, 3–9. (Russian)

3. ———, *To the problem of representability of analytical functions*, Soobšč. Inst. Mat. i Mech. Akad. Nauk Armjan. SSR **2** (1948), 3–40. (Russian)

4. ———, *The parametric representation of some general classes of meromorphic functions in the unit circle*, Dokl. Akad. Nauk SSSR **157** (1964), 1024–1027. (Russian) MR **29** #6020.

5. ———, *Integral transforms and representations of functions in the complex domain*, "Nauka", Moscow, 1966. (Russian) MR **35** #370.

9. ———, *A generalized Riemann-Liouville operator and some of its applications*, Izv. Akad. Nauk SSSR Ser. Mat. **32** (1968), 1075–1111 = Math. USSR Izv. **2** (1968), 1027–1065. MR **38** #2313.

7. ———, *Theory of factorization of functions meromorphic in the disk*, Mat. Sb. **79** (**121**) (1969), 517–615 = Math. USSR Sb. **8** (1969), 493–592. MR **41** #3772.

8. R. S. Galoyan, *On the meromorphic functions of the class $N\{\omega\}$*, Izv. Akad. Nauk Armjan. SSR Mat. **7** (1972), no. 5, 334–360. (Russian)

9. M. M. Džrbašjan and V. S. Zaharjan, *Boundary properties of subclasses of meromorphic functions of bounded type*, Izv. Akad. Nauk SSSR Ser. Mat. **34** (1970), 1262–1339 = Math. USSR Izv. 4 (1970), 1273–1354. MR **46** #9352a.

10. ———, *Boundary properties of subclasses of meromorphic functions of bounded type*, Izv. Akad. Nauk Armjan. SSR Mat. **6** (1971), no. 2/3, 182–194. (Russian)

MATHEMATICS INSTITUTE OF THE ARMENIAN ACADEMY OF SCIENCE
YEREVAN, U.S.S.R.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Some Metric Properties of Quasi-Conformal Mappings

F. W. Gehring*

1. Introduction. This article is a summary of some recent research concerned with the following problem. Suppose that $f: D \to D'$ is an n-dimensional quasi-conformal mapping and that E is a set in D. How do E and $f(E)$ compare in size? By the size of a set we mean either the Hausdorff p-dimensional measure or the Hausdorff dimension of the set.

For $p \in [0, \infty)$ the *normalized Hausdorff p-dimensional outer measure* of a set E in Euclidean n-space R^n is defined as

$$\mathscr{H}^p(E) = \lim_{t \to 0} \left(\inf \sum_j \alpha(p)\, 2^{-p} \operatorname{dia}(E_j)^p \right),$$

where the infimum is taken over all countable coverings of E by sets E_j with $\operatorname{dia}(E_j) < t$, and where

$$\alpha(p) = \frac{\Gamma(\frac{1}{2})^p}{\Gamma(p/2 + 1)}.$$

As p varies from 0 to n, the measures $\mathscr{H}^p$ interpolate in a natural way between the counting measure and Lebesgue measure in R^n. In particular, if T is a p-dimensional hyperplane in R^n, then $\mathscr{H}^p \mid T$ coincides with the Lebesgue p-dimensional outer measure m_p in T. The *Hausdorff dimension* of E is defined as

$$\mathscr{H}\text{-dim}(E) = \inf \{p: \mathscr{H}^p(E) = 0\}.$$

Clearly $0 \leqq \mathscr{H}\text{-dim}(E) \leqq n$ and $\mathscr{H}\text{-dim}(E) = p$ whenever E is an open p-dimensional planar set.

2. Hausdorff dimension under quasi-conformal mappings. Suppose that $f: D \to D'$

*This research was supported in part by the National Science Foundation, Contract GP-28115.

is a homeomorphism where D and D' are domains in R^n. Then

$$\mathcal{H}\text{-dim}\,(f(E)) = \mathcal{H}\text{-dim}\,(E) \tag{1}$$

for all $E \subset D$ if f is a diffeomorphism or, more generally, if both f and f^{-1} are locally Lipschitzian. The homeomorphism f is said to be an n-dimensional *quasi-conformal mapping* if the function

$$H(x,f) = \limsup_{r\to 0} \frac{|f(y) - f(x)|}{|f(z) - f(x)|}, \qquad |y - x| = |z - x| = r,$$

is bounded in D, in which case it follows that f has a nonsingular differential at almost all points of D. The following example shows, however, that (1) fails for this more general class of mappings [6].

EXAMPLE 1. *For each $p, q \in (0, n)$ there exist an n-dimensional quasi-conformal mapping $f\colon R^n \to R^n$ and a compact set $E \subset R^n$ such that $\mathcal{H}$-dim $(E) = p$ and $\mathcal{H}$-dim $(f(E)) = q$.*

On the other hand, we have the following result relating the Hausdorff dimensions of E and $f(E)$ with the *linear dilatation* $K(f)$, $K(f) = \text{ess sup}_{x\in D} H(x, f)$, for a quasi-conformal mapping f ([1] and [6]).

THEOREM 1. *If $f\colon D \to D'$ is an n-dimensional quasi-conformal mapping and if $p \in [0, n]$, then*

$$cp/(c + n - p) \leqq \mathcal{H}\text{-dim}\,(f(E)) \leqq (c + n)p/(c + p) \tag{2}$$

for each set $E \subset D$ with $\mathcal{H}$-dim $(E) = p$, where c is a positive constant which depends only on $K(f)$ and n.

Inequality (2) shows that (1) holds for an n-dimensional quasi-conformal mapping f whenever E is of Hausdorff dimension 0 or n.

Theorem 1 is a consequence of the following higher dimensional analogue of an important result due to Bojarski [1].

THEOREM 2. *If $f\colon D \to D'$ is an n-dimensional quasi-conformal mapping, then the partial derivatives of f are locally L^p-integrable in D for $p \in [n, n + c)$, where c is a positive constant which depends only on $K(f)$ and n.*

Theorem 2, in turn, depends on the fact that the maximal stretching

$$L(x,f) = \limsup_{y\to x} \frac{|f(y) - f(x)|}{|y - x|}$$

for the quasi-conformal mapping f satisfies a reversed Hölder inequality on all small n-cubes $Q \subset D$ and on the following result [1].

LEMMA 1. *Suppose that $q, b \in (1, \infty)$ and that Q is an n-cube in R^n. If $g\colon Q \to [0, \infty]$ is L^q-integrable in Q and if*

$$\frac{1}{m(Q')} \int_{Q'} g^q\, dm \leqq b \left(\frac{1}{m(Q')} \int_{Q'} g\, dm \right)^q$$

for all parallel n-cubes $Q' \subset Q$, then g is L^p-integrable in Q with

$$\frac{1}{m(Q)} \int_Q g^p \, dm \leqq \frac{c}{q + c - p} \left(\frac{1}{m(Q)} \int_Q g^q \, dm \right)^{p/q}$$

for $p \in [q, q + c)$, where c is a positive constant which depends only on q, b and n.

3. $\mathcal{H}^p$-absolute continuity. Suppose that $f: D \to D'$ is an n-dimensional quasi-conformal mapping and that $E \subset D$. Since any similarity mapping $g: R^n \to R^n$ is quasi-conformal, there can be no meaningful relation between $\mathcal{H}^p(E)$ and $\mathcal{H}^p(f(E))$ without some further normalization for f. On the other hand, it is well known that f is $\mathcal{H}^n$-absolutely continuous, i.e., $\mathcal{H}^n(E) = 0$ implies that $\mathcal{H}^n(f(E)) = 0$. Since $\mathcal{H}^0(E)$ is equal to the cardinality of E, f is also $\mathcal{H}^0$-absolutely continuous. The following example shows, however, that $p = 0$ and $p = n$ are the only values of $p \in [0, n]$ for which each such f is $\mathcal{H}^p$-absolutely continuous ([3] and [6]).

EXAMPLE 2. *For each $p \in (0, n)$ there exist an n-dimensional quasi-conformal mapping $f: R^n \to R^n$ and a compact set $E \subset R^n$ such that $\mathcal{H}^p(E) = 0$ and $\mathcal{H}^p(f(E)) = \infty$. Moreover when p is an integer, we can choose E so that, in addition, it lies in a p-dimensional hyperplane T.*

On the other hand, we can establish the following result [4].

THEOREM 3. *Suppose that T is a p-dimensional hyperplane in R^n where $p \in (1, n]$. If $f: D \to D'$ is an n-dimensional quasi-conformal mapping and if $\mathcal{H}^p(f(E)) < \infty$ for each compact set $E \subset D \cap T$, then $f \mid (D \cap T)$ is $\mathcal{H}^p$-absolutely continuous.*

Hence, in particular, the restriction of an n-dimensional quasi-conformal mapping f to an open p-dimensional planar set G is $\mathcal{H}^p$-absolutely continuous if $\mathcal{H}^p(f(G)) < \infty$.

An example due to Beurling and Ahlfors can be used to exhibit a quasi-conformal mapping $f: R^n \to R^n$ such that $f(R^1) = R^1$ and such that $f \mid R^1$ is not $\mathcal{H}^1$-absolutely continuous [3]. Thus the hypothesis that $p > 1$ is essential in Theorem 3.

The proof of Theorem 3 is based on the following inequality for the $\mathcal{H}^p$-measure of an open p-dimensional planar set under an n-dimensional quasi-conformal mapping [4].

THEOREM 4. *Suppose that T is a p-dimensional hyperplane in R^n where $p \in [1, n]$. If $f: D \to D'$ is an n-dimensional quasi-conformal mapping and if $x \in D \cap T$, then*

$$\mathcal{H}^p(f(D \cap T)) \geqq c \operatorname{dist}(f(x), \partial D')^p,$$

where c is a positive constant which depends only on $K(f)$ and n.

The proof for Theorem 4 depends, in turn, on the following lower bound for the Hausdorff measure of sets which link in R^n [2].

LEMMA 2. *Suppose that p is an integer in $[1, n)$. If A and B are disjoint compact sets in R^n and if B is a topological $(n-p-1)$-sphere which is not contractible in $R^n \sim A$, then $\mathcal{H}^p(A) \geqq c \operatorname{dist}(A, B)^p$, where c is a positive constant which depends only on n.*

4. Modulus inequality. Suppose that Γ is a family of curves in $\bar{R}^n = R^n \cup \{\infty\}$. For $p, q \in (0, \infty)$ we define the *p-dimensional q-modulus* of the family Γ by

$$M_q^p(\Gamma) = \inf \int_{R^n} h^q \, d\mathcal{H}^p,$$

where the infimum is taken over all Borel measurable functions $h: R^n \to [0, \infty]$ such that $\int_\gamma h \, ds \geqq 1$ for all locally rectifiable curves $\gamma \in \Gamma$. For convenience we write $M(\Gamma) = M_n^n(\Gamma)$.

The *inner* and *outer dilatations* of an n-dimensional homeomorphism $f: D \to D'$ are given by

$$K_I(f) = \sup (M(f(\Gamma))/M(\Gamma)), \qquad K_O(f) = \sup (M(\Gamma)/M(f(\Gamma))),$$

where the suprema are taken over all curve families Γ in D. The homeomorphism f is quasi-conformal if and only if these dilatations are finite, i.e., if and only if the n-dimensional n-moduli of curve families are quasi-invariant under f.

The analysis leading to the proof of Theorem 3 yields the following result on the behavior of the p-dimensional p-moduli of curve families under an n-dimensional quasi-conformal mapping [4] (cf. also [5]).

THEOREM 5. *If f satisfies the hypotheses of Theorem* 3, *then*

$$M_p^p(\Gamma) \leqq K_O(f) M_p^p(f(\Gamma))$$

for each curve family Γ in $D \cap T$.

References

1. F. W. Gehring, *The L^p-integrability of the partial derivatives of a quasiconformal mapping*, Acta Math. **130** (1973), 265–277.

2. ———, *The Hausdorff measure of sets which link in euclidean space*, Contributions to Analysis: A Collection of Papers Dedicated to Lipman Bers, Academic Press, New York, 1974.

3. ———, *Absolute continuity properties of quasiconformal mappings*, Symposia Math. (to appear).

4. ———, *Lower dimensional absolute continuity properties of quasiconformal mappings*, Math. Proc. Cambridge Philos. Soc. **77** (1975) (to appear).

5. F. W. Gehring and J. Väisälä, *The coefficients of quasiconformality of domains in space*, Acta Math. **114** (1965), 1–70. MR **31** #4905.

6. ———, *Hausdorff dimension and quasiconformal mappings*, J. London Math. Soc. (2) **6** (1973), 504–512. MR **48** #2380.

UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN 48104, U.S.A.

Труды Международного Конгресса Математиков
Ванкувер, 1974

О Представлении Аналитических Функций Рядами Дирихле

А. Ф. Леонтьев

1. Пусть $\bar{D}$—конечная замкнутая выпуклая область, $K(\varphi)$—её опорная функция, $h(\varphi) = K(-\varphi)$, $L(\lambda)$—целая функция экспоненциального типа с индикатрисой роста $h(\varphi)$, λ_k $(k \geqq 1)$ —нули $L(\lambda)$. Допустим, что все нули—простые. Обозначим через $\psi_k(t)$ $(k \geqq 1)$ функции, ассоциированные по Борелю с $[L'(\lambda_k)(\lambda - \lambda_k)]^{-1} L(\lambda)$, они регулярны вне $\bar{D}$. Имеем

$$\frac{1}{2\pi i} \int_C \psi_k(t) e^{\lambda_m t}\, dt = \delta_{km},$$

где C—замкнутый контур, охватывающий $\bar{D}$. В силу этого функции $f(z)$, аналитической на $\bar{D}$, сопоставим [1] ряд Дирихле

$$(1) \qquad f(z) \sim \sum_{k=1}^{\infty} a_k e^{\lambda_k z}, \qquad a_k = \frac{1}{2\pi i} \int_C f(t)\psi_k(t)\, dt.$$

Имеет место [2] теорема единственности: если все $a_k = 0$, то $f(z) \equiv 0$. Ряд (1) вообще не может сходиться в области $G \supset \bar{D}$. Доказано [3]: 1) для того, чтобы ряд (1) сходился в D (открытой части $\bar{D}$), какова бы ни была функция $f(z)$, аналитическая на $\bar{D}$, необходимо и достаточно, чтобы выполнялось условие

$$(2) \qquad \ln\left|L'(\lambda_k)\right| > [h(\varphi_k) - \varepsilon]\left|\lambda_k\right|, \qquad \lambda_k = \left|\lambda_k\right| e^{i\varphi_k},\ k > k_0(\varepsilon),\ \forall\, \varepsilon > 0;$$

2) для того, чтобы ряд (1) всегда сходился в D к функции $f(z)$, необходимо и достаточно, чтобы функция $L(\lambda)$ была функцией вполне регулярного роста и выполнялось условие (2). Функции $L(\lambda)$ вполне регулярного роста со свойством (2) для любой конечной выпуклой области $\bar{D}$ существуют. Более того,

имеются функции вполне регулярного роста со свойством (2), которые удовлетворяют дополнительному условию

$$|L(re^{i\varphi})| < e^{h(\varphi)r}/r^{\mu}, \qquad \mu > 1. \tag{3}$$

При этом условии функции $\phi_k(t)$ аналитичны вне $\bar{D}$ и непрерывны вне D. В этом случае ряд (1) можно сконструировать для функций $f(z)$, аналитических в D и непрерывных в $\bar{D}$ (в формуле для коэффициентов в качестве C надо взять границу области D). Если $L(\lambda)$ вполне регулярного роста и удовлетворяет условиям (2) и (3), то ряд (1) и в новой ситуации сходится в D к $f(z)$. Вопрос о разложении в ряд Дирихле функций, аналитических лишь в области D, решается положительно на основании следующей теоремы [4]: пусть $F(z)$ регулярна в конечной выпуклой открытой области D. Имеются функция $f(z)$, аналитическая в D и непрерывная в $\bar{D}$, и целая функция $M(\lambda) = \sum_0^\infty b_k\lambda^k$ с ростом не выше первого порядка минимального типа такие, что

$$M(D)f(z) \equiv \sum_0^\infty b_k f^{(k)}(z) = F(z).$$

Представив $f(z)$ рядом (1), в силу указанной теоремы

$$F(z) = \sum_0^\infty a_k M(\lambda_k)e^{\lambda_k z}, \qquad z \in D.$$

Как следствие этого результата получаем: произвольную функцию $F(z)$, аналитическую в выпуклом многоугольнике D с вершинами в точках $\gamma_1, \cdots, \gamma_p$, можно представить в виде

$$F(z) = \sum_{k=1}^{p} F_k(z),$$

где $F_k(z)$ регулярна в полуплоскости D_k, ограниченной прямой, проходящей через вершины γ_k и γ_{k+1} ($D \subset D_k$), и периодична с периодом равным $\gamma_{k+1} - \gamma_k$ ($\gamma_{p+1} = \gamma_1$).

Если на функцию $L(\lambda)$ не накладывать никаких ограничений кроме того, что она экспоненциального типа с индикатрисой роста $h(\varphi)$, то ряд (1) может не сходиться ни в какой области. Тем не менее его коэффициенты a_k ($k \geqq 1$), в силу теоремы единственности, принципиально должны определять функцию $f(z)$. В [5] рассмотрена ситуация, когда все λ_k вещественны и $L(\lambda)$ вещественна на вещественной оси. В этой ситуации при некоторых $r_k \uparrow \infty$ определены функции

$$f^+(z) = \lim_{k\to\infty} \sum_{0<\lambda_\nu<r_k} a_\nu e^{\lambda_\nu z}, \qquad f^-(z) = \lim_{k\to\infty} \sum_{0<-\lambda_\nu<r_k} a_\nu e^{\lambda_\nu z}$$

соответственно в углах $|\arg(z - \alpha^+) - \pi| < \varphi_0 < \pi/2$, $|\arg(z - \alpha^-)| < \varphi_0$, где $\alpha^+ < 0 < \alpha^-$. Речь идёт о том, как восстановить $f(z)$ с помощью $f^+(z)$ и $f^-(z)$. Ответ состоит в следующем. Возможны три случая: (1) функции $f^+(z)$ и $f^-(z)$ аналитически продолжаются на некоторый интервал вещественной оси и на этом интервале $f^+(x) + f^-(x) = f(x)$; (2) функции $f^+(z)$ и $f^-(z)$ квазианалити-

чески продолжаются на некоторый интервал вещественной оси и на нём $f^+(x) + f^-(x) = f(x)$; (3) функции $f^+(z)$ и $f^-(z)$ аналитически продолжаются до начала координат по вещественной оси (в этом случае $\bar{D}$—начало координат и начало координат—особая точка для функций $f^+(z), f^-(z)$), причём существуют пределы $\lim_{x\to-0} f^{+(k)}(x)$, $\lim_{x\to+0} f^{-(k)}(x)$ и

$$\lim_{x\to-0} f^{+(k)}(x) + \lim_{x\to+0} f^{-(k)}(x) = f^{(k)}(0)$$

так, что $f(z)$ можно восстановить с помощью её ряда Тейлора. И.Ф. Красичков-Терновский [6] в общей ситуации и иным способом указал, как с помощью коэффициентов a_n можно вычислить величины $f^{(k)}(0)$ $(k \geqq 0)$.

2. Можно рассмотреть вопрос о представлении функции рядом Дирихле в замкнутой области, если функция аналитична в открытой области и имеет определённую гладкость в замкнутой области. В [7] рассмотрен случай, когда

$$L(\lambda) = \int_C e^{\lambda t}\, d\sigma(t), \qquad C = \partial\bar{D},$$

где $\sigma(t)$—функция ограниченной вариации на C. Положим

$$\psi_k(t) = \frac{-1}{L'(\lambda_k)}\, e^{-\lambda_k t} \int_{\alpha_k}^{t} e^{\lambda_k \eta} d\sigma(\eta), \qquad \alpha_k \in C;$$

функции $f(z)$, аналитической в D и непрерывной в $\bar{D}$, сопоставляется ряд (1). Доказано, что если $L(\lambda)$—функция вполне регулярного роста и

$$\sum_{k=1}^{\infty} \left|\lambda_k^{-2}[L'(\lambda_k)]^{-1}\, e^{\lambda_k z}\right| < \infty, \qquad z \in \bar{D},$$

а $f(z)$ имеет непрерывные производные $f'(z), f''(z)$ в $\bar{D}$ и удовлетворяет условию

$$\int_C f(t)\, d\sigma(t) = 0,$$

то ряд (1) абсолютно сходится в замкнутой области $\bar{D}$ к $f(z)$. Если $\bar{D}$—выпуклый многоугольник, верны оценки

$$\begin{aligned} &\left|L'(\lambda_k)\right| > Ae^{h(\varphi_k)|\lambda_k|}, && \varphi_k = \arg\lambda_k, \\ &\left|L(re^{i\varphi})\right| > Be^{h(\varphi)r}, && r_k - p \leqq r \leqq r_k + p,\ r_k \uparrow \infty, \end{aligned}$$

а $f(z)$ имеет непрерывную производную $f'(z)$ в $\bar{D}$, то ряд (1) равномерно сходится к $f(z)$ в области $\bar{D}_\varepsilon$, получаемой из $\bar{D}$ выбрасыванием ε-окрестностей вершин.

3. Здесь речь пойдёт о представлении аналитических функций рядами Дирихле в бесконечных областях. Имеет место следующая теорема [8]: пусть D—левая полуплоскость $\mathrm{Re}\, z < 0$. Каково бы ни было $\rho > 1$, имеется последовательность положительных чисел $\{\lambda_k\}_1^\infty$, $\lim_{k\to\infty} k/\lambda_k^\rho = \tau$, $0 < \tau < \infty$, такая, что любую функцию $f(z)$, аналитическую в D, можно представить в виде

$$f(z) = \sum_1^\infty A_k e^{\lambda_k z} + \text{целая функция}.$$

Что касается целых функций, то относительно их верно утверждение [1]: любую целую функцию можно представить во всей плоскости рядом Дирихле, причём показатели ряда можно выбрать лежащими на трёх лучах.

Заметим, что любую функцию, аналитическую в выпуклой бесконечной многоугольной области (ограниченной конечным числом отрезков прямых) можно представить в виде суммы функций, каждая из которых регулярна в соответствующей полуплоскости. Отсюда, согласно вышеизложенному, следует положительное решение вопроса о разложении в ряды Дирихле аналитических функций в такого рода областях.

4. Пусть $L(\lambda)$—целая функция экспоненциального типа, $\gamma(t)$—функция, ассоциированная по Борелю с $L(\lambda)$, $\bar{D}$—наименьшее выпуклое замкнутое множество, содержащее все особенности $\gamma(t)$. Обозначим через $\lambda_1, \lambda_2, \cdots$ — нули $L(\lambda)$ и предположим ради простоты, что все они—простые. Если ряд (1) (или последовательность конечных линейных комбинаций $e^{\lambda_k z}$) сходится в области $G \supset \bar{D}$, то его сумма (предельная функция последовательности) удовлетворяет уравнению свёртки

$$M_L(f) \equiv \frac{1}{2\pi i} \int_C \gamma(t) f(t+z)\, dt = 0 \tag{4}$$

для z достаточно малых по модулю. Экспоненты $e^{\lambda_k z}$ $(k \geqq 1)$—элементарные решения уравнения (4). Основной вопрос в теории уравнений (4)—это вопрос о том, как и где можно выразить решение уравнения (4) через элементарные. Решению $f(z)$ уравнения (4) сопоставляем ряд (1). В простом случае, когда $L(\lambda)$—функция вполне регулярного роста, известен следующий результат Диксона [9] (см. также [10]): Пусть $f(z)$ решение уравнения (4), регулярное в области $G = \bigcup_{\alpha \in Q} \bar{D}_\alpha$, где Q—открытая область и $\bar{D}_\alpha$—смещение $\bar{D}$ на вектор α. Тогда внутри области G равномерно

$$f(z) = \lim_{k \to \infty} \sum_{|\lambda_\nu| < r_k} a_\nu e^{\lambda_\nu z}, \qquad r_k \uparrow \infty. \tag{5}$$

При плохих оценках снизу для $L(\lambda)$ представление в виде (5) или совсем не получается, или получается в области меньшей области G. В связи с этим возник вопрос о суммировании ряда (1). В ряде случаев суммирование действительно можно осуществить. Все эти случаи относятся к ситуации, когда решение $f(z)$ является регулярным в односвязной бесконечной области G. Область G будем называть областью типа левой полуплоскости, если она обладает свойствами: каково бы ни было $N > 0$, имеются точки $z = x + iy$ в G с $y > N$ и $y < -N$; если $z_0 \in G$, то и все точки горизонтального луча, идущего из z_0 влево, принадлежат G. Область, которая получается из области типа левой полуплоскости путём поворота вокруг начала координат, будем называть областью типа полуплоскости. В [11] доказано, что если решение $f(z)$ уравнения (4) регулярно в области G типа полуплоскости, то тогда ряд (1) суммируется определенным методом к $f(z)$ во всей области G. Ряд удается просуммировать к $f(z)$ ещё в двух случаях: в случае, когда G есть область

типа полосы и в случае, когда G есть область типа полуполосы. Отметим, что любая бесконечная выпуклая область принадлежит к одному из указанных трёх типов. В случае конечных областей И.Ф. Красичкову-Терновскому [13] удалось доказать аппроксимационную теорему: Если решение $f(z)$ уравнения (4) регулярно в выпуклой области $G \supset \bar{D}$, то $f(z)$ внутри G можно аппроксимировать с любой точностью линейными комбинациями элементарных решений.

5. Результаты предыдущего пункта можно использовать для выяснения структуры подпространств аналитических функций, инвариантных относительно дифференцирования. Пусть G—область комплексной плоскости, H—пространство функций, аналитических в G, $W \subset H$—замкнутое подпространство, инвариантное относительно дифференцирования. Примером инвариантного подпространства W может служить совокупность функций, аналитических в G, которые удовлетворяют одному или нескольким уравнениям свёртки. Пусть W—инвариантное подпространство. Рассмотрим всевозможные экспоненциальные одночлены $z^k e^{\lambda z}$, которые содержатся в W. Л. Шварц в 1947 году сформулировал задачу: допускает ли каждое замкнутое инвариантное подпространство W спектральный синтез, т.е. совпадает ли с замыканием линейной оболочки экспоненциальных одночленов, в нём содержащихся. В случае, когда G—вся плоскость, ответ положительный, что показал ещё сам Л. Шварц. Систематически эту задачу в последние годы исследовал И.Ф. Красичков-Терновский [12]—[14]. Он показал, что если G—бесконечная выпуклая область, то спектральный синтез имеет место всегда. В [11] установлено, что спектральный синтез всегда допустим, если G—область типа полуплоскости (она не обязательно выпукла) или типа полуполосы, причём произвольная функция $f(z) \in W$ представляется с помощью просуммированного ряда из экспоненциальных одночленов, которые принадлежат W. Область G типа полуплоскости или типа полуполосы обладает тем характерным свойством, что имеется направление l, удовлетворяющее условию: если $z_0 \in G$, то и весь луч, выходящий из z_0 в направлении l, принадлежит G. Имеется пример бесконечной области G без этого свойства и пример подпространства W, когда W—не тривиальное инвариантное подпространство и W не содержит в себе экспонент. И.Ф. Красичков-Терновский доказал, что какова бы ни была ограниченная выпуклая область, всегда существует не тривиальное инвариантное подпространство, не допускающее спектрального синтеза. Он же показал, при каких дополнительных условиях спектральный синтез имеет место. Задача сводится к эквивалентной задаче о замкнутых подмодулях в топологическом модуле целых функций экспоненциального типа.

Литература

1. А.Ф. Леонтьев, Успехи матем. наук **24** (1969), 97–164.
2. ———, *Об одном свойстве единственности*, Мат. сборник **72** (1967), 237–249.
3. ———, Известия АН СССР, сер. матем. **36** (1972), 1282–1295.

4. А.Ф. Леонтьев, Мат. сборник **81** (1970), 552–579.
5. ———, Известия АН СССР, сер. матем. **35** (1971), 125–153.
6. И.Ф. Красичков-Терновский, Функц. анализ и его применения **7** (1973), 38–43.
7. А.Ф. Леонтьев, Известия АН СССР, сер. матем. **37** (1973), 577–592.
8. ———, Мат. сборник **85** (1971), 563–580.
9. D. G. Dickson, Trans. Amer. Math. Soc. **110** (1964), 361–374.
10. А.Ф. Леонтьев, Мат. сборник **70** (1966), 132–144.
11. ———, Труды матем. ин-та им. В.А. Стеклова **112** (1971), 125–153.
12. И.Ф. Красичков-Терновский, Мат. сборник **87** (1972), 459–489.
13. ———, Мат. сборник **88** (1972), 3–30.
14. ———, Мат. сборник **88** (1972), 331–352.

Башкирский Филиал АН СССР
Уфа, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Classification of Kleinian Groups

Bernard Maskit*

We present here a complete classification of those Kleinian groups which have an invariant region of discontinuity and which, in their action on hyperbolic 3-space, have a finite-sided fundamental polyhedron. This classification is complete in the same sense that finitely-generated Fuchsian groups of the first kind are completely classified, i.e., there is a countable collection of topologically distinct classes, each such class can be described by a finite set of numbers called the signature; all the groups belonging to any one topological class appear (infinitely often) in the deformation space (defined using quasi-conformal mappings) of any one group in the class; this deformation space can be parametrized as a complex manifold.

Our results can also be regarded as a classification of all uniformizations of any finite Riemann surface (i.e., a closed Riemann surface from which a finite number of points have been deleted), where the uniformizing group has a finite-sided fundamental polyhedron (see §9).

The proofs of the theorems are based on the combination theorems [12], the planarity theorem [13], Bers' technique of variation of parameters using quasi-conformal mappings [4], and Marden's isomorphism theorem [11]. Details will appear elsewhere.

1. A Kleinian group is a discrete subgroup of PSL(2; C) which acts discontinuously at some point of $\hat{C} = C \cup \{\infty\}$. The set of points at which G acts discontinuously is denoted by $\Omega = \Omega(G)$; its complement, the limit set, is denoted by $\Lambda = \Lambda(G)$.

The components of Ω are called *components* of G. G is a function group if there is a component Δ which is kept invariant by G.

*Research supported in part by NSF grant P019972000.

From here on we consider only Kleinian groups which have a finite-sided fundamental polyhedron, and which are function groups, with invariant component $\Delta = \Delta(G)$. Under these conditions, Ahlfors showed that $\Lambda(G)$ has 0 measure [2], and that $\Delta'/G = (\Delta - \{\text{elliptic fixed points}\})/G$ is a finite Riemann surface [1].

2. A parabolic element $g \in G$ is called *accidental* if there is a conformal map $f : \Delta(G) \to \Delta(G^*)$, where f conjugates G into G^*, but $f \circ g \circ f^{-1}$ is loxodromic.

3. Combining the criterion of Beardon and Maskit [3] with the results in [14] and [15] (see also Bers [5] and Kra and Maskit [10]), one sees that if Δ is simply-connected and if G contains no accidental parabolic elements, then G is either elementary (i.e., Λ contains at most one point), or quasi-Fuchsian (i.e., G is a perhaps trivial quasi-conformal deformation of a Fuchsian group). In either case, the action of G on Δ can be topologically completely described by a collection of "integers" called the signature $(g, n; \nu_1, \cdots, \nu_n)$ as follows:

Δ'/G is a surface of genus g with n distinct points removed; lifting a small loop about the ith puncture yields an elliptic or parabolic element of order ν_i.

For topological reasons, the signatures $(0,1; \nu)$ and $(0, 2; \nu_1, \nu_2)$, $\nu_1 \neq \nu_2$, cannot occur. All other signatures with $g \geqq 0, n \geqq 0, 2 \leqq \nu_i \leqq \infty$, do occur.

4. Returning to the general case, a subgroup $H \in G$ is a *factor subgroup* if H is a maximal subgroup for which $\Delta(H)$ is simply connected, H contains no accidental parabolic elements, and H contains every parabolic element of G whose fixed point lies in $\Lambda(H)$.

5. Every factor subgroup of G is finitely generated, and up to conjugacy in G, there are only finitely many [15]. Let $H_1, \cdots, H_s$ be a complete list of nonconjugate factor subgroups, and let $(g_i, n_i; \nu_{i1}, \cdots, \nu_{in_i})$ be the signature of H_i; these signatures are called the *factor signatures* of G.

There is a partial pairing P among the ν_{ij} given by intersections of factor subgroups. If H' is a factor subgroup conjugate to H_k, then $J = H_i \cap H'$ is either trivial or is a common maximal elliptic or parabolic cyclic subgroup. In H_i, J corresponds to some ν_{ij}, and after conjugation, J corresponds to some ν_{kl} in H_k. This intersection pairs ν_{ij} with ν_{kl}, and of course $\nu_{ij} = \nu_{kl}$. (The partial pairing P can be regarded as a symmetric incidence matrix where there is at most one 1 in any row.)

Using the combination theorems, we can construct a minimal subgroup G_0 having $H_1, \cdots, H_s$ as factor subgroups, and with partial pairing P; G is then constructed from G_0 as the free product of G_0 with a Schottky group of some rank $t \geqq 0$ [15].

The *signature* of G is the collection $\{(g_1, n_1; \nu_{11}, \cdots, \nu_{1n_1}), \cdots, (g_s, n_s; \nu_{s1}, \cdots, \nu_{sn_s}), P, t\}$. The group G does not uniquely determine the signature, for we may permute the H_i, and for fixed i we may permute the ν_{ij}. We identify two signatures if they differ by such a permutation.

6. Not all signatures can actually occur. Certain factor signatures cannot occur

(see §3); the factor signature (0, 0) can occur only if $s = 1$; if there is a factor signature (0, 2; ν_{i1}, ν_{i2}), then ν_{i1} can only be paired with ν_{i2} and then only if $\nu_{i1} < \infty$; if there are two factor signatures (0, 3; 2, 2, ∞), then the parabolic elements cannot be paired with each other. A signature satisfying these rules is called admissible.

THEOREM 1 [14], [16]. *Every admissible signature is the signature of a Kleinian group.*

7. There is a technique due to Bers [4] which shows that once we have a group G realizing a certain signature, then the conformal structure on Δ'/G can be chosen at will, i.e., there is a quasi-conformal homeomorphism $w : \hat{\mathbf{C}} \to \hat{\mathbf{C}}$, where $w \circ G \circ w^{-1} = G^*$ is also a Kleinian group (i.e., G^* is a quasi-conformal deformation of G), so that $\Delta(G^*)/G^*$ has a preassigned conformal structure.

THEOREM 2. *G and G^* have the same signature if and only if G^* is a quasi-conformal deformation of G.*

Proof of this theorem will appear elsewhere.

8. Bers [6] defined the deformation space $T(G)$ as the space of isomorphism classes of normalized quasi-conformal deformations of G. Kra [9] proved that $T = T_1 \times T_2$, where T_1 is supported in Δ and T_2 is supported in $\Omega - \Delta$. T_2 is a product of Teichmüller spaces, hence (Bers [7]) a domain in some $\mathbf{C}^n$. T_2 is a complex mainfold [17] which in special cases is known to be a domain in some $\mathbf{C}^n$; one expects that this is generally true. One can factor the deformation space by a discontinuous group to obtain the space of Kleinian groups of the same signature; Bers [8] showed that it is a normal complex space.

9. In general, a uniformization of a finite Riemann surface X' is a Kleinian group G (not necessarily a function group), where X' is conformally a connected component of Ω/G. Our results can be viewed as a classification of all uniformizations of X' by function groups with a finite-sided fundamental polyhedron, where $X' = \Delta'/G$ (each such uniformization can be described by an admissible signature, every admissible signature comes from such a uniformization, and every other such uniformization corresponding to this signature is a quasi-conformal deformation, of the first, supported on $\Omega - \Delta$).

The signature describes the regular covering $p : \Delta' \to X'$ and describes which elements of G are accidental parabolic. This information can also be given by a set of branch numbers for the deleted points of X', together with a set $w_1, \cdots, w_n$ of simple disjoint loops, and "integers" $\alpha_1, \cdots, \alpha_n$, $1 \leqq \alpha_i \leqq \infty$ [13], [14]. These loops divide X' into subsurfaces $Y_1, \cdots, Y_s$; the factor subgroups are precisely the stabilizing subgroups of the connected components of the sets $p^{-1}(Y_i)$. Lifting a loop w_i yields an elliptic or accidental parabolic cyclic subgroup of order α_i; if $\alpha_i > 1$ then this cyclic subgroup is the intersection of two factor subgroups, and every nontrivial intersection of factor subgroups can be obtained in this manner.

References

1. L. V. Ahlfors, *Finitely generated Kleinian groups*, Amer. J. Math. **86** (1964), 413–429. MR **29** #4890.

2. ———, *Fundamental polyhedrons and limit point sets of Kleinian groups*, Proc. Nat. Acad. Sci. U.S.A. **55** (1966), 251–254. MR **31** #3175.

3. A. Beardon and B. Maskit, *Limit points of Kleinian groups and finite-sided fundamental polyhedra*, Acta Math. **132** (1974), 1–12.

4. L. Bers, *Uniformization by Beltrami equations*, Comm. Pure Appl. Math. **14** (1961), 215–228. MR **24** #A2022.

5. ———, *On boundaries of Teichmüller spaces and on Kleinian groups*. I, Ann. of Math. (2) **91** (1970), 570–600. MR **45** #7044.

6. ———, *Spaces of Kleinian groups*, Several Complex Variables, I (Proc. Conf., Univ. of Maryland, College Park, Md., 1970), Springer, Berlin, 1970, pp. 9–34. MR **42** #6216.

7. ———, *On moduli of Riemann surfaces*, mimeographed Lecture Notes, Eidgenossische Technische Hochschule, Zurich, 1964.

8. ———, *On moduli of Kleinian groups*, Uspehi Mat. Nauk **29** (1974), 86–102.

9. I. Kra, *On spaces of Kleinian groups*, Comment. Math. Helv. **47** (1972), 53–69. MR **46** #5611.

10. I. Kra and B. Maskit, *Involutions on Kleinian groups*, Bull. Amer. Math. Soc. **78** (1972), 801–805. MR **46** #348.

11. A. Marden, *The geometry of finitely generated Kleinian groups*, Ann. of Math. (2) **99** (1974), 383–462.

12. B. Maskit, *On Klein's combination theorem*. III, Advances in the Theory of Riemann Surfaces, Ann. of Math. Studies, no. 66, Princeton Univ. Press, Princeton, N. J., 1971, pp. 297–316. MR **44** #6955.

13. ———, *A theorem on planar covering surfaces with applications to* 3*-manifolds*, Ann. of Math. (2) **81** (1965), 341–355. MR **30** #2472.

14. ———, *On boundaries of Teichmüller spaces and on Kleinian groups*. II, Ann. of Math. (2) **91** (1970), 607–639. MR **45** #7045.

15. ———, *Decomposition of certain Kleinian groups*, Acta Math. **130** (1973), 243–263.

16. ———, *Construction of Kleinian groups*, Proc. Conf. Complex Analysis (Minneapolis, 1964), Springer, Berlin, 1965, pp. 281–296. MR **31** #2963.

17. ———, *Self-maps on Kleinian groups*, Amer. J. Math. **93** (1971), 840–856. MR **45** #544.

SUNY AT STONY BROOK
STONY BROOK, NEW YORK 11794, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Intrinsic Metrics on Teichmüller Space

H. L. Royden

1. My purpose here is to describe some recent results about metrics on the Teichmüller space T_g, that is on the space of moduli of compact Riemann surfaces of a given genus $g > 1$. If W is a differentiable surface of genus g and $\mathcal{S}$ the family of all conformal structures on W, then two structures in $\mathcal{S}$ are said to be *equivalent* if there is a conformal mapping φ from W with one structure onto W with the other structure such that φ is homotopic to the identity map on W. The Teichmüller space T_g is the set $\mathcal{S}$ of conformal structures modulo this equivalence. The space T_g carries a natural complex analytic structure which makes it a $(3g - 3)$-dimensional complex manifold which can be embedded in $\boldsymbol{C}^{3g-3}$ as a bounded domain of holomorphy [6]. It is homeomorphic to a ball.

A conformal structure on W is usually given by specifying a collection of local uniformizing variables on W, i.e., a collection $\{z_\alpha\}$ of homeomorphisms of open sets on W into $\boldsymbol{C}$ such that $z_\alpha \circ z_\beta^{-1}$ is a holomorphic function wherever defined.

Let W_0 be W with a fixed conformal structure. A Beltrami differential μ on W_0 is an entity which changes under holomorphic coordinate changes in such a way that $\mu d\bar{z}/dz$ is invariant. In modern terminology μ is a differential on W_0 of type (0, 1) with values in the (holomorphic) tangent bundle of W_0. The absolute value $|\mu|$ of a Beltrami differential is well defined independently of coordinates, and we make the space of Beltrami differentials into a normed space by using the L^∞ norm $\|\mu\|_\infty = \sup_W |\mu|$. Given a Beltrami differential μ with $\|\mu\| < 1$, we can define a new conformal structure by taking as local uniformizing variables functions w which satisfy the equation $\partial w/\partial \bar{z} = \mu \partial w/\partial z$. The Riemann surface obtained from μ in this way is denoted by W_μ, and the Beltrami differential μ is just the complex dilation of the identity map of W_0 into W_μ, that is $\mu = (\partial w/\partial \bar{z})/(\partial w/\partial z)$, where z and w are local uniformizers on W_0 and W_μ, respectively. Conversely, given a con-

formal structure making W into a Riemann surface W', the complex dilation μ of the identity map from W_0 onto W' is a Beltrami differential on W_0 and $W_\mu = W'$. We say that two Beltrami differentials μ and ν are equivalent if W_μ and W_ν are equivalent, i.e., if there is a conformal map φ of W_μ onto W_ν which is homotopic to the identity. Thus T_g is the unit ball in the space of Beltrami differentials modulo this equivalence, and a point $x \in T_g$ is a class $[W_\mu]$ of equivalent structures on W. The space M of Beltrami differentials is a normed linear space and thus has a natural topological and complex analytic structure. We take the quotient topology which T_g inherits from M for the topology of T_g, and will see that T_g also inherits a complex analytic structure from M.

The preceding representation of T_g depends on the base point W_0 chosen, but if we take a different conformal structure W_μ for our base point, then we have a natural map between the set M_0 of Beltrami differentials of norm less than one on W_0 and the set M_μ of those on W_μ. This map is given by

$$\nu \frac{\overline{dz}}{dz} \to \frac{\nu - \mu}{1 - \nu\bar{\mu}} \frac{\partial w/\partial z}{\overline{\partial w/\partial z}} \frac{\overline{dw}}{dw}.$$

Thus to consider local properties of T_g in a neighborhood of a point $x \in T_g$, we may always use a representation in which the base structure is an element of the equivalence class x.

2. The tangent and cotangent spaces for T_g. A flow on W is a family φ_t of diffeomorphisms of W onto itself depending smoothly on a parameter t. If φ_0 is the identity, then in terms of a local uniformizer z we have $\varphi_t(z) = z + th(z) + O(t^2)$ where h is a vector field on W_0. If we let μ_t be the complex dilation of φ_t, then $\mu_t = t\partial h/\partial \bar{z} + O(t^2)$. Thus we say that a Beltrami differential μ is infinitesimally trivial if there is a tangent vector field on W_0 such that $\mu = \partial h/\partial \bar{z}$. If μ is a Beltrami differential and η a quadratic differential on W_0, then $\mu\eta$ is a form of type (1, 1), and hence $\int_{W_0} \mu\eta$ is well defined. It can be shown that μ is infinitesimally trivial if and only if $\int \mu\eta = 0$ for all holomorphic quadratic differentials η on W_0.

Since the tangent space at 0 to the space M of Beltrami differentials is just M itself, we see that the tangent space at $[W_0]$ to T_g is the space of Beltrami differentials on W_0 modulo the infinitesimally trivial ones. Since $\int \mu\eta = 0$ for all quadratic differentials η if and only if μ is infinitesimally trivial, it follows that the space Q of quadratic differentials is the cotangent space to T_g at W_0.

3. Differential metrics on T_g. A differential metric on a manifold V is given by specifying a real-valued function $F = F(x, \xi)$ defined for points $\langle x, \xi \rangle$ in the tangent bundle of V ($x \in V$, $\xi \in (TV)_x$), such that $F(x, \xi) > 0$ for $\xi \neq 0$ and $F(x, \alpha\xi) = |\alpha| F(x, \xi)$. Arc length for a curve $x(t)$ is defined by $l = \int F(x(t), \dot{x}(t))\, dt$, and the distance between any two points is defined to be the infimum of the lengths of curves joining them. The metric is Hermitian if for each x the form $F(x, \xi)$ is a Hermitian form on $(TV)_x$.

We can also define a differential metric by giving a form $G(x, \eta)$ on the cotangent bundle of V with $G(x, \eta) > 0$ for $\eta \neq 0$ and $G(x, \alpha\eta) = |\alpha| G(x, \eta)$. We define F

by duality as $F(x, \xi) = \sup \{[\xi, \eta]: G(x, \eta) = 1\}$, and proceed as before.

4. The Teichmüller metric and its properties. Since the cotangent space to T_g at $[W_\mu]$ is the space Q_μ of holomorphic quadratic differentials on W_μ, each differential metric on T_g is obtained by assigning a norm in each of the linear spaces Q_μ. One natural norm to take is the L_1 norm $\|\eta\| = \int_{W_\mu} |\eta|$. If we take this norm for the form $G(\mu, \eta)$ on the cotangent bundle, we obtain the infinitesimal form of the Teichmüller metric.

If we use this to define a norm in the tangent space at W_0, then it is readily seen that for this norm $\|\mu\|_T = \inf \|\nu\|_\infty$, as ν ranges over all Beltrami differentials with $\mu - \nu$ infinitesimally trivial.

If η is a holomorphic quadratic differential on W_0 and k a constant less than 1, then $\mu = k\bar{\eta}/|\eta|$ is a Beltrami differential, and such a Beltrami differential is called a Teichmüller differential. It is known [1], [5] that every Beltrami differential ν with $\|\nu\|_\infty < 1$ is equivalent to a unique Teichmüller differential μ, and $\|\mu\|_\infty \leqq \|\nu\|_\infty$ for every ν equivalent to μ with equality only if $\mu = \nu$. It follows from this that for a Teichmüller differential μ the mapping $t \to t\mu$ is a geodesic in the Teichmüller metric and that every geodesic has this form.

The norm $G(x, \eta) = \int |\eta|$ on the space Q of holomorphic quadratic differentials is not smooth, but has first derivatives which satisfy a Hölder condition with exponent depending on the order of the largest zero of η. This fact can be used to show the following [9]: If $\Phi: Q_W \to Q_{W'}$ is a linear isometry, then there is a conformal map φ: $W' \to W$ such that $\Phi = \alpha\varphi^*$ with $|\alpha| = 1$. From this it follows that every isometry of T_g with the Teichmüller metric maps a class $[W_\mu]$ of conformal structures into a class $[W_\nu]$ with W_ν conformally equivalent to W_μ, i.e., there exists a conformal mapping between them, not necessarily homotopic to the identity.

The set of orientation preserving diffeomorphisms of W onto itself is a group, and those homotopic to the identity form a normal subgroup. The quotient group is called the mapping class group or the Teichmüller modular group. This group acts on T_g as follows: Given a diffeomorphism φ, we consider φ a map from W_0 to W_μ. Define $\varphi\mu$ to be the complex dilation of this map. Then the equivalence class $[\varphi\mu]$ of $\varphi\mu$ depends only on the mapping class of φ and the equivalence class of μ. Thus we have an action of the Teichmüller modular group on T_g, and it is not difficult to show that the action of $[\varphi]$ is a biholomorphic map of T_g onto itself isometric with respect to the Teichmüller metric. Our previous statement implies that every isometry arises from the action of the Teichmüller modular group.

Kobayashi [7], [8] has introduced an invariant pseudo-metric on each complex manifold V which can be characterized as the largest pseudo-metric on V such that each holomorphic map φ of the unit disk (with the Poincaré metric) into V is distance decreasing with this metric. This pseudo-metric is, by its definition, invariant under biholomorphic self maps of V. From the characterization of the geodesics in the Teichmüller metric it can be shown [9] that the Teichmüller metric on T_g is the infinitesimal form [10] of the Kobayashi metric for T_g. Thus the Teichmül-

ler metric of T_g is invariant under biholomorphic self maps of T_g, and we see that the biholomorphic self maps of T_g are the actions of the Teichmüller modular group.

5. The Bers embedding of T_g. If $\mu_1, \cdots, \mu_N$, $N = 3g - 3$, are any set of Beltrami differentials on W_0 which are independent modulo the infinitesimally trivial ones, we obtain local holomorphic ooardinates for a neighborhood of W_0 in T_g by assigning $\langle t_1, \cdots, t_N \rangle$ to $[W_\mu]$ with $\mu = \sum t_i \mu_i$. If η is a holomorphic quadratic differential on W_0 and $ds^2 = \lambda^2 |dz|^2$ is the Poincaré metric on W_0, then $\lambda^{-2}\bar{\eta}$ is a Beltrami differential, and one of this form is called a Bers differential. By using the $3g - 3$ independent holomorphic quadratic differentials to get our μ_i, we obtain a special set of holomorphic coordinates at $[W_0]$ known as Bers coordinates. It turns out [4], [5] that these coordinates can be extended to a global set of coordinates on T_g and that they then embed T_g as a bounded domain of holomorphy in $\boldsymbol{C}^{3g-3}$.

6. The Weil-Petersson metric. To prescribe a Hermitian metric on T_g we need to define a Hermitian form on the cotangent space at each point, i.e., on the space of holomorphic quadratic differentials on each surface. If $ds^2 = \lambda^2 |dz|^2$ is the Poincaré non-Euclidean distance on a surface W_0, we can define such a Hermitian form by setting $(\eta_1, \eta_2) = \int_{W_0} \lambda^{-2} \eta_1 \bar{\eta}_2$. This metric is the Weil-Petersson metric, and is readily seen to be invariant under biholomorphic self maps of T_g onto itself.

Given a Hermitian metric $ds^2 = g_{\alpha\bar{\beta}}\, dz^\alpha\, d\bar{z}^\beta$ on a complex manifold, we say that it is Kähler if $\partial g_{\alpha\bar{\beta}}/\partial z^\gamma = \partial g_{\gamma\bar{\beta}}/\partial z^\alpha$. A metric is Kähler if and only if we can find normal coordinates at each point, i.e., holomorphic coordinates such that at the point the metric tensor has the development $g_{\alpha\bar{\beta}} = \delta_{\alpha\bar{\beta}} + O(|z|^2)$. If the metric is Kähler and real analytic, one can introduce a set of canonical coordinates at a point which are characterized by the property that the power series for $g_{\alpha\bar{\beta}}$ contains no terms which are products only of unbarred variables (or only of barred variables). Ahlfors [2] has shown that the Bers coordinates are normal coordinates for the Weil-Petersson metric and hence that it is Kähler. I have recently extended this result to show that the Bers coordinates are canonical.

In terms of canonical coordinates,

$$g_{\alpha\bar{\beta}} = \delta_{\alpha\bar{\beta}} - \tfrac{1}{2} r_{\alpha\bar{\beta}\gamma\bar{\delta}} z^\gamma \bar{z}^\delta + O(|z|^3)$$

where $r_{\alpha\bar{\beta}\gamma\bar{\delta}}$ is the Riemann curvature tensor. If ξ^α and η^α are unit tangent vectors, the holomorphic bisectional curvature in the directions ξ, η is given by $K_{\xi\eta} = r_{\alpha\bar{\beta}\gamma\bar{\delta}} \xi^\alpha \bar{\xi}^\beta \eta^\gamma \bar{\eta}^\delta$ and the holomorphic sectional curvature in the direction ξ is given by

$$K_{\xi\xi} = r_{\alpha\bar{\beta}\gamma\bar{\delta}}\, \xi^\alpha \bar{\xi}^\beta \xi^\gamma \bar{\xi}^\delta.$$

Ahlfors [3] has shown that the holomorphic bisectional curvature of the Weil-Petersson metric is nonpositive and the sectional curvature is negative. I have recently established the bound $K_{\xi\xi} < -(55/18)[\pi(g-1)]^{-1}$. There is heuristic evidence that 55/18 should be replaced by 4.

Bibliography

1. L. V. Ahlfors, *On quasiconformal mappings*, J. Analyse Math. **3** (1953), 1–58; correction, ibid. **4** (1954), 207–208. MR **16**, 348.

2. ———, *Some remarks on Teichmüller's space of Riemann surfaces*, Ann. of Math. (2) **74** (1961), 171–191. MR **14** #4480.

3. ———, *Curvature properties of Teichmüller's space*, J. Analyse Math. **9** (1961/62), 161–176. MR **25** #192.

4. L. V. Ahlfors and G. Weil, *A uniqueness theorem for Beltrami equations*, Proc. Amer. Math. Soc. **13** (1962), 975–978. MR **26** #6393.

5. L. Bers, *Quasiconformal mappings and Teichmüller's theorem*, Analytic Functions, Princeton Univ. Press, Princeton, N. J., 1960, pp. 89–119. MR **22** #5716.

6. ———, *Spaces of Riemann surfaces as bounded domains*, Bull. Amer. Math. Soc. **66** (1960), 98–103; correction, ibid. **67** (1961), 465–466. MR **22** #2695; **24** #A826.

7. S. Kobayashi, *Invariant distances on complex manifolds and holomorphic mappings*, J. Math. Soc. Japan **19** (1967), 460–489. MR **13** #736.

8. ———, *Hyperbolic manifolds and holomorphic mappings*, Pure and Appl. Math., 2, Dekker, New York, 1970. MR **43** #3503.

9. H. L. Royden, *Automorphisms and isometries of Teichmüller space*, Advances in the Theory of Riemann Surfaces (Proc. Conf., Stony Brook, N.Y., 1969), Ann. of Math. Studies, no. 66, Princeton Univ. Press, Princeton, N. J., 1971, pp. 369–383. MR **44** #5452.

10. ———, *Remarks on the Kobayashi metric*, Several Complex Variables. II (Proc. Internat. Conf., Univ. of Maryland, College Park, Md., 1970), Lecture Notes in Math., vol. 185, Springer-Verlag, Berlin and New York, 1971, pp. 125–137. MR **46** #3826.

11. ———, *Invariant metrics on Teichmüller space*, Contributions to Analysis, Academic Press, New York and London, 1974.

STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

On Quadratic Differentials and Extremal Quasi-Conformal Mappings*

Kurt Strebel

1. Extremal quasi-conformal mappings and Teichmueller mappings. A regular quasi-conformal mapping of a plane domain G onto a domain G' (more generally of a Riemann surface R onto a surface R') is an orientation preserving homeomorphism f with continuous partial derivatives which, except for isolated points, has a nonvanishing Jacobian and takes infinitesimal circles into infinitesimal ellipses of bounded dilatation. To fix the notation, we put $w = f(z)$ and write the differential in complex form: $dw = p(z)dz + q(z)d\bar{z}$, with $p = f_z$, $q = f_{\bar{z}}$ and $|p| > |q| \geqq 0$. Then the dilatation at the point z becomes

$$D(z) = \frac{|p(z)| + |q(z)|}{|p(z)| - |q(z)|} = \frac{1 + |\kappa(z)|}{1 - |\kappa(z)|},$$

with $\kappa(z) = q(z)/p(z)$ the complex dilatation of f. The supremum $K = \sup_{z \in G} D(z)$ is called the maximal dilatation of f; evidently $K = (1 + k)/(1 - k)$ with $k = \sup_{z \in G} |\kappa(z)| < 1$. Any quasi-conformal mapping with a maximal dilatation $K \leqq Q$ is called Q-quasi-conformal.

The first to consider this class of mappings was Herbert Groetzsch in 1928. He also solved the following extremal problem: Given two quadrilaterals S and S', i.e., Jordan domains with four assigned boundary points, map S quasi-conformally onto S' with smallest maximal dilatation K. Such a mapping is called extremal quasi-conformal (möglichst konform). The solution is of the form $f = \Psi^{-1} \circ F \circ \Phi$, where Φ and Ψ are conformal mappings of S, respectively S', onto horizontal rectangles and F is a horizontal stretching with the factor K. For every other mapping $\tilde{f}$ which preserves the vertices we have $\tilde{K} > K$, i.e., f is even unique extremal.

* An extended version of this article will be published elsewhere.

If we introduce the modulus M of a quadrilateral as one of the ratios of the sides of a conformally equivalent rectangle, the above result shows that for any Q-quasi-conformal mapping we have $Q^{-1}M \leqq M' \leqq QM$. It was an extremely fruitful idea to replace the original definition of quasi-conformality by just this property: Any homeomorphism f which satisfies the above inequality is called Q-quasi-conformal.

One of the big gains by introducing this larger family is normality or even compactness. To fix the ideas, let $\tilde{f}$ be a quasi-conformal mapping of the disk $D: |z| < 1$ onto $D': |w| < 1$. In the class of all quasi-conformal mappings of D onto D' which agree with $\tilde{f}$ on the circumference $\Gamma: |z| = 1$, there exists an extremal quasi-conformal mapping f. However, it is in general not unique [4].

In order to get uniqueness of the solution, we go back to the example of Groetzsch. The complex dilatation of this mapping is

$$\kappa(z) = k\bar{\Phi}'^2/|\Phi'|^2 = k\,\bar{\varphi}/|\varphi|,$$

with $k = (K-1)/(K+1)$ and $\varphi = \Phi'^2$. Under a change of the variable, the function φ evidently transforms like a quadratic differential. For an arbitrary holomorphic (or meromorphic) quadratic differential φ and positive number k, $0 < k < 1$, we can form $\kappa = k\bar{\varphi}/|\varphi|$. A quasi-conformal mapping f with this complex dilatation is called a Teichmueller mapping associated with the quadratic differential φ. It is locally, except for the zeroes and poles of φ, of the form "conformal $\circ$ affine $\circ$ conformal". It is known ([5], [3]) that a Teichmueller mapping of the disk, more generally of a compact bordered Riemann surface, which is associated with a quadratic differential φ of finite norm $\|\varphi\| = \iint |\varphi|(z)|\,dx\,dy$, is unique extremal. But now we have lost the former general existence theorem. In order to get existence *and* uniqueness, one either has to impose certain conditions on the general solution (without losing the existence theorem) or else prove the existence of Teichmueller mappings. It is the purpose of this article to do the latter.

2. Quasi-conformal mappings with equal boundary values. Following [3], let f and $\tilde{f}$ be quasi-conformal mappings of D onto D' which agree on the circumference Γ. Set $f_1 = \tilde{f}^{-1}$. Let φ be an arbitrary holomorphic quadratic differential in D with norm $\|\varphi\| = 1$. Any (noncritical) vertical trajectory β of φ is mapped by the composition $f_1 \circ f$ onto an arc with the same endpoints on Γ. Therefore the length inequality

$$\int_{\beta} |\varphi(z)|^{1/2}\,|dz| \leqq \int_{f_1 \circ f(\beta)} |\varphi(\tilde{z})|^{1/2}\,|d\tilde{z}|$$

holds. We now decompose D into open vertical strips S with respect to φ. Each S is mapped by any branch of $\Phi(z) = \int (\varphi(z))^{1/2}\,dz$ onto a domain S^* swept out by vertical straight segments. We introduce $\zeta = \xi + i\eta = \Phi(z)$ as a conformal parameter in S. With $d\tilde{z} = p_1 dw + q_1 d\bar{w}$ the differential of f_1, we get

$$\int d\eta \leqq \int \frac{|\varphi(\tilde{z})|^{1/2}}{|\varphi(z)|^{1/2}} \left| p_1 p + q_1 \bar{q} - (p_1 q + q_1 \bar{p}) \frac{\varphi}{|\varphi|} \right| d\eta$$

where the integral is taken over the vertical segment in S^* which corresponds to β. Integration over ξ, transition to S by the mapping $\zeta = \Phi(z)$ and then summation over the strips of the exhaustion gives

$$1 \leqq \iint_D |\varphi(\tilde{z})|^{1/2} |\varphi(z)|^{1/2} \left| p_1 p + q_1 \bar{q} - (p_1 q + q_1 \bar{p}) \frac{\varphi}{|\varphi|} \right| dx\, dy.$$

An application of the Schwarz inequality and subsequent arranging of terms finally yields the result:

$$1 \leqq \iint_D |\varphi(z)| \frac{|1 - \kappa\varphi/|\varphi||^2}{1 - |\kappa|^2} \frac{|1 - \kappa_1(\bar{p}/p)(\varphi/|\varphi|)(1 - \bar{\kappa}\bar{\varphi}/|\varphi|)/(1 - \kappa\varphi/|\varphi|)|^2}{1 - |\kappa_1|^2} dx\, dy$$

for every holomorphic quadratic differential φ in D with norm $\|\varphi\| = 1$.

The inequality contains the uniqueness theorem I mentioned: A Teichmueller mapping f associated with a quadratic differential φ of finite norm is unique extremal. This simply follows by putting $\kappa = k\bar{\varphi}/|\varphi|$ ($\|\varphi\| = 1$) into the inequality and looking at the right-hand factor of the integrand.

On the other hand, an easy estimation of this factor leads to

$$\frac{1}{K_1} = \frac{1}{\tilde{K}} \leqq \iint_D |\varphi('z)| \frac{|1 - \kappa\varphi/|\varphi||^2}{1 - |\kappa|^2} dx\, dy$$

from which it is not difficult to get the following sufficient condition for extremality: If there is a sequence of holomorphic quadratic differentials φ_n, $\|\varphi_n\| = 1$, such that

$$\lim_{n\to\infty} \iint_D \varphi_n \kappa\, dx\, dy = k,$$

then κ is extremal (i.e., the mapping f which is determined by κ up to a linear transformation is extremal for its boundary values). It is known ([1], [3]) that this condition is also necessary for extremality. Let us call a sequence (φ_n) with the above limit property a Hamilton sequence for κ. Then the result can be expressed in the following way: A complex dilatation κ is extremal if and only if it admits a Hamilton sequence.

3. Existence of Teichmueller mappings. Let us assume that there exists a Hamilton sequence (φ_n) for κ which converges in norm to some φ, i.e., $\|\varphi_n - \varphi\| \to 0$. Then φ is holomorphic, has norm one and $\iint_D \varphi\kappa\, dx\, dy = k$; hence $\kappa = k\bar{\varphi}/|\varphi|$. The mapping f is thus aTeichmueller mapping associated with φ. For any sequence (φ_n) we can of course take a subsequence which converges locally uniformly in D to some φ, $0 \leqq \|\varphi\| \leqq 1$. If we can give conditions under which $\|\varphi\| < 1$ is impossible, we have the desired existence theorem.

Let $\varphi = \lim_{n\to\infty} \varphi_n = 0$, and let $\tilde{f}$ be a quasi-conformal mapping on an annulus $\tilde{r} < |z| < 1$ which agrees with f on Γ (but not necessarily on $|z| = \tilde{r}$). We choose r_0, $0 < r_0 < 1$, such that the composition $f_1 \circ f$, $f_1 = \tilde{f}^{-1}$, makes sense. It maps the circle $|z| = r_0$ onto some Jordan curve Γ_0 in D. We apply the earlier procedure to the mapping $f_1 \circ f$ in the annulus $r_0 < |z| < 1$ and to the quadratic differentials φ_n. A vertical trajectory which has one or both of its endpoints on $|z| = r_0$ now is transformed into a curve with one or both of its endpoints on Γ_0. But as the φ_n's

go to zero, the length inequality will, for these arcs, eventually hold up to some arbitrarily small term. Moreover, the circle $|z| = r_0$ itself will become arbitrarily short. With this in mind, we get

$$\frac{1}{\tilde{K}} - \varepsilon \leqq \iint\limits_{r_0<|z|<1} |\varphi_n(z)| \frac{|1 - \kappa\varphi_n/|\varphi_n||^2}{1 - |\kappa|^2}\, dx\, dy$$

for arbitrarily small ε and all sufficiently large n. Using the fact that

$$\lim_{n\to\infty} \iint\limits_{r_0<|z|<1} \varphi_n \kappa\, dx\, dy = k$$

and letting $\varepsilon \to 0$ we finally get $k \leqq \tilde{k}$. We have proved

THEOREM 1. *If there is a Hamilton sequence (φ_n) for κ which tends to zero locally uniformly in D (degenerating Hamilton sequence), then f is extremal compared with any quasi-conformal mapping $\tilde{f}$ of an arbitrarily small annulus $\tilde{r} < |z| < 1$ which agrees with f on Γ: $|z| = 1$.*

The case $0 < \|\varphi\| < 1$ is reduced to the above by choosing $\tilde{\varphi}_n = (\varphi_n - \varphi)/\|\varphi_n - \varphi\|$. This is a degenerating Hamilton sequence for κ, as is easily seen. We can therefore say: If κ admits a Hamilton sequence (φ_n) which tends locally uniformly to φ, $0 < \|\varphi\| < 1$, then it also admits a degenerating Hamilton sequence.

Let now f with complex dilatation κ be an extremal mapping for the given boundary values. If $k = 0$, it is conformal, hence unique extremal. So let $k > 0$. If the boundary homeomorphism admits a quasi-conformal extension $\tilde{f}$ into an annulus $\tilde{r} < |z| < 1$ with $\tilde{k} < k$, any Hamilton sequence for κ must converge in norm. Hence f is a Teichmueller mapping. We have proved

THEOREM 2. *Given a homeomorphism of Γ: $|z| = 1$ onto Γ': $|w| = 1$ which admits a quasi-conformal extension $\tilde{f}$ into an arbitrarily small annulus $\tilde{r} < |z| < 1$ with a maximal dilatation $\tilde{K}$ arbitrarily close to one, then there is a unique extremal quasi-conformal mapping f with the given boundary values. It is a Teichmueller mapping associated with a quadratic differential φ of norm one, and every Hamilton sequence (φ_n) converges in norm to φ.*

The same is true if the maximal dilatation $\tilde{K}$ of the "barrier" $\tilde{f}$ can be chosen smaller than the maximal dilatation K of the extremal mapping f.

A simple example of a boundary correspondence of the first kind is an analytic homeomorphism of Γ onto Γ'; but it is elementary to show that, e.g., a homeomorphism with a continuous second derivative and a first derivative bounded away from zero also has the extension property.

On the other hand, there can exist degenerating Hamilton sequences for an affine mapping of a bounded domain, as was recently shown by E. Reich in [2], where somewhat related ideas are used to characterize the "substantial" boundary points of a domain. Therefore our condition for the extremal to be of Teichmueller form is only sufficient. But it can of course be generalized to compact bordered Riemann surfaces.

References

1. R.S. Hamilton, *Extremal quasiconformal mappings with prescribed boundary values*, Trans. Amer. Math. Soc. **138** (1969), 399–406. MR **19** #7093.

2. E. Reich, *On the relation between local and global properties of boundary values for extremal quasiconformal mappings*, Ann. of Math. Studies, no. 79, Princeton Univ. Press, Princeton, N. J., 1974, pp. 391–407.

3. E. Reich and K. Strebel, *Extremal quasiconformal mappings with given boundary values*, Contributions to Analysis, Academic Press, New York, 1974, pp. 375–391.

4. K. Strebel, *Zur Frage der Eindeutigkeit extremaler quasikonformer Abbildungen des Einheitskreises*, Comment. Math. Helv. **36** (1961/62), 306–323. MR **29** #1331.

5. ———, *Zur Frage der Eindeutigkeit extremaler quasikonformer Abbildungen des Einheitskreises*. II, Comment. Math. Helv. **39** (1964), 77–89. MR **31** #346.

UNIVERSITY OF ZURICH
ZURICH, SWITZERLAND

Section 13

Partial Differential Equations

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Some Recent Advances in the Multidimensional Parametric Calculus of Variations

William K. Allard

Introduction. Suppose Ω is an open subset of $\boldsymbol{R}^n$, k is an integer and $0 \leqq k \leqq n$. Let $\boldsymbol{G}(n, k)$ be the Grassmann manifold of k-planes through 0 in $\boldsymbol{R}^n$. Suppose Φ is a *positive k-dimensional parametric integrand on* Ω; by this we mean that Φ maps the product $\Omega \times \boldsymbol{G}(n, k)$ smoothly into the positive real numbers. Let R be a k-dimensional surface in Ω (we will be precise momentarily). If x is a regular point of R let $R_x \in \boldsymbol{G}(n, k)$ be the tangent plane to R at x. We study the existence and regularity of surfaces R satisfying conditions on how the integral $\int_R \Phi(x, R_x)\, dH^k x$ changes when R is replaced by an image of R; here H^k is the k-dimensional Hausdorff measure on Ω, which we define and discuss below. Our treatment of this problem and problems closely related to it is by no means complete. The theorems discussed in this article are treated in detail in the works listed in the References, among which is Herbert Federer's treatise *Geometric measure theory* which comprehensively exposes the geometric approach with which we attack the problem.

Hausdorff measure, surfaces. Let $\alpha(k) = L^k\{x \in \boldsymbol{R}^k : |x| < 1\}$, where L^k is the Lebesgue measure on $\boldsymbol{R}^k$. Given $A \subset \Omega$ and $\delta > 0$, let

$$\varphi_\delta(A) = \inf\left\{\sum_{i=1}^{\infty} \alpha(k)(\operatorname{diam} A_i/2)^k\right\}$$

where $A \subset \bigcup_{i=1}^{\infty} A_i \subset \Omega$ and $\operatorname{diam} A_i \leqq \delta$. Let $H^k(A) = \lim_{\delta \downarrow 0} \varphi_\delta(A)$ and call this set function the *k-dimensional Hausdorff measure on* Ω. H^k is Borel regular and $H^n = L^n$. We set $(H^k \mathbin{\lfloor} A)(B) = H^k(A \cap B)$ whenever $A \subset \Omega$, $B \subset \Omega$.

Let $\boldsymbol{M}_k(\Omega)$ be the class of C^1 properly imbedded k-dimensional submanifolds of Ω of locally finite k-area. Thus $M \in \boldsymbol{M}_k(\Omega)$ if $M \subset \Omega$, $H^k \mathbin{\lfloor} M$ is locally finite and, for each $a \in M$, there is $F : \Omega \to \boldsymbol{R}^k$ such that F is C^1, rank $DF(a) = k$ and $M = \{x : F(x) = F(a)\}$ near a. Let $\boldsymbol{R}_k(\Omega)$ be the class of *k-rectifiable* subsets of Ω; by

definition, $R \in \boldsymbol{R}_k(\Omega)$ if $R \subset \Omega$, $H^k \mathbin{\lfloor} R$ is locally finite and, for H^k almost all $a \in R$, there is $R_a \in G(n, k)$ such that

$$\lim_{s \to \infty} s^k(H^k \mathbin{\lfloor} \{x : s^{-1}(x - a) \in R\}) = H^k \mathbin{\lfloor} R_a$$

weakly as Radon measures on $\boldsymbol{R}^n$. These are the "surfaces" R of the Introduction and the "regular points" are the points a for which R_a exists. The relationship between $\boldsymbol{M}_k(\Omega)$ and $\boldsymbol{R}_k(\Omega)$ is as follows:

THEOREM (SEE FEDERER). *$R \in \boldsymbol{R}_k(\Omega)$ if and only if $R \subset \Omega$ and, for each $\varepsilon > 0$, there is $M \in \boldsymbol{M}_k(\Omega)$ such that $H^k[(M \sim R) \cup (R \sim M)] < \varepsilon$.*

If $F : \Omega \to \boldsymbol{R}^\nu$ is C^1 we let the *Jacobian* $J_k F$: $\Omega \times \boldsymbol{G}(n, k) \to \{s : 0 \leqq s < \infty\}$ have at $(x, S) \in \Omega \times \boldsymbol{G}(n, k)$ the value $(\det \langle v_i, DF(x)\rangle \cdot \langle v_j, DF(x)\rangle)^{1/2}$, where $v_1, \cdots, v_k \in S$ and $v_i \cdot v_j = \delta_{ij}$. The following mapping formula is basic for our purposes:

THEOREM (SEE FEDERER). *If $R \in \boldsymbol{R}_k(\Omega)$, Γ is an open subset of $\boldsymbol{R}^\nu$ and $F : \Omega \to \Gamma$ is C^1 and proper, then* (a) *$F(R) \in \boldsymbol{R}_k(\Gamma)$*; (b) *for H^k almost all $b \in F(R)$, $F(R)_b = DF(a)(R_a)$ whenever $a \in R \cap F^{-1}\{b\}$*;

$$\text{(c)} \qquad \int_R J_k F(x, R_x)\, dH^k x = \int_\Gamma \operatorname{card}(R \cap F^{-1}\{y\})\, dH^k y.$$

Minimizing conditions and the Euler-Lagrange operator. We now specify two types of "conditions on how the integral changes".

Suppose that Φ is an integrand on Ω; $0 < \delta \leqq \infty$; $\beta : [0, \delta) \to [0, \infty)$, $\beta \in C^\infty$, $\beta(0) = 0$, $\beta' \geqq 0$. We say that $R \in \boldsymbol{R}_k(\Omega)$ is (Φ, δ, β) *minimizing in Ω* if whenever $F : \Omega \to \Omega$ is C^∞, $K = \operatorname{Closure}\{x : F(x) \neq x\}$ is compact in Ω and $r = \operatorname{diam}[K \cup F(K)] \leqq \delta$ we have that

$$\int_{R \cap K} \Phi(x, R_x)\, dH^k x \leqq (1 + \beta(r)) \int_{F(R \cap K)} \Phi(y, F(R)_y)\, dH^k y.$$

Rectifiable set solutions to a wide variety of geometric variational problems with constraints are (Φ, δ, β) minimizing. Surfaces which arise from capillary action, surfaces which minimize an integral avoiding an obstacle and surfaces which solve variational problems with partially free boundary are examples.

Our second type of condition is more infinitesimal. Let $\mathscr{X}(\Omega) = \{g : \Omega \to \boldsymbol{R}^n,$ g is C^∞, spt g is compact$\}$; think of the members of $\mathscr{X}(\Omega)$ as vector fields on Ω. Given $g \in \mathscr{X}(\Omega)$, set $h_t(x) = x + tg(x)$ for $(t, x) \in \boldsymbol{R} \times \Omega$ and let

$$\delta(\Phi; x, S)(g) = \frac{d}{dt} \Phi(h_t(x), Dh_t(x)(S)) J_k h_t(x, S)\Big|_{t=0} \quad \text{for } (x, S) \in \Omega \times \boldsymbol{G}(n, k).$$

Note that $\delta(\Phi; x, S(g))$ depends linearly on g. Given $R \in \boldsymbol{R}(\Omega)$ with $\int_R \Phi(x, R_x)\, dH^k x < \infty$, we calculate

$$\begin{aligned} \frac{d}{dt} \int_{h_t(R)} \Phi(y, h_t(R)_y)\, dH^k y &= \frac{d}{dt} \int_R \Phi(h_t(x), Dh_t(x)(R_x)) J_k h_t(x, S)\, dH^k x\Big|_{t=0} \\ &= \int_R \delta(\Phi; x, R_x)(g)\, dH^k x. \end{aligned}$$

Thus for any $R \in \boldsymbol{R}_k(\Omega)$ we have a linear functional $\delta(\Phi; R) : \mathscr{X}(\Omega) \to \boldsymbol{R}$, the *Euler-Lagrange operator*, whose value at $g \in \mathscr{X}(\Omega)$ is the last term in the above calculation.

If $1 < p \leqq \infty$ and $1 = 1/p + 1/q$, we let

$$\|\delta(\Phi; R)\|_p = \sup\left\{\delta(\Phi; R)(g) : \int_R |g|^q \, dH^k \leqq 1\right\}.$$

For example, if $M \in \boldsymbol{M}_k(\Omega)$, M is smooth, M is closed relative to Ω and H is the mean curvature vector of M then

$$\|\delta(\Phi; R)\|_p = k\left(\int_M |H|^p \, dH^k\right)^{1/p}.$$

The relationship between these two types of conditions is not understood at this time. It obvious, however, that if R is $(\Phi, \delta, 0)$ minimizing then $\delta(\Phi; R) = 0$.

Some theorems about (Φ, δ, β) minimizing sets. We say an integrand Φ on Ω is *elliptic* if there is $c > 0$ such that for all $(a, T) \in \Omega \times \boldsymbol{G}(n, k)$ we have

$$\int_R \Phi(a, R_x) \, dH^k x - \int_D \Phi(a, D_x) \, dH^k x \geqq c[H^k(R) - H^k(D)]$$

where $D = T \cap \{x : |x| < 1\}$ and $R \in \boldsymbol{R}_k(\boldsymbol{R}^n)$ is compact and does not retract onto $T \cap \{x : |x| = 1\}$.

The *area* integrand A on Ω defined by $A(x, S) = 1$ for $(x, S) \in \Omega \times \boldsymbol{G}(n, k)$ is obviously elliptic as are the integrands which are close to it in the C^2 topology. Ellipticity is a natural sufficient condition for the regularity of (Φ, δ, β) minimizing sets as well as for lower semicontinuity theorems for the integral of Φ over a rectifiable chain. Ellipticity is invariant under diffeomorphisms. In case $k = n - 1$, Φ is elliptic if and only if, for each $a \in \Omega$, $\boldsymbol{R}^n \sim \{0\} \ni \xi \to |\xi| \, \Phi(a, \{v : v \cdot \xi = 0\})$ defines a uniformly convex norm on $\boldsymbol{R}^n$.

A basic fact in the theory is the

REGULARITY THEOREM (ALMGREN). *If*

(a) *Φ is elliptic on Ω,*

(b) $0 < \delta \leqq \infty$,

(c) $\int_0^\delta r^{-(1+\alpha)} \beta(r)^{1/2} \, dr < \infty$ *for some α with* $0 \leqq \alpha < 1$,

(d) $R \in \boldsymbol{R}_k(\Omega)$ *is (Φ, δ, β) minimizing in Ω*

then $\operatorname{spt}(H^k \cap R) = M \cup S$ *where $M \in \boldsymbol{M}_k(\Omega)$, S is closed and $H^k(S) = 0$. (S is the singular set.) In case $\alpha > 0$, the tangent map of M is locally Hölder continuous with exponent α.*

It is not true in general that $S = \emptyset$; except in special cases, not much is known about S.

The partitioning problem which we are about to discuss is a geometric minimization problem with a constraint. Fix an integer $\mu \geqq 2$. We say a sequence $G_1, \cdots, G_\mu$ of disjoint open subsets of $\boldsymbol{R}^n$ is an *admissible L^n partition* of $\boldsymbol{R}^n$ if G_i is bounded, $1 \leqq i \leqq \mu - 1$; $L^n(\boldsymbol{R}^n \sim \bigcup_{i=1}^{\mu} G_i) = 0$ and $\operatorname{Bdry} G_i \in \boldsymbol{R}_{n-1}(\boldsymbol{R}^n)$, $1 \leqq i \leqq \mu$.

Minimization problem. *Suppose* $0 < V_i < \infty$, $1 \leqq i \leqq \mu - 1$; $0 < W_{i,j} < \infty$, $1 \leqq i < j \leqq \mu$; *and* Φ *is an* $(n-1)$*-integrand on* $\boldsymbol{R}^n$. *Among all admissible* L^n *partitions* $G_1, \cdots, G_\mu$ *of* $\boldsymbol{R}^n$ *with* $L^n(G_i) = V_i$, $1 \leqq i \leqq \mu - 1$, *is it possible to find one for which*

$$\sum_{1 \leqq i < j \leqq \mu} W_{i,j} \int_{B_{i,j}} \Phi(x, (B_{i,j})_x)\, dH^{n-1}x$$

is least? (*Here* $B_{i,j} = \text{Bdry } G_i \cap \text{Bdry } G_j$.)

THEOREM (ALMGREN). *If* (a) $\Phi = \Phi(x, S)$ *is elliptic and depends trivially on* x, (b) *for each* $i = 1, \cdots, \mu$ *and each nonzero* μ*-tuple* $(b_1, \cdots, b_\mu)$ *of nonnegative real numbers there exists* $j \neq i$ *such that* $b_j W_{i,j} > \sum_{k \notin (i,j)} b_k (W_{k,j} - W_{i,k})$ (*here we have written* $W_{i,j} = W_{j,i}$ *in case* $i > j$), *then* (c) *the minimization problem has a solution*; (d) *if* $G_1, \cdots, G_\mu$ *solves the minimization problem,* $\bigcup_{i=1}^{\mu} \text{Bdry } G_i$ *is* H^{n-1} *almost equal a member of* $\boldsymbol{M}_{n-1}(\boldsymbol{R}^n)$ *which is* C^∞.

Heuristically, (b) says that it is always advantageous to eliminate a particular G_i by adding it to a judiciously chosen G_j. The proof of (d) depends on the regularity theorem.

In case $k = 2$, $n = 3$, and we are dealing with the area integrand A, a very satisfactory regularity theorem is known. We introduce notation to facilitate its formulation. Whenever $A \subset \boldsymbol{R}^3$ and $b \in \boldsymbol{R}^3$, let $\Upsilon_b(A)$ be the class of subsets B of $\boldsymbol{R}^3$ for which there exist open neighborhoods U of 0 and V of b and a C^1 diffeomorphism $F: U \to V$ such that $F(0) = b$, $DF(0)$ is an isometry and $F(A \cap U) = B \cap V$. Let ρ be a rotation of $\boldsymbol{R}^3$ by $2\pi/3$ and let H be a closed half-plane in $\boldsymbol{R}^3$ whose edge contains the axis of ρ. Let $Y = \bigcup_{i=0}^{2} \rho^i(H)$. Fix a regular tetrahedron in $\boldsymbol{R}^3$ with center of gravity 0. Let T be the set of points tx such that $t \geqq 0$ and x lies in the closure of the 1-skeleton of T.

THEOREM (TAYLOR). (a) *Y and T are, up to rotation and H^2 null sets, the only* $(A, \infty, 0)$ *minimizing cones with vertex* 0 *in* $\boldsymbol{R}_2(\boldsymbol{R}^3)$ *besides the members of* $\boldsymbol{G}(3, 2)$.

(b) *Suppose* $\Omega \subset \boldsymbol{R}^3$, $\delta > 0$, $\int_0^\delta r^{-1}\beta(r)^{1/2}\, dr < \infty$, $R \in \boldsymbol{R}_2(\Omega)$ *and* R *is* (A, δ, β) *minimizing. Then* spt $H^2 \cap R = M \cup C \cup D$, *where* $M \in \boldsymbol{M}_2(\Omega)$, $C \in \boldsymbol{M}_1(\Omega)$, $D \in \boldsymbol{M}_0(\Omega)$; *for each* $c \in C$, $M \cup C \in \Upsilon_c(Y)$; *for each* $d \in D$, $M \cup C \cup D \in \Upsilon_d(T)$.

EXAMPLE. Let $B = \{x \in \boldsymbol{R}^{2m} : \sum_{i=1}^{m} x_i^2 = \sum_{i=m+1}^{2m} x_i^2 = \sqrt{2}/2\}$ and let $C = \{tx : 0 \leqq t \leqq 1 \text{ and } x \in B\}$.

THEOREM (BOMBIERI, DEGIORGI, GIUSTI). *If* $m = 4, 5, 6, \cdots$, *then* $H^{2m-1}(C) < H^{2m-1}(R)$ *whenever* $R \in \boldsymbol{R}_{2m-1}(\boldsymbol{R}^{2m})$, R *is compact,* R *cannot be retracted on* B *and* R *is not* H^{2m-1} *almost equal* C.

COROLLARY. *If* $m = 4, 5, 6, \cdots$, *and* $M \in \boldsymbol{M}_{2m-1}(\boldsymbol{R}^{2m})$ *has boundary* B, *then* $\delta(A; M)(g) < 0$ *for some* $g \in \mathscr{X}(\boldsymbol{R}^{2m})$ *with* spt $g \cap B = \emptyset$.

LEMMA (SIMONS). *Suppose* $m = 2, 3$. *Then* $\delta(A; C)(g) = 0$ *whenever* $g \in \mathscr{X}(\boldsymbol{R}^{2m})$ *and* spt $g \cap B = \emptyset$. *However, there exists* $g \in \mathscr{X}(\boldsymbol{R}^{2m})$ *with* spt $g \cap B = \emptyset$ *such that* $(d/dt)^2 H^{2m-1}[h_t(C)]\big|_{t=0} < 0$, *where* $h_t(x) = x + tg(x)$ *for* $(t, x) \in \boldsymbol{R} \times \boldsymbol{R}^{2m}$.

THEOREM. *Suppose* $n = 3, 4, 5, 6, 7$. *Suppose* $B \in \boldsymbol{M}_{n-2}(\boldsymbol{R}^n)$ *is compact. Among all*

$M \in \boldsymbol{M}_{n-1}(\boldsymbol{R}^n)$ with boundary B, there is at least one which minimizes H^{n-1} measure. Any such manifold is analytic away from B and (smooth) analytic at B if B is (smooth) analytic.

From the first theorem we learn that singularities arise even in codimension 1 minimizing problems. From the corollary we see that a singular extremal, unlike a smooth extremal, need not locally minimize.

Some theorems about the Euler-Lagrange operator.

INTERIOR REGULARITY THEOREM (ALLARD). *Whenever $0 < \varepsilon < 1$ and $k < p < \infty$ there is $\eta > 0$ with the following property: If $\Omega = \{x : |x| < 1\}$, $R \in \boldsymbol{R}_k(\Omega)$, $0 \in \operatorname{spt} H^k \cap R$, $H^k(R) \leqq (1 + \eta)\alpha(k)$ and $\|\delta(A; R)\|_p \leqq \eta$ then $M = \{x : |x| < 1 - \varepsilon\} \cap \operatorname{spt}(H^k \cap R) \in \boldsymbol{M}_k(\Omega)$, $H^k(M \sim R) = 0$ and $\|M_x - M_a\| \leqq \varepsilon |x - a|^{1-k/p}$ whenever $x, a \in M$.*

Roughly speaking, one might say that if R is like a k-disc in area and generalized mean curvature, it is like a k-disc in the $C^1_{1-k/p}$ topology. We have a relative version of this theorem in which the model is a half disc instead of a disc. If $B \in \boldsymbol{M}_{k-1}(\{x : |x| < 1\})$, $0 < R < \infty$, we say *B is of reach R* if, whenever $b \in B$, ν is normal to B at b and $|\nu| = R$, the ball $\{x : |x - (b + \nu)| < R\}$ does not meet B.

BOUNDARY REGULARITY THEOREM (ALLARD). *Whenever $0 < \varepsilon < 1$ and $k < p < \infty$ there is $\eta > 0$ with the following property: If $B \in \boldsymbol{M}_{k-1}(\{x : |x| < 1\})$, B is smooth, B is closed relative to $\{x : |x| < 1\}$, $0 \in B$, B is of reach $1/\eta$, $\Omega = \{x : |x| < 1\} \sim B$, $R \in \boldsymbol{R}_k(\Omega \sim B)$, distance$(0, \operatorname{spt} H^k \cap R) \leqq \eta$, $H^k(R) \leqq (1 + \eta)\alpha(k)/2$ and $\|\delta(A; R)\|_p \leqq \eta$ then $M = \{x : |x| < 1 - \varepsilon\} \cap \operatorname{spt}(H^k \cap R) \in \boldsymbol{M}_k(\Omega)$, $H^k(M \sim R) = 0$, M is a C^1 manifold with boundary $B \cap \{x : |x| < 1 - \varepsilon\}$ in $\{x : |x| < 1 - \varepsilon\}$ and $\|M_x - M_a\| \leqq \varepsilon |x - a|^{1-k/p}$ whenever $x, a \in M$.*

It is not known for what class of integrands these theorems are true. In dealing with the area integrand it is extremely useful to study $\delta(A, R)(g)$ where $g(x) = \varphi(|x|)x$ for $x \in \boldsymbol{R}^n$ because one obtains thereby information about $r^{-k}H^k(R \cap \{x : |x| < r\})$ as a function of r. Little is known about how this quantity changes in terms of $\delta(\Phi; R)$ when Φ is not the area integrand.

Does every compact Riemannian submanifold contain a closed submanifold of everywhere zero mean curvature having a given dimension? The answer is not known in general. Using Morse theory methods Almgren has proved the following.

EXISTENCE THEOREM. *Suppose N is a compact Riemannian manifold and $0 < k < \dim N$. There are a positive integer μ and $R_1, \cdots, R_\mu \in \boldsymbol{R}_k(N)$ such that $R_i = \operatorname{spt}(H^k \cap R_i) \neq \emptyset$, $1 \leqq i \leqq \mu$; $R_1 \supset \cdots \supset R_\mu$ and $\sum_{i=1}^{\mu} \delta(A; R_i) = 0$.*

Let us now take $R_1, \cdots, R_\mu$ as in the conclusion of the existence theorem. We would like to apply the interior regularity theorem at H^k almost all points of R_1 but cannot since we do not know if $\delta(A; R_1) = 0$. Since the R_i are closed, simple point set theoretic considerations allow us to conclude that $\delta(A; R_1) = 0$ near the points of dense subset of R_1. Thus for some $M \in \boldsymbol{M}_k(N)$ we have that Closure M

$= R_1$. It is our hope that some recent deep unpublished work of Almgren will be carried to the point where one can conclude that $H^k(R_1 \sim M) = 0$.

It has been know for some time that, in case $k = 1$, one can take $\mu = 1$ and $R_1 \in \boldsymbol{M}_1(N)$ in the conclusion of the existence theorem. Very recently, this result has been extended to the case $k = 2$, $n = 3$.

THEOREM (PITTS). *Suppose* dim $N = 3$. *There is* $M \in \boldsymbol{M}_2(N)$ *such that* M *is closed and* $\delta(A; M) = 0$.

The theorem is proved by combining the proof of the existence theorem with the interior regularity theorem and some classical 2-dimensional differential geometry.

References

1. W. K. Allard, *On the first variation of a varifold*, Ann. of Math. (2) **95** (1972), 417–491. MR **46** #6136.

2. ———, *On the first variation of a varifold: boundary behavior*, Ann. of Math. (to appear).

3. F. J. Almgren, Jr., *Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints*, Mem. Amer. Math. Soc. (preprint).

4. E. Bombieri, E. DeGiorgi and E. Giusti, *Minimal cones and the Bernstein problem*, Invent. Math. **7** (1969), 243–268. MR **40** #3445.

5. H. Federer, *Geometric measure theory*, Die Grundlehren der math. Wissenschaften, Band 153, Springer-Verlag, New York, 1969. MR **41** #1976.

6. J. Pitts, *Every compact three dimensional manifold contains two dimensional minimal submanifolds*, Ph.D. Thesis, Princeton University, Princeton, N.J., 1974.

7. J. Simons, *Minimal varieties in Riemannian manifolds*, Ann. of Math. (2) **88** (1968), 62–105. MR **38** #1617.

8. J. Taylor, *The structure of singularities in soap-bubble-like and soap-film-like minimal surfaces*, Rutgers University, New Brunswick, N.J. (preprint).

PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08540, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Free Boundary Problems in the Theory of Fluid Flow Through Porous Media

Claudio Baiocchi

A. Introduction. It is well known that variational inequalities, and in particular inequalities with obstacles, give rise to *free boundary problems*, the free boundary being the boundary of the "coincidence set"; and conversely in some free boundary problems the free boundary can be obtained as the boundary of the coincidence set for the solution of a variational inequality (see [22]; and the articles in these PROCEEDINGS of Kinderlehrer [21] and Lions [23]).

However, in many concrete free boundary problems, the solution *cannot be directly* seen as the solution of a variational inequality (see [23] again, in particular §3); and this is the case of the free boundary problems connected with the fluid flow through porous media. In order to study this type of problem I introduced in [2] a change of unknown functions which still enables one to see the free boundary as the boundary of the coincidence set for the solution of a variational problem; more precisely, by means of this transformation, the free boundary problem is reduced to *one* variational inequality in the simplest cases (see [2], [5], [6], [7], [9], [24], [25]); into *a family*, depending on a real parameter, of variational inequalities, in more complicated situations (see [5], [6], [7], [13], [14], [26], [28]); and into a *quasi*-variational inequality in the general case (see [3]).

Here we will limit ourselves to describe the transformation for some free boundary problems arising from hydraulics; however the method seems to have more general applications. It has been adapted to problems of fluid dynamics (subsonic flow around an obstacle, see [11], [12]) and of thermodynamics (a Stefan-type problem, see [18]). We will also limit ourselves to stating *theoretical* results for our problem; however we want to point out that the method is very well adapted to the construction of *numerical* solutions: Indeed it suggests algorithms which are

justified from the mathematical point of view and which compete very well with the ones already known[1] both in simplicity of programming and in speed of execution (see [4], [6], [7], [14], [15], [16], [28]); although in many cases (see [6], [7], [13], [14], [28]) some theoretical results have been obtained by using numerical analysis, say by passing to the limit on approximate solutions.

B. The physical problem. Two water reservoirs, of different levels, are separated by an earth dam; water flows from the highest level to the lowest one; and we look for the quantities (like flow region, streamlines, velocity, pressure, discharge,...) associated with the motion.

We will limit ourselves to the case of a homogeneous isotropic dam, on a horizontal impervious basis, for a steady, irrotational, incompressible and bidimensional[2] flow; moreover we neglect capillarity and evaporation effects. However the method has been adapted to more general situations like variable permeability ([5], [9]), sloping basis ([5], [13], [14]), evolution problems [27], tridimensional dam [25], evaporation [24]; problems like infiltration from a channel [26] and presence of many liquids with different densities[3] [5] can also be treated.

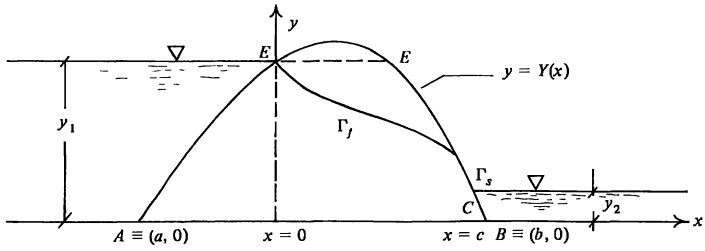
We will suppose that the cross section D of the dam has the form

$$D = \{(x, y) \mid a < x < b; 0 < y < Y(x)\} \tag{B. 1}$$

where $a < 0 < b$, Y is a "regular" function, e.g.,

$$Y(x) \text{ is } C^2 \text{ and concave on } [a, b] \tag{B. 2}$$

and, for a suitable c with $0 < c < b$,[4] we have $Y(0) = y_1$, $Y(c) = y_2$; y_i denoting the heights of the reservoirs.



The problem is usually[5] stated as follows: We look for a subset Ω of D and for a function $u(x, y)$ such that

$$\Delta u = 0 \quad \text{in } \Omega \qquad (\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2); \tag{B. 3}$$

[1] See, e.g., [17] for a survey on numerical treatments of free boundary problems.

[2] I.e., along the direction orthogonal to the figure the dam is infinitely extended and has a constant section.

[3] E.g., fresh and salt water in a coastal aquifer.

[4] A case where ECB is vertical, so $c = b$, will be treated in §D.

[5] See, e.g., [1], [8], [20]. Ω is the flow region; Γ_f, Γ_s are respectively the "free line" and the "seepage line"; $u(x, y)$ is a "potential velocity", say the velocity V of water in Ω is proportional to the gradient of u: $V = -k\nabla u$ (k being the permeability coefficient). For the physical meaning of (B. 3), ···, (B. 7) see [1], [8], [20] again; (B. 8) means that there is no water coming into Ω through Γ_s.

(B. 4) $$u\big|_{AF} = y_1; \quad u\big|_{BC} = y_2; \quad u\big|_{\Gamma_s} = y,$$

where $\Gamma_s = \partial\Omega \cap \partial D \backslash FABC$;

(B. 5) $$u\big|_{\Gamma_f} = y, \quad \text{where } \Gamma_f = \partial\Omega \cap D;$$

(B. 6) $$(\partial u/\partial y)\big|_{AB} = 0;$$

(B. 7) $$(\partial u/\partial n)\big|_{\Gamma_f} = 0 \quad (\partial/\partial n \text{ outward normal derivative});$$

in [3], in order to exclude some "nonphysical" solutions, we proposed to impose also the relation:

(B. 8) $$(\partial u/\partial n)\big|_{\Gamma_s} \leqq 0.$$

REMARK B.1. Obviously Ω cannot be "any" subset of D, e.g., setting $D_1 = \{(x, y) \mid (x, y) \in D, \ a < x \leqq 0\}$, $D_2 = \{(x, y) \in D \mid c \leqq x < b\}$, $D_3 = \{(x, y) \mid (x, y) \in D, \ 0 < x < c\}$ and denoting by φ the "upper boundary" of Ω in D_3, say

(B. 9) $$\varphi(x) = \sup\{y \mid (x, y) \in \Omega\}, \qquad 0 < x < c,$$

we must have

(B. 10) $$\Omega = D_1 \cup D_2 \cup \{(x, y) \mid (x, y) \in D_3, \ y < \varphi(x)\}.$$

Regularity assumptions on φ, u and the meaning of (B. 3), ···, (B. 8) will be summarized in § C.

REMARK B.2. Problem (B. 3), ···, (B. 7) is a typical free boundary problem: In a domain Ω with boundary partially unknown we must solve a problem with "too many" conditions (see (B. 5), (B. 7)) on the unknown part Γ_f of $\partial\Omega$.

C. Reduction to a quasi-variational inequality. We start with a definition of weak solutions for our problem.

DEFINITION C.1. A couple $\{\Omega, u\}$ is a weak solution of the problem (B. 3), ···, (B. 8) if

(C.1) Ω is an open connected subset of D such that, with the definition (B.9) of φ, (B.10) holds;

(C.2) $u \in C^0(\bar{\Omega}) \cap H^1(\Omega)$,[6] and satisfies (B.4), (B.5) in the sense of the continuous functions;

(C.3) for any $\psi \in C^1(\bar{D})$ vanishing near $AF \cup BC$ and satisfying $\psi(x, Y(x)) \geqq 0$ for $0 < x < c$, we have $\int_\Omega \nabla u \cdot \nabla\psi \, dx \, dy \leqq 0$. (Note that (C.3) contains (B.3) in the distribution sense and (B.6), (B.7), (B.8) in the usual variational sense.)

REMARK C.1. Condition (C.2) is "almost optimal" because we can prove that any solution must have unbounded gradient in Ω; on the contrary, condition (C.1) is very weak: From the physical point of view we could impose a hypothesis of the type

(C.4) φ is continuous strictly decreasing on $[0, c]$;

a weak solution satisfying(C.4) will be called a *strong* solution.

We perform now a *change of unknown functions*, setting

[6] Say u is continuous on $\bar{\Omega}$ and its first derivatives (in the distribution sense) are square-integrable on Ω.

$$\text{(C. 5)}\qquad \begin{aligned} \tilde{u}(x, y) &= u(x, y) && \text{in } \bar{\Omega}, \\ &= y && \text{in } \bar{D}\backslash\bar{\Omega}; \\ z(x, y) &= \int_0^y [t - \tilde{u}(x, t)]\, dt && \text{in } \bar{D}. \end{aligned}$$

We can prove (for the details see [3]) that $\{\Omega, u\}$ *is a weak solution if and only if* $z(x, y)$ *satisfies*:

$$\text{(C. 6)}\qquad z \in C^1(\bar{D});$$

$$\text{(C. 7)}\qquad z\big|_{AB} = 0;\quad z_{y|_{AF}} = y - y_1;\quad z_{y|_{BC}} = y - y_2;\quad z_{y|_{FBEC}} = 0;$$

$$\text{(C. 8)}\qquad [0, c] \ni x \mapsto z(x, Y(x)) \text{ is concave};$$

$$\text{(C. 9)}\qquad \Delta z \in \chi_{D_1} + \chi_{D_2} + \chi_{D_3} \cdot H(z - z(x, Y(x))),$$

where D_i are defined in Remark B.1, χ_{D_i} is the characteristic function of D_i, and H is the maximal monotonic graph associated with the Heaviside function.[7]

REMARK C.2. Formula (C.5) gives the map $\{\Omega, u\} \to z$; the inverse map can be given in the form

$$\text{(C. 10)}\qquad \begin{aligned} \Omega &= D_1 \cup D_2 \cup \{(x, y) \in D_3 \,|\, z(x, y) > z(x, Y(x))\}, \\ u(x, y) &= y - z_y(x, y) \quad \text{in } \bar{\Omega}. \end{aligned}$$

In order to study problem (C. 6), ···, (C. 9) we first remark that on setting

$$\text{(C. 11)}\qquad \begin{aligned} a(u, v) &= \int_D [(u_x - Y'u_y)v_x + (u_y + Y'u_x)v_y - Y''u_y v]\, dx\, dy \quad \forall\, u, v \in H^1(D),^8 \\ L(V) &= -\int_{D_1 \cup D_2} v\, dx\, dy + \int_a^0 (1 + Y'^2)(Y(x) - y_1)v(x, Y(x))\, dx \\ &\quad + \int_c^b (1 + Y'^2)(Y(x) - y_2)v(x, Y(x))\, dx \qquad \forall\, v \in H^1(D),^9 \\ j(W, v) &= \int_{D_3} (v(x, y) - W(x, Y(x)))^+\, dx\, dy \qquad \forall\, v, w \in H^1(D),^{10} \end{aligned}$$

under hypothesis (B.2) on $Y(x)$, we have

(C. 12) For any $W \in C^1(\bar{D})$ there exists a unique $z_W \in H^1(D)$ with $z_{W|_{AB}} = 0$ such that, for any $v \in H^1(D)$ with $v\big|_{AB} = 0$, we have

$$a(z_W, v - z_W) + j(W, v) \geqq j(W, z_W) + L(v - z_W);$$

moreover, for problem (C. 12) we have a "regularity result":

(C. 13) For any $W \in C^1(\bar{D})$, $z_W \in C^1(\bar{D})$; moreover z_W satisfies (C. 7) and $\Delta z_W \in \chi_{D_1} + \chi_{D_2} + \chi_{D_3} \cdot H(z - W(x, Y(x)))$,

so that the solutions of (C.6), (C.7), (C.9) are the fixed points for the map $W \to z_W$. In other words we must solve:

---

[7] I.e., H is the multivalued function defined by $H(t) = \{0\}$ for $t < 0, H(t) = \{1\}$ for $t > 0$, $H(0) = \{h \,|\, 0 \leqq h \leqq 1\}$.

[8] $a(u, v)$ is the bilinear from associated to the Laplace operator for which the "natural condition" on $AFECB$ is $\partial/\partial y$.

[9] L takes into account the terms $\chi_{D_1} + \chi_{D_2}$ in (C. 9), and the nonhomogeneous conditions on z_y in (C. 7).

[10] We set $t^+ = t$ for $t \geqq 0$ and $t^+ = 0$ for $t \leqq 0$; remark that $H(t)$ is the subdifferential of $t \to t^+$.

$$\text{(C.14)}\qquad \begin{array}{l} z \in H^1(D); \qquad z\big|_{AB} = 0; \\ \text{for any } v \in H^1(D) \text{ with } v\big|_{AB} = 0, \quad a(z, v - z) + j(z, v) \geqq j(z, z) + L(v - z). \end{array}$$

Problem (C.14) is a *quasi-variational inequality*,[11] and we can prove that there exist a maximal solution $z_{\max}$ and a minimal solution $z_{\min}$ of (C.14). Remark however that we do not need any solution of (C.14) (say of (C.6), (C.7), (C.9)) but only the solutions of (C.14) which satisfy (C.8). Actually we can prove that $(z_{\max}(x, Y(x)))'' \leqq 0$ in $]0, c[$, while, in general, $(z_{\min}(x, Y(x)))'' \not\leqq 0$ in $]0, c[$; however, in the family of the solutions z of (C.14) which satisfy (C.8) there exists a minimum element that will be denoted $\tilde{z}_{\min}$.

By means of these results we get an existence theorem for weak solutions (in the sense of Definition C.1) of the free boundary problem; a uniqueness result, and the validity of (C.4), will be proved in § D for the case when "the right wall is vertical"; remark that, in general, the uniqueness of the solution is equivalent to the validity of the relation

$$\text{(C. 15)}\qquad z_{\max} = \tilde{z}_{\min},$$

while the regularity of the solution, i.e., the validity of (C.4), is a problem of regularity for the boundary of the "coincidence set".[12]

REMARK C.3. Our theory seems to be very well adapted to numerical computations; in particular we make use of efficient and rigorous algorithms in order to get approximations of $z_{\max}$ and $\tilde{z}_{\min}$ (see [**15**]); numerical results suggest the conjecture that, in any case, (C.15) holds.

D. The case when the right wall is vertical. In this section we will study the problem assuming that ECB is vertical, say $c = b$ and $ECB = \{(c, y) \big| 0 \leqq y \leqq y_1\}$.

Let us first remark that in this case, for the strong solutions, we have

$$\text{(D. 1)}\qquad \int_0^{\varphi(x)} u_x(x, t)\, dt \text{ is constant (a. e.) on } [0, c].$$

In terms of z this relation gives

$$\text{(D. 2)}\qquad \exists q : z(x, Y(x)) = z(E) + q(x - c) \qquad \forall\, x \in [0, c];$$

moreover some of the *oblique* derivative conditions in (C.7) become *tangential*, so that we get

$$\text{(D. 3)}\qquad z\big|_{AB} = 0; \quad z\big|_{BC} = \frac{y^2 - 2y_2 y}{2}; \quad z\big|_{CE} = -\, y_2^2 / 2$$

and from (D.2), (D.3) (with an unknown q!):

$$\text{(D. 4)}\qquad z(x, Y(x)) = -\,(y_2^2/2) + q(w - c) \qquad \forall\, x \in [0, c];$$

[11] Quasi-variational inequalities have been introduced in [**10**] in order to study some impulsive-control problems; also in this case it was a free boundary problem (see the article of J. L. Lions in these PROCEEDINGS [**23**]).

[12] For this problem, in general, see the article of Kinderlehrer in these PROCEEDINGS [**21**]; for a result related to infiltration problems see [**19**].

moreover we still have

$$z_{y|_{AF}} = y - y_1, \tag{D. 5}$$

$$z_{y|_{FE}} = 0. \tag{D. 6}$$

Suppose now that q is known; on setting

$$\begin{aligned} K_q = \{v \mid v \in H^1(D);\ v|_{AB} = 0;\ v|_{BC} = (y^2 - 2y_2 y)/2; \\ v|_{CE} = -y_2^2/2;\ v(x, Y(x)) = -y_2^2/2 + q(x - c)\ (0 \leqq x \leqq c)\}; \\ j_q(v) = \int_{D_3} ((v + y_2^2/2) - q(x - c))^+ \, dx\, dy, \end{aligned} \tag{D. 7}$$

we still can put the problem (C.6), (C.7), (C.9) into a variational form:

$$z \in K_q,\ \forall\, v \in K_q, \qquad a(z, v - z) + j_q(v) \geqq j_q(z) + L(v - z) \tag{D. 8}$$

($a(u, v)$ and $L(v)$ being similar to the ones defined in (C.11)); (D.8) contains (D.3), (D.4), (D.5), (C.9); moreover (C.8) is obviously satisfied (see (D.4)); so that, in order to solve (C.6), ⋯, (C.9) we only need a value q^* of q such that the corresponding solution z_{q^*} of (D.8) satisfies (D.6), (C.6). In [5], [13] it is proved that such a q^* exists and is unique; moreover z_{q^*} is such that the corresponding φ defined through (C.9), (B.9) satisfies (C.4); so that we have an existence and uniqueness theorem for strong solutions.

We can also prove that in this case the procedure of § C gives $z_{\max} = \tilde{z}_{\min} = z_{q^*}$; this result can be regarded as a uniqueness and regularity theorem for the quasi-variational problem.

Finally we consider a more particular case, assuming that D is a rectangle: $D =]0, c[\times]0, y_1[$. In this case also (D.5) is a tangential condition; from (D.3), (D.4), (D.5), z being continuous, we can deduce

$$z|_{AF} = (y^2 - 2y_1 y)/2, \quad q = (y_1^2 - y_2^2)/2c, \tag{D. 10}$$

so that the values of z are known *on the whole* ∂D. Let $g(x, y)$ be a function defined on ∂D and assume the same values of z; on setting

$$K = \{v \mid v \in H^1(D);\ v|_{\partial D} = g\} \tag{D. 11}$$

we can write the problem in the form:

$$\begin{aligned} z \in K,\ \forall\, v \in K, \quad a(z, v - z) + \int_D (v - g(x, y_1))^+ \, dx\, dy \\ \geqq \int_D (z - g(x, y_1))^+ \, dx\, dy \end{aligned} \tag{D. 12}$$

where a can now be *any* bilinear form associated to the Laplace operator; for instance we can choose $a(u, v) = \int_D \nabla u \cdot \nabla v \, dx\, dy$, and in this case (D.12) is equivalent to:

$$\begin{aligned} &z \in K;\ z \text{ minimizes on } K \text{ the functional } J \text{ given by} \\ &J(v) = \frac{1}{2} \int_D |\nabla v|^2 \, dx\, dy + \int_D (v - g(x, y_1))^+ \, dx\, dy. \end{aligned} \tag{D. 13}$$

We still have that the (unique, q being known; see (D.10)) solution z of (D.12)

(or (D.13)) is regular and satisfies (D.6); and we can reach a conclusion as in the previous case (for details see [2], [5]).

Bibliography

1. V. I. Aravin and S. A. Numerov, *Theory of fluid flow in undeformable porous media*, Moscow, 1953; English transl., Israel Program for Scientific Translations, Jerusalem, 1955.
2. C. Baiocchi, *Sur un problème à frontière libre traduisant le filtrage de liquides à travers des milieux poreux*, Ann. Mat. Pura Appl. **92** (1972), 107–127; C. R. Acad. Sci. Paris Sér. A-B **273** (1971), A1215–A1217. MR **45** #6265.
3. ———, Boll. Un. Mat. Ital. (to appear); Prelim. note C. R. Acad. Sci. Paris Sér. A-B **278** (1974), A1201–A1204.
4. C. Baiocchi, V. Comincioli, L. Guerri and G. Volpi, Calcolo **10** (1973), 1–86.
5. C. Baiocchi, V. Comincioli, E. Magenes and G. A. Pozzi, Ann. Mat. Pura Appl. **97** (1973), 1–82.
6. C. Baiocchi and E. Magenes, Uspehi Mat. Nauk **29**, 3 (1974), 50–69.
7. ———, *Atti del convegno: Metodi valutativi della Fisica Matematica*, Accademia Nazionale, Lincei, 1972.
8. J. Bear, *Dynamics of fluid in porous media*, American Elsevier, New York, 1972.
9. V. Benci, Ann. Mat. Pura Appl. **100** (1974), 191–209.
10. A. Bensoussan and J. L. Lions, *Nouvelle formulation de problèmes de contrôle impulsionnel et applications*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A1189–A1192. MR **47** #5690.
11. H. Brézis and G. Duvaut, *Écoulements avec sillages autour d'un profil symmétrique sans incidence*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A875–A878. MR **47** #4520.
12. H. Brézis and G. Stampacchia, *Une nouvelle méthode pour l'étude d'écoulements stationnaires*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A129–A132. MR **47** #4521.
13. V. Comincioli, Ann. Mat. Pura Appl. **100** (1974), 211–238.
14. ———, Appl. Math. and Optimization (to appear).
15. ———, Pubbl. del L.A.N. Pavia, no. 79, 1974.
16. V. Comincioli, L. Guerri and G. Volpi, Pubbl. del L.A.N. Pavia, no. 17, 1971.
17. C. W. Cryer, J. Assoc. Comput. Math. **17** (1970), 397–411.
18. G. Duvaut, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A1461–A1463.
19. A. Friedman (to appear).
20. M. E. Harr, *Groundwater and seepage*, McGraw-Hill, New York, 1962.
21. D. Kinderlehrer, *Elliptic variational inequalities*, these PROCEEDINGS.
22. H. Lewy and G. Stampacchia, *On the regularity of the solution of a variational inequality*, Comm. Pure Appl. Math. **22** (1969), 153–188. MR **40** #816.
23. J. L. Lions, these PROCEEDINGS.
24. G. A. Pozzi, Boll. Un. Mat. Ital. **9** (1974), 416–440.
25. G. Stampacchia, Uspehi Mat. Nauk **29**, 4 (1974), 88–101.
26. A. Torelli, Rend. Sem. Mat. Univ. Padova (to appear).
27. ———, Boll. Un. Mat. Ital. (to appear).
28. G. Volpi (in preparation).

UNIVERSITÀ DI PAVIA AND LABORATORIO DI ANALISI NUMERICA DEL C.N.R.
PAVIA, ITALY

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

On a Class of Fuchsian Type Partial Differential Operators

M. S. Baouendi*

We shall mainly discuss in this article a characteristic local Cauchy problem for differential operators of the following form:

$$P(t, x, D_t, D_x) = t^k D_t^m + \sum_{|\alpha|+p\leq m;\ 0\leq p\leq m-1} a_{p,\alpha}(t, x) D_t^p D_x^\alpha \tag{1}$$

where k is a nonnegative real number, $t \in \boldsymbol{R}$, $x \in \boldsymbol{R}^n$, $a_{p,\alpha}$ being defined in a neighborhood of 0 in $\boldsymbol{R}^{n+1}$; the initial surface is $\{t = 0\}$.

It is easily seen that no natural Cauchy problems are "well-posed" for such operators (1). We need some restrictions. We first define the notion of *Fuchsian weight* of a differential monomial. If

$$Q = a(t, x) t^l D_t^p D_x^\alpha$$

where a is a continuous function defined in a neighborhood of 0 in $\boldsymbol{R}^{n+1}$, l is a nonnegative real number, the Fuchsian weight of Q is said to be less than or equal to $p - l$. It is exactly $p - l$ if $a(0, 0) \neq 0$.

A differential operator (1) is called of *Fuchsian type* if the Fuchsian weight of any $a_{\alpha,p} D_t^p D_x^\alpha$ is less than or equal to $m - k$, and is strictly less than $m - k$ if $|\alpha| \neq 0$. That is to say that P can be written in the following form:

$$P = t^k D_t^m + \sum_{|\alpha|+p\leq m;\ 0\leq p\leq m-1} b_{p,\alpha}(t, x) t^{e_{p,\alpha}} D_t^p D_x^\alpha; \tag{2}$$

the coefficients $b_{p,\alpha}$ are continuous functions near 0 in $\boldsymbol{R}^{n+1}$ and $e_{p,\alpha}$ are nonnegative real numbers satisfying $e_{p,\alpha} \geq p - m + k$ and $e_{p,\alpha} > p - m + k$ if $|\alpha| \neq 0$. For more simplicity, we restrict ourselves here, unless otherwise mentioned,

*Supported by NSF grant GP-35825. The contribution of the author is in a joint work with C. Goulaouic (University of Paris).

to the case $k \in \boldsymbol{N}$, $e_{p,\alpha} \in \boldsymbol{N}$ and we assume the coefficients $b_{p,\alpha}$ to be $\mathscr{C}^\infty$ functions. If $k > m$, t^{k-m} is a common factor so that, without loss of generality, we can assume from now on that $0 \leqq k \leqq m$. With the assumptions made above, the operator (2) can be written

$$P = Q + R \tag{3}$$

where

$$Q = t^k D_t^m + a_{m-1}(x) t^{k-1} D_t^{m-1} + \cdots + a_{m-k}(x) D_t^{m-k},$$
$$R = \sum_{0<p\leqq m-1} \sum_{|\alpha|\leqq m-p} t^{\max(0,\, k+p-m+1)}\, a_{p,\alpha}(t, x) D_t^p D_x^\alpha.$$

We say that Q is the principal Fuchsian part of P. It is an ordinary differential operator in t where x is a parameter. The characteristic polynomial associated with (3) is

$$\mathscr{C}(\lambda, x) = \lambda(\lambda - 1) \cdots (\lambda - m + 1) + \cdots + a_{m-k}(x)\lambda(\lambda - 1) \cdots (\lambda - m + k + 1)$$

and its roots, called characteristic roots, are denoted by $\lambda_1(x), \cdots, \lambda_k(x)$, $\lambda_{k+1} = 0$, $\cdots, \lambda_m = m - k - 1$.

The first result is a generalization of the Cauchy-Kovalevsky theorem:

THEOREM 1. *Assume that the coefficients of P are analytic functions. The following two conditions are equivalent*:

(i) $\lambda_j(0) \notin \{l \in \boldsymbol{N}, l \geqq m - k\}$, *for* $1 \leqq j \leqq m$.

(ii) *For any analytic functions* $u_0, \cdots, u_{m-k-1}$ *defined in a neighborhood of* 0 *in* $\boldsymbol{R}^n$ *and for any analytic function f defined in a neighborhood of* 0 *in* $\boldsymbol{R}^{n+1}$, *there is a unique analytic function u, defined in a neighborhood of* 0 *in* $\boldsymbol{R}^{n+1}$, *such that* $Pu = f$ *and* $D_t^j u(0, \cdot) = u_j$ *for* $0 \leqq j \leqq m - k - 1$.

If condition (i) fails, the kernel and the compatibility conditions for the Cauchy problems (ii) can be described.

We shall give now a uniqueness result. We assume that the coefficients of P are defined in a neighborhood of 0 in $\boldsymbol{R}^{n+1}$ of the form $(-T, T) \times \Omega$ where $T > 0$ and Ω is an open set in $\boldsymbol{R}^n$. We have

THEOREM 2. *Assume that the coefficients of P are* $\mathscr{C}^\infty$ *with respect to t valued in the space of analytic functions with respect to x. Let* $h \in \boldsymbol{N}$ *be such that* $\operatorname{Re} \lambda_j(0) < m - k + h$ *for* $1 \leqq j \leqq m$. *If* $u \in \mathscr{C}^\infty((-T, T), \mathscr{D}'(\Omega))$ *satisfies*

$$Pu \equiv 0 \text{ and } D_t^j u(0, \cdot) = 0 \quad \text{for } 0 \leqq j \leqq m - k + h - 1,$$

then u vanishes in a neighborhood of 0 *in* $\boldsymbol{R}^{n+1}$.

The proofs of Theorems 1 and 2 are in [5] and use an abstract method for solving the Cauchy problem (see [20], [17], [19] and others). Theorem 2 is a generalization of the Holmgren uniqueness theorem. Let us note that, in particular in the case $k = 0$, one does not need the analyticity of the coefficients with respect to t, as it is done usually in classical proofs (see [14], [15]). A similar result in the non-characteristic case is obtained also in [18].

It is not possible in general to start with any distribution u vanishing for $t < 0$ in the statement of Theorem 2, but one can consider "regular" distributions at $t = 0$ and use the regularity theorem stated below (Theorem 3).

If $l \in \boldsymbol{N}$, we denote by $\mathscr{C}^l_{-\infty}((-T, T), \mathscr{D}'(\Omega))$ the space of distributions u defined in $(-T, T) \times \Omega$ such that there are $m \in \boldsymbol{N}$ and $v \in \mathscr{C}^l((-T, T), \mathscr{D}'(\Omega))$ satisfying $u = (D_t t)^m v$. If $-l \in \boldsymbol{N}$, we denote by $\mathscr{C}^l_{-\infty}((-T, T), \mathscr{D}'(\Omega))$ the space of distributions u such that there is $w \in \mathscr{C}^0_{-\infty}((-T, T), \mathscr{D}'(\Omega))$ and $u = D_t^{-l} w$. We have:

THEOREM 3. *Let $l \in \boldsymbol{Z}$ be such that* $\operatorname{Re} \lambda_j(x) < l - m + k$ *for any $x \in \Omega$ and $1 \leqq j \leqq k$. If $u \in \mathscr{C}^l_{-\infty}((-T, T), \mathscr{D}'(\Omega))$ and $Pu \in \mathscr{C}^\infty((-T, T), \mathscr{D}'(\Omega))$ then $u \in \mathscr{C}^\infty((-T, T), \mathscr{D}'(\Omega))$.*

The proof of Theorem 3 and applications are given in [6]. Let us observe that the conclusion of Theorem 3 fails to be true for operators P of the form (1) with $\mathscr{C}^\infty$ coefficients if one does not assume that P is of Fuchsian type.

OTHER RESULTS. I would like to mention briefly some recent results closely related to the preceding ones:

1. Alinhac in [1] and [2] studied the Cauchy problem for Fuchsian hyperbolic operators. He proved in particular in [2] that if P is an operator of the form (1), with $\mathscr{C}^\infty$ coefficients, reduced to its principal part, and if there are uniqueness and existence in the space of $\mathscr{C}^\infty$ functions vansining at $t = 0$ as well as all their derivatives, P must be necessarily of Fuchsian type.

2. A class of singular evolution equations is studied in [3].

3. Helffer and Zuily proved in [13] that Fuchsian operators with $\mathscr{C}^\infty$ coefficients with $k > 0$ are not hypoelliptic. However if P is elliptic outside $\{t = 0\}$ a "partial hypoellipticity" and "analyticity" might be true. This problem has been studied in some general cases in [4], [8], [9] and [10] extending [11] and [12] to the characteristic case.

4. Some results for singular nonlinear Cauchy problems of Fuchsian type have been recently obtained [7]. This work is based on an adaptation to the singular case of some ideas of Nirenberg [16] used in the noncharacteristic case.

References

1. S. Alinhac, *Systèmes hyperboliques singuliers* (to appear).

2. ———, *Problèms de Cauchy pour des opérateurs singuliers*, Bull. Soc. Math. France **102** (1974), 289–315.

3. C. Baiocchi and M. S. Baouendi, *Équations d'évolution à coéfficients singuliers*, Goulaouic-Schwartz Seminar, École Polytechnique Paris, 1972–73, no. 28.

4. M. S. Baouendi and C. Goulaouic, *Étude de l'analyticité et de la régularité Gevrey pour une classe d'opérateurs elliptiques dégénérés*, Ann. Sci. École Norm. Sup. (4) **4** (1971), 31–46. MR **44** #4366.

5. ———, *Cauchy problems with characteristic initial hypersurface*, Comm. Pure Appl. Math. **26** (1973), 455–475.

6. ———, *Cauchy problems with multiple characteristics in spaces of regular distributions*, Uspehi Mat. Nauk **29** (1974), 70–76. (Russian)

7. ———, *Singular non linear Cauchy problems*, J. Differential Equations (to appear).

8. P. Bolley and J. Camus, *Hypoellipticité partielle et hypoanalyticité d'une classe d'opérateurs elliptiques et dégénérés*, Asterisque (to appear).

9. P. Bolley, J. Camus and B. Hanouzet, *Étude de l'analyticité et de la régularité Gevrey pour une classe de problèmes aux limites elliptiques et dégénérés*, Astérisque (to appear).

10. P. Bolley, J. Camus and B. Helffer, *Hypoellipticité partielle d'une classe d'opérateurs elliptiques et dégénérés*, C. R. Acad. Sci. Paris Sér. A-B **775** (1974) (to appear).

11. V. V. Grušin, *On a class of hypoelliptic operators*, Mat. Sb. **83 (125)** (1970), 456-473 = Math. USSR Sb. **12** (1970), 458–476. MR **43** #5158.

12. ———, *On a class of elliptic pseudo-differential operators degenerate on a submanifold*, Mat. Sb. **84 (126)** (1971), 163–195 = Math. USSR Sb. **13** (1971), 155–186. MR **44** #860.

13. B. Helffer and C. Zuily, *Non hypoellipticité d'une classe d'opérateurs différentiels*, C. R. Acad. Sci. Paris Sér. A-B **277** (1973), A1061–A1063.

14. L. Hörmander, *Linear partial differential operators*, Die Grundlehren der math. Wissenschaften, Band 116, Academic Press, New York; Springer-Verlag, Berlin, 1963. MR **28** #4221.

15. F. John, *On linear partial differential equations with analytic coefficients. Unique continuation of data*, Comm. Pure Appl. Math. **2** (1949), 209–253. MR **12**, 185.

16. L. Nirenberg, *An abstract form of the non linear Cauchy-Kowalewski theorem*, J. Differential Geometry **6** (1972), 561–576.

17. L. V. Ovsjannikov, *A singular operator in a scale of Banach spaces*, Dokl. Akad. Nauk SSSR **163** (1965), 819–822 = Soviet Math. Dokl. **6** (1965), 1025–1028. MR **32** #8164.

18. J. Persson and T. Yamanaka, *On an extension of Holmgren's uniqueness theorem*, Comment. Math. Univ. St. Paul. **22** (1973), 19–30.

19. F. Treves, *Ovchyannikov theorem and hyperdifferential operators*, Notas de Matermática, no. 46, Instituto de Mathematica Pura e Aplicada, Conselho Nacional de Pesquisas, Rio de Janeiro, 1968. MR **44** #7386.

20. T. Yamanaka, *Note on Kowalevskaya's system of partial differential equations*, Comment. Math. Univ. St. Paul. **9** (1961), 7–10. MR **23** #A3341.

PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Monotone Operators, Nonlinear Semigroups and Applications

Haim Brezis

Our purpose is to discuss some recent progress in the area of monotone operators and nonlinear semigroups as well as some applications to nonlinear partial differential equations. The first papers on this subject (by G. Minty, F. Browder, J. Leray and J. L. Lions, M. Visik; see [3], [8], [13] for a complete bibliography) were mainly concerned with the question of existence of a solution for nonlinear equations of Hammerstein type or elliptic type or parabolic type. These existence results are by now well known and I would like to describe further properties.

Essentially we are going to consider operators in a real Hilbert space H (for the latest news in Banach spaces, the reader is referred to the article of M. Crandall). Let A be a multivalued mapping from H into H, i.e., for every $u \in H$, Au is a subset of H; let $D(A) = \{u \in H;\ Au \neq \emptyset\}$ and $R(A) = \bigcup_{u \in H} Au$. A is monotone if

$$(f_1 - f_2, u_1 - u_2) \geqq 0, \quad f_1 \in Au_1, \quad f_2 \in Au_2,$$

and *maximal monotone* if A has no proper monotone extension. A characterization due to G. Minty asserts that a monotone operator A is maximal monotone iff $(I + \lambda A)$ is surjective for every $\lambda > 0$, in which case $J_\lambda = (I + \lambda A)^{-1}$ (the resolvent of A) is an everywhere defined contraction on H.

A fundamental class of maximal monotone operators consists of *subdifferentials of convex functions*. Let $\phi : H \to (-\infty, +\infty]$ be a convex l.s.c. function such that $\phi \not\equiv +\infty$; let $D(\phi) = \{u \in H;\ \phi(u) < +\infty\}$ and for $u \in D(\phi)$ let $\partial\phi(u) = \{f \in H;\ \phi(v) - \phi(u) \geqq (f, v - u),\ \forall\, v \in H\}$. Then $\partial\phi$ is maximal monotone.

We recall that when A is maximal monotone then, for every $u \in D(A)$, Au is closed and convex (A^0u denotes the projection of 0 on Au) and also $\overline{D(A)}$ is a convex set.

1. A strange property of $R(A + B)$. Let A and B be maximal monotone operators; we have always $R(A + B) \subset R(A) + R(B)$ and in general $R(A) + R(B)$ is a much larger set. However it turns out that in many cases we have "almost" equality. Here is a typical result in that direction.

THEOREM 1. *Let A be maximal monotone and let $B = \partial\phi$. Assume that*

$$\phi((I + \lambda A)^{-1}u) \leqq \phi(u), \qquad \lambda > 0,\ u \in H. \tag{1}$$

Then $\operatorname{Int}[R(A + B)] = \operatorname{Int}[R(A) + R(B)]$ *and* $\overline{R(A + B)} = \overline{R(A) + R(B)}$.

SKETCH OF THE PROOF. Given $f \in \operatorname{Int}[R(A) + R(B)]$ we would like to solve $Au + Bu \ni f$. It follows from (1) that $A + B$ is maximal monotone (see [2]). Therefore there exists u_ε such that $\varepsilon u_\varepsilon + Au_\varepsilon + Bu_\varepsilon \ni f$.

Using (1) it is easy to verify that $|Au_\varepsilon| \leqq C$, $|Bu_\varepsilon| \leqq C$ with C independent of ε. We have now to show that $|u_\varepsilon|$ remains bounded as $\varepsilon \to 0$. Instead of obtaining a bound on $|u_\varepsilon|$ (as usually done) we rely on the uniform boundedness principle. Let $r > 0$ be such that $B(f, r) \subset R(A) + R(B)$. Given $h \in H$ with $|h| < r$ we have $f + h \in Av + Bw$. By the monotonicity of A and B we get

$$(Au_\varepsilon - Av, u_\varepsilon - v) \geqq 0, \qquad (Bu_\varepsilon - Bw, u_\varepsilon - w) \geqq 0.$$

Therefore $(h, u_\varepsilon) \leqq (Av - Au_\varepsilon, v) + (Bw - Bu_\varepsilon, w) \leqq C'$ and so $|u_\varepsilon|$ remains bounded.

In case we assume just $f \in R(A) + R(B)$, then we can only prove that $\sqrt{\varepsilon}\,|u_\varepsilon|$ remains bounded and at the limit $f \in \overline{R(A + B)}$.

REMARK. The conclusion of Theorem 1 still holds true if (1) is replaced by one of the following assumptions:

(2) Both A and B are subdifferentials of convex functions and $A + B$ is maximal monotone.

(3) $D(B) = H$.

The last result can be used to solve equations of *Hammerstein type*. Let K be a maximal monotone operator with $D(K) = H$ and let $B = \partial\phi$ with $D(B) = H$; then $R(I + KB) = H$. Indeed for $f \in H$, the equation $u + KBu \ni f$ can be written as $Au + Bu \ni 0$ where $Au = -K^{-1}(u - f)$ and $R(A) = H$. So $\operatorname{Int} R(A + B) = \operatorname{Int}[R(A) + R(B)] = H$. A similar result has been proved in [5] for mappings in Banach spaces.

We now illustrate Theorem 1 by two examples.

EXAMPLE 1. Consider a periodic evolution equation of the form

$$du/dt + \mathscr{B}u = f(t) \quad \text{on } (0, T) \text{ with } u(0) = u(T) \tag{4}$$

where $u(t)$ takes its values in a Hilbert space $\mathscr{H}$ and $\mathscr{B}$ is the subdifferential of a convex function in $\mathscr{H}$. It is clear that if (4) has a solution, then necessarily

$$\frac{1}{T}\int_0^T f(t)\,dt \in \overline{\operatorname{conv}}\, R(\mathscr{B}) = \overline{R(\mathscr{B})}.$$

Conversely we have

THEOREM 2 (HARAUX). *Suppose*

$$\frac{1}{T}\int_0^T f(t)\,dt \in \operatorname{Int} R(\mathscr{B});$$

then (4) *has a solution.*

It suffices to apply Theorem 1 with $H = L^2(0, T; \mathscr{H})$, $Au = du/dt$, $D(A) = \{u \in H; u' \in H \text{ and } u(0) = u(T)\}$ and B is the canonical extension of $\mathscr{B}$ to H. It is clear that $R(A) + R(B) \supset \{f \in H;\ T^{-1}\int_0^T f(t)\,dt \in R(\mathscr{B})\}$.

EXAMPLE 2. Consider the nonlinear boundary value problem

$$\Delta u = 0 \quad \text{on } \Omega, \qquad \partial u/\partial n + \beta(u) \ni f \quad \text{on } \partial\Omega, \tag{5}$$

where Ω is a smooth bounded domain and β is a maximal monotone graph in $R \times R$. It is clear that if (5) has a solution then $|\partial\Omega|^{-1}\int_{\partial\Omega} f(\sigma)\,d\sigma \in \overline{R(\beta)}$. Conversely we have

THEOREM 3 (SCHATZMAN, HESS). *Suppose*

$$\frac{1}{|\partial\Omega|}\int_{\partial\Omega} f(\sigma)\,d\sigma \in \operatorname{Int} R(\beta);$$

then (5) *has a solution.*

REMARKS. (1) In case $D(\beta) = [0, +\infty)$, $\beta(r) = 0$ for $r > 0$, $\beta(0) = (-\infty, 0]$, the boundary condition in (5) can be written $u \geqq 0$, $\partial u/\partial n - f \geqq 0$, $u(\partial u/\partial n - f) = 0$ on $\partial\Omega$ and a solution exists provided $\int_{\partial\Omega} f(\sigma)\,d\sigma < 0$ (a similar result can be found in [15]).

(2) Theorems 2 and 3 are comparable to some results of Landesman and Lazer and L. Nirenberg (see [16]). However the techniques are totally different.

2. Evolution equations and nonlinear semigroups. Let A be a maximal monotone operator and consider the evolution equation

$$du/dt + Au \ni 0 \quad \text{on } [0, +\infty),\ u(0) = u_0. \tag{6}$$

We recall first a well-known result

THEOREM 4 (KATO [11], KOMURA [12]). *Given $u_0 \in D(A)$, there exists a unique solution of* (6) *such that $u(t) \in D(A)$ for all $t \geqq 0$ and u is Lipschitz continuous on* $[0, +\infty)$.

REMARK. The same method can be used to solve $du/dt + Au \ni f(t)$, $u(0) = u_0$, provided f is smooth enough. In addition $u(t)$ is differentiable from the right at every $t \geqq 0$ and $d^+u/dt + A^0u = 0$ for all $t \geqq 0$ (see Kato [11], Crandall and Pazy [9], Dorroh [10]). If $u_0, \hat{u}_0 \in D(A)$, the corresponding solutions $u(t)$ and $\hat{u}(t)$ satisfy $|u(t) - \hat{u}(t)| \leqq |u_0 - \hat{u}_0|$ for all $t \geqq 0$. Thus the mapping $u_0 \to u(t)$ can be extended by continuity to $\overline{D(A)}$. We denote the extension by $S(t)$; $S(t)$ is called the *semigroup generated by* $-A$.

A number of results about linear semigroups are still valid for nonlinear semigroups. For instance we have nonlinear analogues of the theorems of Hille-Yosida-Phillips and Trotter-Kato-Neveu.

THEOREM 5 (KOMURA [12], CRANDALL AND PAZY [9]). *Let C be a closed convex set in H and let $S(t)$ be a semigroup of contractions on C (i.e., $S(0) = I$, $S(t_1 + t_2) = S(t_1) \circ S(t_2)$, $|S(t)u - S(t)v| \leqq |u - v|$ and $|S(t)u - u| \to 0$ as $t \to 0$). Then there exists a unique maximal monotone operator A such that $\overline{D(A)} = C$ and $S(t)$ coincides with the semigroup generated by $-A$.*

THEOREM 6. *Let A_n, A be maximal monotone operators and let $S_n(t)$, $S(t)$ be the corresponding semigroups. The following properties are equivalent:*

(7) $\forall\, x \in D(A)$, $\exists\, x_n \in D(A_n)$ *such that* $x_n \to x$ *and* $S_n(t)\, x_n \to S(t)x$ *uniformly on bounded t intervals.*

(8) $\forall\, x \in D(A)$, $\forall\, \lambda > 0$, $(I + \lambda A_n)^{-1} x \to (I + \lambda A)^{-1} x$.

(9) $\forall\, x \in D(A)$, $\exists\, x_n \in D(A_n)$ *such that* $x_n \to x$ *and* $A_n^0 x_n \to A^0 x$.

Related results were obtained by Miyadera and Oharu, Brezis and Pazy, Benilan, Goldstein, Kurtz, etc.

3. Smoothing action and asymptotic behavior of semigroups generated by subdifferentials. In general when $u_0 \in \overline{D(A)}$, (6) has no "real" solution and $S(t)u_0$ represents a generalized solution of (6). But in case $A = \partial\phi$, $S(t)u_0$ turns out to be a "classical" solution of (6) even for $u_0 \in \overline{D(A)}$. More precisely

THEOREM 7. *Let $A = \partial\phi$ and let $u_0 \in \overline{D(A)}$. Then $u(t) = S(t)u_0 \in D(A)$ for all $t > 0$, $u(t)$ is Lipschitz continuous on every interval $[\delta\ ,\ +\infty)$ $(\delta > 0)$ and $u(t)$ satisfies (6). In addition one has*

$$|A^0 S(t)u_0| = \left|\frac{d^+u}{dt}(t)\right| \leqq \frac{1}{t}\,|u_0 - u(t)| \quad \text{for all } t > 0. \tag{10}$$

For the proof see [2] (estimate (10) is new).

Since $S(t)$ maps $\overline{D(A)}$ into $D(A)$ for positive t we can say that $S(t)$ has a smoothing action on the initial data. To illustrate the smoothing action, consider the following nonlinear heat equation

$$\begin{aligned} &\partial u/\partial t - \Delta u + \beta(u) \ni 0 \text{ on } \Omega x(0, +\infty), \\ &u = 0 \text{ on } \partial\Omega x(0, +\infty), \qquad u(x, 0) = u_0(x) \text{ on } \Omega, \end{aligned} \tag{11}$$

where β is a monotone function (or graph). Given $u_0 \in L^2(\Omega)$, the solution $u(\cdot, t)$ lies in the Sobolev space $H^2(\Omega)$ for every $t > 0$.

REMARKS. (1) In fact (11) has a stronger smoothing action. Starting with $u_0 \in L^1(\Omega)$ one can show that $u(\cdot, t) \in W^{2,p}(\Omega)$ for every $p < +\infty$ and $t > 0$. The proof is quite technical and is not a consequence of an abstract result about smoothing action in Banach spaces.

(2) When β is not smooth there is also an "unsmoothing" action of $S(t)$ (a typically nonlinear phenomenon, well known in variational inequalities): Even if $u_0 \in C^\infty$, it may happen that $u(\cdot, t) \notin C^2$.

Suppose now ϕ achieves its minimum and let $K = \{v \in H;\ \phi(v) = \operatorname{Min} \phi\}$. The trajectories $S(t)u_0$ are orthogonal to the level curves of ϕ (as in the steepest descent method) and it is natural to conjecture that $S(t)u_0$ converges to some limit in K

as $t \to +\infty$. So far it is not known whether the strong limit exists. There are only partial answers:

(a) For every $u_0 \in \overline{D(A)}$, $S(t)u_0$ converges weakly as $t \to +\infty$ to some limit in K (R. Bruck).

(b) If $(I + A)^{-1}$ is compact (i.e., maps bounded sets into compact sets), then $S(t)u_0$ converges strongly as $t \to +\infty$.

(c) If ϕ is even, then $S(t)u_0$ converges strongly as $t \to +\infty$ (R. Bruck).

4. Interpolation classes. Let A be maximal monotone; for $0 < \alpha \leqq 1$ and $1 \leqq p \leqq +\infty$ define

$$\mathscr{B}_{\alpha,p} = \left\{u_0 \in \overline{D(A)};\ \left|\frac{(I + tA)^{-1}u_0 - u_0}{t^\alpha}\right| \in L^p_* = L^p\left(0, 1; \frac{dt}{t}\right)\right\}.$$

When A is a linear operator the $\mathscr{B}_{\alpha,p}$'s coincide with the interpolation spaces between $D(A)$ and H of Lions and Peetre [14]. These intermediate classes (not spaces!) can be characterized in various ways (D. Brezis [1]):

(a) *The trace method.*

$$\mathscr{B}_{\alpha,p} = \{v(0); v \in C([0,1]; H) \text{ with } t^{1-\alpha}\left|dv/dt\right| \in L^p_*,\ t^{1-\alpha}\left|A^0 v\right| \in L^p_*\}.$$

(b) *The method K.* Let $K(t, u_0) = \operatorname{Inf}_{v \in D(A)} \{\left|v - u_0\right| + t\left|A^0 v\right|\}$; then

$$\mathscr{B}_{\alpha,p} = \{u_0 \in \overline{D(A)};\ t^{-\alpha}K(t, u_0) \in L^p_*\}.$$

(c) *The semigroup method.* Let $S(t)$ be the semigroup generated by $-A$; then

$$\mathscr{B}_{\alpha,p} = \left\{u_0 \in \overline{D(A)};\ \left|\frac{S(t)u_0 - u_0}{t^\alpha}\right| \in L^p_*\right\}.$$

To prove the last result one can use the following simple inequalities:

$$\begin{aligned} &\left|S(t)u_0 - u_0\right| \leqq 3\left|J_t u_0 - u_0\right|, \\ &\left|J_t u_0 - u_0\right| \leqq \frac{2}{t}\int_0^t \left|S(\tau)u_0 - u_0\right| d\tau. \end{aligned} \tag{12}$$

In case $A = \partial\phi$ one has also

$$\left|J_t u_0 - u_0\right| \leqq (1 + 2^{-1/2})\left|S(t)u_0 - u_0\right|. \tag{13}$$

Rephrasing this fact we can say that if we consider the singular perturbation problem $u_\varepsilon + \varepsilon A u_\varepsilon = u_0$ and the evolution problem $du/dt + Au = 0$, $u(0) = u_0$, then the rate of convergence of $\left|u_\varepsilon - u_0\right|$ as $\varepsilon \to 0$ is the same as the rate of $\left|u(t) - u_0\right|$ as $t \to 0$. When $Au = -\Delta u + \left|u\right|^k \operatorname{sign} u$ one can describe $\mathscr{B}_{\alpha,p}$ in terms of Besov and Lorentz spaces.

5. Compact supports. We conclude with a surprising property which is specific to nonlinear problems and has no analogue in the linear theory. Consider the evolution equation

$$du/dt + Au \ni f \quad \text{on } (0, +\infty), \qquad u(0) = u_0. \tag{14}$$

Question. Under what conditions does u have a compact support, i.e., $u(t) \equiv 0$ for $t \geqq T$? A necessary condition is that $f(t) \in A0$ for $t \geqq T$, but in general this is not a sufficient condition. If A is linear it just means that f has a compact support, and of course this will not imply that u has a compact support. When A is multivalued one can give a sufficient condition which is very close to the necessary condition.

THEOREM 8. *Suppose there is some* $\rho(t) \geqq 0$ *such that* $B(f(t), \rho(t)) \subset A0$ *for* $t \geqq t_0$ *and* $\int_{t_0}^{\infty} \rho(t)\, dt = \infty$. *Then the solution of* (14) *has a compact support.*

We illustrate this fact by an example. Suppose $g(t)$ is the given trajectory of a gangster chased by a policeman $p(t)$. The strategy of the policeman is simple: He runs with speed V (as fast as he can!) towards $g(t)$. Thus we have

$$\frac{dp}{dt} = V \frac{g(t) - p(t)}{|g(t) - p(t)|} \quad \text{as long as } p(t) \neq g(t),$$

i.e., $dp/dt \in A(g(t) - p(t))$ where $Au = Vu/|u|$ for $u \neq 0$ and $A0 = \bar{B}(0, V)$ (so that A is maximal monotone). Let $u(t) = g(t) - p(t)$, and we get $du/dt + Au \ni dg/dt$. Here Theorem 8 tells us that if $|dg(t)/dt| \leqq V$ and $\int_0^{\infty} V - |dg(t)/dt|\, dt = +\infty$, then p reaches g in a finite time.

A similar property holds true in nonlinear parabolic equations even though we cannot apply Theorem 8. Consider for example

$$(15) \qquad \frac{\partial t}{\partial u} - \Delta u + \beta(u) \ni f \quad \text{on } R^n \times (0, +\infty), \qquad u(x, 0) = u_0(x) \text{ on } R^n,$$

where β is a monotone function with a jump at 0, $\beta(0) = [\gamma^-, \gamma^+]$.

In [6] we show that if $\gamma^- + \varepsilon \leqq f \leqq \gamma^+ - \varepsilon$ for some $\varepsilon > 0$, and if u_0 has a compact support, then (15) has a solution with compact support (both in x and t). If u_0 does not have a compact support, but $u_0 \to 0$ at infinity then, for every $t > 0$, $u(\cdot, t)$ has a compact support in x. In other words the support of u "shrinks" instantaneously (in sharp contrast with what happens for the linear heat equation!). Similar results for elliptic problems are considered in [4]. The original motivation for looking at the compact support property came from the study of the flow past a given profile. In [7] we show that in the hodograph plane the problem can be stated as a free boundary value problem which turns out to be solvable by the techniques of variational inequalities. The size of the support corresponds to the maximum velocity of the flow and plays an important role.

References

1. D. Brezis, *Classes d'interpolation associées à un opérateur monotone*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), 1553–1556.

2. H. Brezis, *Monotonicity methods in Hilbert spaces...*, Contributions to Nonlinear Functional Analysis (E. Zarantonello ed.), Academic Press, New York, 1971.

3. ———, *Opérateurs maximaux monotones et semigroupes de contractions dans les espace de Hilbert*, Math. Studies, no. 5, North-Holland, Amsterdam, 1973.

4. ———, *Solutions with compact support of variational inequalities*, Uspehi Mat. Nauk **29** (1974), 103–108.

5. H. Brézis and F. Browder, *Équations intégrales non linéaires du type Hammerstein*, C. R. Acad. Sci. Paris **279** (1974), 1–2.

6. H. Brézis and A. Friedman, *Estimates on the support of solutions of parabolic variational inequalities* (to appear).

7. H. Brézis and G. Stampacchia, *Une nouvelle méthode pour l'étude d'écoulements stationnaires*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A129-A132. MR **47** #4521.

8. F. E. Browder, *Existence theorems for nonlinear partial differential equations*, Proc. Sympos. Pure Math., vol. 16, Amer. Math. Soc., Providence, R. I., 1970, pp. 1–60. MR **42** #4855.

9. M. G. Crandall and A. Pazy, *Semi groups of nonlinear contractions and dissipative sets*, J. Functional Analysis **3** (1969), 376–418; *On accretive sets in Banach spaces*, J. Functional Analysis **5** (1970), 204–217. MR **39** #4705; **41** #4201.

10. J. R. Dorroh, *A nonlinear Hille-Yosida-Phillips theorem*, J. Functional Analysis **3** (1969), 345–353. MR **39** #2019.

11. T. Kato, *Accretive operators and nonlinear evolution equations in Banach spaces*, Proc. Sympos. Pure Math., vol. 18, part I, Amer. Math. Soc., Providence, R. I., 1970, pp. 138–161. MR **42** #6663.

12. Y. Komura, *Nonlinear semi groups in Hilbert space*, J. Math. Soc. Japan **19** (1967), 493–507; ibid. **21** (1969), 375–402. MR **35** #7176.

13. J. -L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969. MR **41** #4326.

14. J. -L. Lions and J. Peetre, *Sur une classe d'espaces d'interpolation*, Inst. Hautes Études Sci. Publ. Math. No. 19 (1964), 5–68. MR **29** #2627.

15. J. -L. Lions and G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math. **20** (1967), 493–519. MR **35** #7178.

16. L. Nirenberg, *Generalized degree and nonlinear problems*, Contributions to Nonlinear Functional Analysis (Zarantonello ed.), Academic Press, New York, 1971.

UNIVERSITÉ PARIS VI
PARIS 5e, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Semigroups of Nonlinear Transformations and Evolution Equations*

Michael G. Crandall

Over the past several years a substantial theory concerning certain nonlinear evolution equations in general Banach spaces has been developed which recent investigations indicate will provide a convenient framework for the treatment of many interesting problems in nonlinear partial differential equations. Here we give a brief report on some of these developments.

No attempt is made to state the sharpest known results as this would involve distracting technicalities. The heart of the matter is illustrated by more easily stated cases which are sufficient for most applications.

A. The general theory. Throughout we denote by $\mathscr{I}$ a subinterval of $\boldsymbol{R}$, by X a Banach space with norm $\|\ \|$, by A a mapping $A : X \to 2^X$ and by f an element of $L^1_{\text{loc}}(\mathscr{I}, X)$. We consider the equation

$$(\mathrm{E}_f) \qquad du/dt + Au \ni f.$$

A function $u:\mathscr{I} \to X$ is a *strong solution* of (E_f) on $\mathscr{I}$ if $u \in C(\mathscr{I}, X)$, u is absolutely continuous and differentiable a.e. on compact subintervals of int $\mathscr{I}$ and $du(t)/dt + Au(t) \ni f(t)$ a.e. on $\mathscr{I}$. A *weak solution* u of (E_f) on $\mathscr{I}$ is a function $u \in C(\mathscr{I}, X)$ for which there is a sequence $\{[u_n, f_n]\}$ such that u_n is a strong solution of (E_{f_n}) on $\mathscr{I}$ and $[u_n, f_n] \to [u, f]$ in $C(\mathscr{I}, X) \times L^1_{\text{loc}}(\mathscr{I}, X)$.

For existence results we will have to restrict A. Let $\tau: X \times X \to \boldsymbol{R}$ be defined by

$$\tau(x, y) = \lim_{\lambda \to 0+} \lambda^{-1}(\|x + \lambda y\| - \|x\|) = \inf_{\lambda > 0} \lambda^{-1}(\|x + \lambda y\| - \|x\|).$$

Then τ is upper semicontinuous and $\|x + \lambda y\| \geqq \|x\|$ for $\lambda > 0$ iff $\tau(x, y) \geqq 0$. The

*This work was supported in part by NSF grant GP-38519.

operator A is *accretive* if $\tau(x_1 - x_2, y_1 - y_2) \geqq 0$ whenever $y_i \in Ax_i$ for $i = 1, 2$. Equivalently, A is accretive iff the operator $J_\lambda = (I + \lambda A)^{-1}$ defined for $\lambda > 0$ by $J_\lambda(x + \lambda y) = x$ if $y \in Ax$ is a contraction mapping $R(I + \lambda A) = \{x + \lambda y : x \in X, y \in Ax\}$ into X. The following is a basic existence theorem:

THEOREM I. *Let A be accretive, $x_0 \in \overline{D(A)}$ and $R(I + \lambda A) \supset \overline{D(A)}$ for $\lambda > 0$. Then the unique solution $u_\varepsilon(t) = J_\varepsilon^{([t/\varepsilon]+1)} x_0$ of*

$$(u_\varepsilon(t) - u_\varepsilon(t - \varepsilon))/\varepsilon + Au_\varepsilon(t) \ni 0 \quad \text{for } t \geqq 0, \qquad u_\varepsilon(t) = x_0 \quad \text{for } t < 0, \tag{1}$$

converges locally uniformly as $\varepsilon \downarrow 0$. Moreover, if $S_A(t)x_0 = \lim_{\varepsilon \to 0+} J_\varepsilon^{([t/\varepsilon]+1)} x_0$ for $x_0 \in \overline{D(A)}$, $t \geqq 0$, then $S_A(t)$ is a semigroup of contractions on $\overline{D(A)}$.

This result was proved in Crandall and Liggett **[14]** by explicit estimation of the quantities $\|J_\mu^n x - J_\lambda^m x\|$ for $x \in D(A)$, $\mu, \lambda > 0$, $n, m \in \mathbf{Z}^+$. Of course, $u(t) = S_A(t)x_0$ is supposed to be a solution of the Cauchy problem (E_0), $u(0) = x_0$. If X is reflexive and A is closed, $u(t)$ is a strong solution of (E_0) on $[0, \infty)$ for $x_0 \in D(A)$ and a weak solution for $x_0 \in \overline{D(A)}$. The proof uses the inequalities

$$\|u(t) - x\| - \|u(s) - x\| \leqq \int_s^t \tau(x - u(\alpha), y)\, d\alpha \tag{2}$$

for $0 \leqq s \leqq t$, $x \in D(A)$ and $y \in Ax$. Inequalities equivalent to (2) were established in **[14]** under an additional technical condition which was removed by Miyadera **[22]**. Unfortunately, $S_A(t)x_0$ may fail to be differentiable for any $t \geqq 0$, $x_0 \in \overline{D(A)}$, if X is not reflexive and it is not known if $S_A(t)x_0$ is a weak solution of (E_0). A notion of solution of (E_f) was given by Benilan for which $S_A(t)x_0$ is the unique solution of (E_0), $u(0) = x_0$. Generalizing (2), u is called an *integral solution* of (E_f) on $\mathscr{I}$ if $u \in C(\mathscr{I}, X)$ and

$$\|u(t) - x\| - \|u(s) - x\| \leqq \int_s^t \tau(x - u(\alpha), y - f(\alpha))\, d\alpha \tag{3}$$

whenever $s \leqq t$, $x, t \in \mathscr{I}$ and $y \in Ax$. It is easily shown that weak solutions of (E_f) are integral solutions if A is accretive. Benilan's uniqueness result states that if $u \in C(\mathscr{I}, X)$ is obtained as the uniform limit of solutions of implicit difference schemes converging to the problem (E_f), then

$$\|u(t) - v(t)\| - \|u(s) - v(s)\| \leqq \int_s^t \tau(v(\alpha) - u(\alpha), g(\alpha) - f(\alpha))\, d\alpha \tag{4}$$

whenever v is an integral solution of (E_g) on $\mathscr{I}$, etc. This strongly generalizes (2). See **[3]** for the precise formulations and proof as well as the related notion of a "good" solution which is defined *via* (4). In particular, (1) is an implicit scheme of the appropriate kind and $S_A(t)x$ is the unique integral solution of (E_0) satisfying $u(0) = x$.

Existence theorems for (E_f) are easily obtained from Theorem I. For example, if A is accretive and $R(I + A) = X$ (i.e., A is *m-accretive*), $f \in L^1([0, T], X)$, $x \in \overline{D(A)}$,

then (E_f) has a unique integral solution u on $[0, T]$ satisfying $u(0) = x$ (and the inequalities (4)). This result was established in Benilan [2]. The existence is also proved in Crandall and Pazy [15], while investigating a more general time dependence than $A(t) = A - f(t)$.

The above concerns just one aspect of the general theory important for applications. Namely, it provides existence theorems and a notion of solution for problems of the form (E_f) which may not admit any strong solutions. Among other types of results useful in applications are the convergence and perturbation theorems. Convergence theorems deal with the continuity of solutions of (E_f) in A and have been proved by many authors (see, e.g., [3], [9], [19], [23], [25]). We take an example from [3]: Let A_n be m-accretive, $f_n \in L^1([0, T], X)$, $x_n \in D(A_n)$ and u_n be the integral solution of (E_{f_n}) satisfying $u_n(0) = x_n$. Let $x_n \to x_0$ and $f_n \to f_0$ in $L^1([0, T], X)$. If $\xi_n \in D(A_n)$, $\xi_n \to \xi_0 \in D(A_0)$ implies $(I + \lambda(A_n - f(t)))^{-1}\xi_n \to (I + \lambda(A_0 - f(t)))^{-1}\xi_0$ for $\lambda > 0$ a.e. $t \in [0, T]$, then $u_n \to u_0$ uniformly on $[0, T]$. Useful general perturbation theorems valid in unrestricted spaces are rare. A nice one states that $A + B$ is m-accretive if A is m-accretive and $B : X \to X$ is continuous and accretive. This was proved by Webb [29] if A is linear, Barbu [1], and the simplest proof is in Pierre [26]. For another result see [11]. It is probably the case that most useful perturbation theorems in nonreflexive spaces will involve hypotheses (maximum principles, etc.) reflecting a particular class of applications as in [5], [10], and [18].

There are many individual contributors to, and interesting results in, this area we are not able to mention. In particular, we have limited ourselves to results which are valid in nonreflexive spaces X. A good bibliography for the subject through 1971 and a commentary is available in the book of Brezis [8]. The best source covering arbitrary Banach spaces is the paper of Benilan [3], which treats the topics we have touched upon and much more in considerable generality. A couple of further remarks may make the earlier literature more intelligible. Considerable interest was attached to the nature of the map $A \to S_A$ and infinitesimal generators of semigroups of contractions. A report on the map $A \to S_A$ is given in [12]. Benilan has shown that the situation here is improved if one deals with the map assigning to A the solution operator associated with the family of problems (E_f) as f is varied rather than just S_A (at least if A is m-accretive). The question of whether there is a useful notion of an infinitesimal generator such that each semigroup of contractions has a nontrivial one is not settled. (It is known that X-valued notions are too restrictive.) See, however, Neuberger [24]. As a carry-over from studies in Hilbert spaces, $\frac{1}{2}\| \ \|^2$ and its directional derivatives appear in many papers rather than $\| \ \|$ and the function called τ here. It is now clear that $\| \ \|$ and τ are the most convenient to use in general spaces, but notation and practice here vary and each paper must be checked for its conventions. Finally, we remark that other notions are associated with the term "weak solution" besides the one used here. See, e.g., the discussion in Oharu [25] and the works referenced there.

B. Applications. The study of applications of the general theory is very young and

developing rapidly. Thus we merely intend to indicate something of its current nature and scope. Aside from the first case below, which is discussed in more detail, we do so by means of reference to selected examples.

We begin with the conservation law

$$\text{(CL)} \qquad u_t + \sum_{i=1}^{N} \frac{\partial}{\partial x_i} \varphi_i(u) = u_t + \varphi(u)_x = 0$$

where $\varphi : \boldsymbol{R} \to \boldsymbol{R}^N$. It is well known (see, e.g., [21]) that the Cauchy problem (CL), $u(0, x) = u_0(x)$, does not admit global (in $t \geqq 0$) continuous solutions in general, no matter how smooth φ and u_0 may be; and that if (CL) is understood in the sense of distributions then solutions are not unique. Kruzkov [20] gave a notion of solution of equations generalizing (CL) and proved existence and uniqueness theorems for the Cauchy problem. In the case of (CL), Kruzkov interprets the equation as the family of inequalities

$$\text{(5)} \qquad \int_0^\infty \int_{\boldsymbol{R}^N} \Big(|u(t, x) - k| f_t(t, x) + \operatorname{sign}(u(t, x) - k) \sum_{i=1}^{N} (\varphi_i(u(t, x)) - \varphi_i(k)) f_{x_i}(t, x) \Big) dx\, dt \geqq 0$$

for $f \in \mathscr{D}((0, \infty) \times \boldsymbol{R}^N)$, $f \geqq 0$, $k \in \boldsymbol{R}$. The relation of (CL) to the abstract theory was suggested in Quinn [27] and established in [13] (and in a special case by Flaschka [16]). If $u, \varphi(u), \psi \in L^1_{\text{loc}}(\boldsymbol{R}^N)$ we say $\operatorname{div}_K \varphi(u) = \psi$ if

$$\text{(6)} \qquad \int \operatorname{sign}(u(x) - k) \Big\{ \sum_{i=1}^{N} (\varphi_i(u(x)) - \varphi_i(k)) f_{x_i}(x) + \psi(x) f(x) \Big\} dx \geqq 0$$

for $f \in \mathscr{D}(\boldsymbol{R}^N)$, $f \geqq 0$, $k \in \boldsymbol{R}$. Let $D(A) = \{u \in L^1(\boldsymbol{R}^N) \cap L^\infty(\boldsymbol{R}^N) : \operatorname{div}_K \varphi(u) = \psi$ for some $\psi \in L^1(\boldsymbol{R}^N)\}$ and $Au = \operatorname{div}_K \varphi(u)$ if $u \in D(A)$. It was proved in [13] that $R(I + A) \supseteq L^1 \cap L^\infty$ if φ is continuous. If also $\limsup_{r \to 0} |\varphi(r)|/|r| < \infty$, it was shown that A is accretive in $L^1(\boldsymbol{R}^N)$. This last condition was weakened to $\lim_{r \to 0} |\varphi(r)|/|r|^{(N-1)/N} = 0$ by Benilan who also proved that $D(A)$ is dense in $L^1(\boldsymbol{R}^N)$. It then follows that the closure $\bar{A}$ of A is m-accretive in $L^1(\boldsymbol{R}^N)$. Moreover, the evolution associated with the operator $\bar{A}$ agrees with Kruzkov's interpretation (5) of (CL). This example shows, in particular, that the difficulties associated with interpreting the solutions of the abstract equation are not due to an irrelevant generality.

The works of Benilan [4] and Oharu [25] treat the technically more complex case where $\varphi = \varphi(x, u)$ depends on x. It should be mentioned that all these treatments use Kruzkov's ingenious arguments.

Other equations give rise to m-accretive operators in L^1 spaces. Examples of a nonlinear hyperbolic system and a nonlinear diffusion equation were given in [12]. The latter topic has had considerable development. Initial boundary value problems for equations of the form $u_t - \Delta\varphi(u) = 0$, where φ is nondecreasing and the boundary conditions are nonlinear, are treated in [2]. The results of Brezis and Strauss [10] give rise to a related class of examples. Consideration of the A associated with $u_t - \Delta\varphi(u) = 0$ in $L^1(\boldsymbol{R}^N)$ leads to interesting problems in anaylsis [7].

Kurtz [19] uses the theory in an interesting way. Let $\mathscr{L} = \{u \in L^1(\boldsymbol{R}) : u \geqq 0\}$. For each $\alpha > 0$ there is an accretive operator A_α in $[L^1(\boldsymbol{R})]^2$ with $\overline{D(A_\alpha)} = \mathscr{L} \times \mathscr{L}$, and $R(I + \lambda A_\alpha) \supset \mathscr{L} \times \mathscr{L}$ for $\lambda > 0$ which governs the evolution in $\mathscr{L} \times \mathscr{L}$ associated with the system $u_t + \alpha u_x + \alpha^2(u^2 - v^2) = 0$, $v_t - \alpha v_x + \alpha^2(v^2 - u^2) = 0$. Kurtz shows that there is also such an operator A in $\mathscr{L}$ corresponding to the equation $w_t - \frac{1}{4}(\log w)_{xx} = 0$. (This last equation does not fall within the scope of the results mentioned above due to the fact that 0 is not in the domain of the log function.) Finally, using his convergence theorem, he proves $\lim_{\alpha \to \infty} S_{A_\alpha}(\alpha t)[u_0, u_0] = [S_A(t)u_0, S_A(t)u_0]$ for $u_0 \in \mathscr{L}$.

Spaces of continuous functions are the natural setting for some problems. Equations of the form $\varphi(u_t) - \Delta u = 0$ give rise to m-accretive operators in $C_0(\bar{\Omega})$ and $L^\infty(\Omega)$ as is shown in Benilan [6] and Konishi [17]. Tamburro [28] obtains an m-accretive operator A in the space of bounded uniformly continuous functions on $\boldsymbol{R}^N$ corresponding to the equation $u_t + f(\nabla u) = 0$ where $f : \boldsymbol{R}^N \to \boldsymbol{R}$ is convex. Similarly, the expression $f(\nabla u) - \Delta u$ defines an m-accretive operator in this space even if f is not convex.

For other examples, the reader can consult the works of Konishi referenced in [17]. As mentioned, the study of applications is just beginning, and there are many open questions concerning both known examples and what other equations might fall within the scope of the theory (or suitable extensions). For instance, (1.3) and (1.4) of Lax [21] correspond to m-accretive operators. Does (1.5)? In the nonlinear hyperbolic system of [12], the associated accretive operator is described as the closure of a relatively simple operator. What is the domain of this closure? And so on. There is much to be done here.

ADDED IN PROOF. The bibliography and text of the paper *On the relation of the operator* $\partial/\partial s + \partial/\partial \tau$ *to evolution governed by accretive operators*, by M. G. Crandall and L. C. Evans, which will appear in the Israel J. Math., provide access to some interesting results which have appeared since the preparation of this paper.

References

1. V. Barbu, *Continuous perturbations of nonlinear m-accretive operators in Banach spaces*, Boll. Un. Mat. Ital. 6 (1972), 270–278.
2. Ph. Benilan, Thesis, Orsay 1972.
3. ———, *Equations d'evolution dans un espace de Banach quelconque*, Ann. Inst. Fourier (to appear.)
4. ———, *Equations quasi-linéaires du* 1er *ordre*, Israel J. Math. (to appear).
5. ———, *Principe du maximum et perturbation d'opérateurs accrétifs dans* $L^1(\Omega)$ (to appear).
6. ———, *Sur le problème* $\Delta u \in \gamma(\cdot, -\partial u/\partial t)$ *dans* $L^\infty(\Omega)$ (in preparation).
7. Ph. Benilan, H. Brezis and M. G. Crandall, *A semilinear elliptic equation in* $L^1(\boldsymbol{R}^N)$, Ann. Scuola Norm. Sup. Pisa (to appear).
8. H. Brezis, *Opérateurs maximaux monotones et semigroupes de contractions dans les espace de Hilbert*, Math. Studies 5, North-Holland, Amsterdam, 1973.
9. H. Brezis and A. Pazy, *Convergence and approximation of semigroups of nonlinear operators in Banach spaces*, J. Functional Analysis 9 (1972), 63–74. MR **45** #2529.

10. H. Brezis and W. Strauss, *Semilinear elliptic equations in L^1*, J. Math. Soc. Japan **25** (1973), 565–590.

11. B. Calvert and K. Gustafson, *Multiplicative perturbation of nonlinear m-accretive operators*, J. Functional Analysis **10** (1972), 149–157.

12. M. G. Crandall, in *Contributions to nonlinear functional analysis*, E. Zarantanello, editor, Academic Press, New York, 1971, pp. 157–179.

13. ———, *The semigroup approach to first order quasilinear equations in several space variables*, Israel J. Math. **12** (1972), 108–132. MR **47** #5473.

14. M. G. Crandall and T. M. Liggett, *Generation of semi-groups of nonlinear transformations on general Banach spaces*, Amer. J. Math. **93** (1971), 265–298. MR **44** #4563.

15. M. G. Crandall and A. Pazy, *Nonlinear evolution equations in Banach spaces*, Israel J. Math. **11** (1972), 57–94. MR **45** #9214.

16. H. Flaschka, Carnegie-Mellon Institute Report, 72–79.

17. Y. Konishi, *On the nonlinear semi-groups associated with $u_t=\Delta\beta(u)$ and $\varphi(u_t)=\Delta u$*, J. Math. Soc. Japan **25** (1973), 622–628.

18. ———, *A remark on fluid flows through porous media*, Proc. Japan Acad. **49** (1973), 102–105.

19. T. Kurtz, *Convergence of sequences of semigroups of nonlinear operators with an application to gas kinetics*, Trans. Amer. Math. Soc. **186** (1973), 259–272.

20. S. N. Kruzkov, *First order quasilinear equations in several independent variables*, Mat. Sb. **81 (123)** (1970), 228–255 = Math. USSR Sb. **10** (1970), 217–243. MR **42** #2159.

21. P. Lax, *Nonlinear partial differential equations of evolution*, Proc. Internat. Congress Math. (Nice, 1970), vol. 2, Gauthier-Villars, Paris, 1971, pp. 831–840.

22. I. Miyadera, *Some remarks on semigroups of nonlinear operators*, Tôhoku Math. J. (2) **23** (1971), 245–258. MR **45** #5805.

23. I. Miyadera and S. Oharu, *Approximation of semigroups of nonlinear operators*, Tôhoku Math. J. (2) **22** (1970), 24–47. MR **41** #7479.

24. J. Neuberger, *Lie generators for one parameter semigroups of transformations*, J. Reine Angew. Math. **258** (1973), 133–136. MR **46** #9796.

25. S. Oharu, *On the semigroup approach to nonlinear partial differential equations of first order*. I (to appear).

26. M. Pierre, *Perturbations localement Lipschitziennes et continues d'opérateurs m-accrétifs* (to appear).

27. B. K. Quinn, *Solutions with shocks: An example of L_1-contractive semigroup*, Comm. Pure Appl. Math. **24** (1971), 125–132. MR **42** #6428.

28. M. Tamburro, Thesis, University of California, Los Angeles, 1974.

29. G. Webb, *Continuous nonlinear perturbations of linear accretive operators in Banach spaces*, J. Functional Analysis **10** (1972), 191–203.

UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Applications of Fourier Integral Operators

J. J. Duistermaat

Fourier integral operators, for the calculus of which I refer to Hörmander [17], have been applied in essentially two ways: as similarity transformations and in the description of the solutions of genuinely nonelliptic (pseudo-) differential equations.

The first application is based on the observation of Egorov [12] that if P, resp. Q, is a pseudo-differential operator with principal symbol equal to p, resp. q, and $P \circ A = A \circ Q$ for an elliptic Fourier integral operator A defined by the homogeneous canonical transformation C, then $p = q \circ C$. This idea, or rather its local version in conic open subsets of the cotangent bundle of the manifolds on which the operators are defined, is a much more powerful tool for bringing operators locally into standard form than merely by coordinate changes in the base space. It has been used not only to reduce the study of wide classes of operators to simple ones like $\partial/\partial x_1, \partial/\partial x_1 + i\partial/\partial x_2, \partial/\partial x_1 + ix_1\partial/\partial x_2$, but also in more subtle problems it has been a helpful trick. A rather complete impression of this sort of application can be obtained by looking at the papers of Egorov [13], [14], Nirenberg and Trèves [26], Hörmander [18], Duistermaat and Hörmander [7], Sato, Kawai and Kashiwara [27], Duistermaat and Sjöstrand [8], Sjöstrand [28], Boutet de Monvel [3], and Weinstein [29].

In this respect the following conjecture of Singer seems interesting. Let $L^m(X)$ denote the space of pseudo-differential operators of order m, $L^{-\infty}(X)$ the space of smoothing operators. I restrict here to operators for which the total symbol has an asymptotic expansion in homogeneous terms of integer order. If A is an elliptic Fourier integral operator defined by a canonical transformation C: $T^*Y\backslash 0 \to T^*X\backslash 0$, then $P \mapsto A^{-1}PA$ is an isomorphism of filtered algebras:

$$\bigcup_{m\in\mathbf{Z}} L^m(X)/L^{-\infty}(X) \to \bigcup_{m\in\mathbf{Z}} L^m(Y)/L^{-\infty}(Y).$$

If moreover A is invertible, then $P \mapsto A^{-1}PA$ is an isomorphism of filtered algebras: $\bigcup_{m\in\mathbf{Z}} L^m(X) \to \bigcup_{m\in\mathbf{Z}} L^m(Y)$.

Now the conjecture is that conversely each isomorphism of filtered algebras of pseudo-differential operators is equal to conjugation by a Fourier integral operator. In this direction Singer and I have proved:

THEOREM. *Let X be compact, $H^1(T^*X\backslash 0, \boldsymbol{C}) = 0$. Then:*

(a) *Any isomorphism of filtered algebras*

$$\bigcup_{m\in\mathbf{Z}} L^m(X)/L^{-\infty}(X) \to \bigcup_{m\in\mathbf{Z}} L^m(Y)/L^{-\infty}(Y)$$

is either equal to conjugation by an elliptic Fourier integral operator, or by one preceded by the automorphism of $\bigcup_{m\in\mathbf{Z}} L^m(X)/L^{-\infty}(X)$ sending the symbol $\sum_j p_j(x,\xi)$ into the symbol $\sum_j e^{-\pi\cdot i\cdot j}\, p_j(x,-\xi)$. Here $p_j(x,\xi)$ denotes the homogeneous term of degree j in the asymptotic expansion of the total symbol.

(b) *Any isomorphism of filtered algebras: $\bigcup_{m\in\mathbf{Z}\cup\{-\infty\}} L^m(X) \to \bigcup_{m\in\mathbf{Z}\cup\{-\infty\}} L^m(Y)$ is equal to conjugation by an invertible continuous linear operator $A: C^\infty(Y) \to C^\infty(X)$.*

(c) *If A is an invertible continuous linear operator: $C^\infty(Y) \to C^\infty(X)$ and $A^{-1}PA \in L^m(Y)$ for every $P \in L^m(X)$, all $m \in \mathbf{Z}$, then A is an elliptic Fourier integral operator.*

The second use of Fourier integral operators, namely as solution operators, goes back to the historical origin of their calculus, because in their local representations

$$(Au)(x) = \int e^{i\varphi(x,y,\vartheta)}\, a(x,y,\vartheta)u(y)\, dy\, d\vartheta; \tag{1}$$

the integrands $e^{i\varphi(x,y,\vartheta)}a(x,y,\vartheta)$ are the asymptotic oscillatory solutions known from geometrical optics and the W.K.B. method in quantum mechanics. Continuous superposition of such waves resembles the construction of Huygens [**19**], and the observation that the major contributions only come from the places where the phase function φ is stationary as a function of ϑ is a counterpart of his well-known principle.

Lax [**21**] and Ludwig [**22**] showed that the solution operators for the Cauchy problem for *strictly* hyperbolic equations have local representations like (1), leading among others to results about the propagation of singularities. (For the case of characteristics of constant multiplicity, see Chazarain [**4**].)

Then in [**16**] Hörmander developed a local theory of Fourier integral operators to give a description for small $|t|$ of the unitary operator $U(t) = e^{-itP}$ = solution of the hyperbolic equation $(i^{-1}\partial/\partial t + P)\circ U = 0$, $U(0) = I$. Here P is a positive elliptic pseudo-differential operator of order 1 on a compact manifold X. From it he obtained the estimate

$$\#\{j;\lambda_j \leqq \lambda\} = (2\pi)^{-n} \operatorname{vol}(B^*X)\cdot\lambda^n + O(\lambda^{n-1}), \qquad \lambda \to \infty, \tag{2}$$

for the spectrum $\lambda_1 \leqq \lambda_2 \leqq \cdots$ of P. Here $B^*X = \{(x,\xi) \in T^*X; p(x,\xi) \leqq 1\}$ and p denotes the principal symbol of P. The improvement over previous results was

that the error term is the best possible for general operators P.

However from the global calculus of [17] it follows that U, regarded as an operator from $C^\infty(X)$ to $C^\infty(\boldsymbol{R} \times X)$, is a Fourier integral operator of order $-\frac{1}{4}$ defined by the canonical relation

$$C = \{(((t, x), (\tau, \xi)), (y, \eta)); \tau + p(x, \xi) = 0, (x, \xi) = \Phi^t(y, \eta)\}. \tag{3}$$

Here Φ^t is the time t flow of the Hamilton vector field H_p ($= \sum_j \partial p/\partial \xi_j \cdot \partial/\partial x_j - \partial p/\partial x_j \cdot \partial/\partial \xi_j$ on local coordinates) defined by p.

In particular $U(t)$ is for each t a Fourier integral operator of order 0 defined by the canonical transformation Φ^t, a fact which can also be proved by first showing that $e^{itP}Qe^{-itP} \in L^m(X)$ if $Q \in L^m(X)$ and then applying the theorem above. This shows also that $t \mapsto e^{-itP}$ cannot be a smooth family of Fourier integral operators if the order of P exceeds 1, because then

$$\frac{d}{dt}(e^{itP}Qe^{-itP})\big|_{t=0} = i[P, Q]$$

is of order $> m$ for most $Q \in L^m(X)$, so the order of $e^{itP}Qe^{-itP}$, if it were pseudo-differential operators, would blow up immediately.

The function

$$t \mapsto \hat{\sigma}(t) \underset{\text{def}}{=} \text{Trace } U(t) = \sum_j e^{-it\lambda_j} \tag{4}$$

is a tempered distribution on $\boldsymbol{R}$ and from the global characterization of U as a Fourier integral operator it follows that it can only have singularities at the periods of periodic H_p-solution curves. The singularity at a period T can be tested by multiplying $\hat{\sigma}$ with a smooth cut-off function $\hat{\rho}$, having its support in a small neighborhood of T, and then investigating the asymptotic behaviour of the inverse Fourier transform

$$\sum_j \rho(\lambda - \lambda_j) = (\rho * \sigma)(\lambda) = (2\pi)^{-1} \int e^{i\lambda t}\, \hat{\rho}(t)\hat{\sigma}(t)\, dt \tag{5}$$

as $\lambda \to \infty$.

The analysis of the singularity at $t = 0$ gives back the asymptotic expansion (2) of Hörmander and implies (see [10]) the Minakshisundaram-Pleijel formula for $\sum_j \exp(-\lambda_j^2 z)$ (asymptotics for $z \searrow 0$), as well as the statements about the poles of the ζ-function $\sum \lambda_j^{-s}$ in Seeley [30], well known in the case that $P^2 =$ Laplace operator. See also the article by Singer in these PROCEEDINGS.

Chazarain [5] determined the nature of the asymptotic expansion for the singularities at the periods $T \neq 0$ under the assumption of clean intersection of the graph of the H_p-flow with the diagonal. In [10] the top order terms of his expansions are computed in terms of the differential of the return map (Poincaré map) of the H_p-flow along the periodic solution curve. There it is also shown that periodicity of the total H_p-flow with period $T > 0$ is equivalent to a strong asymptotic clustering of the spectrum around the points $2\pi k/T + \beta$, $k = 1, 2, \cdots$. (β is a fixed real number.) It was this clustering effect which destroyed the possibility of improving the

error term in (2). If not all H_p-solutions are periodic (and some pathological cases are excluded), then the spectrum is in fact fairly evenly distributed and (2) can be replaced by ($n > 1$):

$$\#\{j;\lambda_j \leqq \lambda\} = (2\pi)^{-n}\cdot \mathrm{vol}(B^*X)\cdot\lambda^n - (2\pi)^{-n}\Big(\int_{S^*X} \mathrm{sub}\ P\Big)\lambda^{n-1} + o(\lambda^{n-1}), \qquad \lambda \to \infty. \tag{6}$$

Here $S^*X = \{(x,\xi) \in T^*X;\ p(x,\xi) = 1\}$, the integration is over the canonical density in T^*X divided by dp, and sub P is the so-called subprincipal symbol of P.

In the asymptotic expansions for (5) a power of i comes up due to the fact that the principal symbol of U is a section of a complex line bundle (called the Maslov bundle) over C, rather than a scalar-valued function. The results of Colin de Verdière [6] suggested that the integer in the exponent should be equal to the Morse index for periodic geodesics if $P = \Delta^{1/2}$. The proof of this resulted in a new approach to the Morse index in variational calculus; see [11].

If the condition of clean intersection is dropped, then the asymptotics of (5) will be that of more general oscillatory integrals, for which the phase functions have degenerate stationary points. Such integrals have been studied by Airy [1], Ludwig [23], and many others. Using Thom's theory of unfoldings of singularities much progress in the understanding of such integrals has been made recently; see Arnol'd [2], Guillemin and Schaeffer [15] and the survey [9]. Malgrange [24] has given an exposition on the relation with the monodromy of singularities.

In [7] parametrices and solutions having their singularities on a bicharacteristic strip have been constructed using Fourier integral operators, for pseudo-differential operators with real principal symbols having only simple zeros or complex principal symbols satisfying an integrability condition.

Recently Melin and Sjöstrand [25] (see also Kučerenko [20]) developed a theory of Fourier integral operators with complex phase functions. They also showed that the projection operators on the kernel and cokernel of operators P such that $\{p, \bar{p}\} \neq 0$ on $p = 0$, constructed in [8], belong to their class. A still more interesting application of their calculus is perhaps the construction of solutions of wide classes of equations with their wave front set on the intersection with the real cotangent bundle of a so-called positive invariant Lagrange manifold.

I hope that I have convinced you that Fourier integral operators now are a well-established tool in the theory of linear partial differential equations. I am sure that its use will continue to have a stimulating effect on the research in this area, at least in the near future.

References

1. G. B. Airy, *On the intensity of light in a neighborhood of a caustic*, Trans. Cambridge Philos. Soc. **6** (1838), 379–403.

2. V. I. Arnol'd, *Integrals of rapidly oscillating functions and singularities of the projections of Lagrangian manifolds*, Functional Anal. Appl. **6** (1972), no. 3, 61–62. See also: *Normal forms for*

functions near degenerate critical points, the Weyl groups A_k, D_k, E_k, and Lagrangian singularities, Functional Anal. Appl. **6** (1972), no. 4, 3–25. See also: Arnol'd's article in these PROCEEDINGS.

3. L. Boutet de Monvel, *Opérateurs hypo-elliptiques à caractéristiques doubles*, Séminaire Goulaouic-Schwartz 1973/74, Exposés I et II.

4. J. Chazarain, *Opérateurs hyperboliques à caractéristiques de multiplicité constante*, Ann. Inst. Fourier (Grenoble) **24** (1974), fasc. 1.

5. ———, *Formule de Poisson pour les variétés riemanniennes*, Invent. Math. **24** (1974), 65–82.

6. Y. Colin de Verdière, *Spectre du laplacien et longueurs des géodésiques périodiques*. II, Compositio Math. **27** (1973), 159–184.

7. J. J. Duistermaat and L. Hörmander, *Fourier integral operators*. II, Acta Math. **128** (1972), 184–269.

8. J. J. Duistermaat and J. Sjöstrand, *A global construction for pseudo-differential operators with non-involutive characteristics*, Invent. Math. **20** (1973), 209–225.

9. J. J. Duistermaat, *Oscillatory integrals, Lagrange immersions and unfoldings of singularities*, Comm. Pure Appl. Math. **27** (1974), 207–281.

10. J. J. Duistermaat and V. W. Guillemin, *The spectrum of positive elliptic operators and periodic geodesics*, Invent. Math. (to appear).

11. J. J. Duistermaat, *On the Morse index in variational calculus*, Advances in Math. (to appear).

12. Ju. V. Egorov, *The canonical transformations of pseudodifferential operators*, Uspehi Mat. Nauk **24** (1969), no. 5 (149), 235–236. (Russian) MR **42** #657.

13. ———, *On nondegenerate hypoelliptic pseudodifferential operators*, Dokl. Akad. Nauk SSSR **186** (1969), 1006–1007 = Soviet Math. Dokl. **10** (1969), 697–699. MR **40** #3049.

14. ———, *Nondegenerate subelliptic pseudodifferential operators*, Mat. Sb. **82 (124)** (1970), 323–342 = Math. USSR Sb. **11** (1970), 291–309. MR **42** #3406.

15. V. W. Guillemin and D. Schaeffer, *Remarks on a paper of D. Ludwig*, Bull. Amer. Math. Soc. **79** (1973), 382–385.

16. L. Hörmander, *The spectral function of an elliptic operator*, Acta Math. **121** (1968), 193–218.

17. ———, *Fourier integral operators*. I, Acta Math. **127** (1971), 79–183.

18. ———, *On the existence and regularity of solutions of linear pseudo-differential equations*, Enseignement Math. **17** (1971), 99–163.

19. C. Huygens, *Traité de la lumière* (1678), Reprint of the English translation of 1690 in Dover Publications, New York, 1962.

20. V. V. Kučerenko, *Maslov's canonical operator of the germ of a complex almost analytic manifold*, Dokl. Akad. Nauk SSSR **213** (1973), 1251–1254 = Soviet Math. Dokl. **14** (1973), 1879–1883.

21. P. D. Lax, *Asymptotic solutions of oscillatory initial value problems*, Duke Math. J. **24** (1957), 627–646. MR **20** #4096.

22. D. Ludwig, *Exact and asymptotic solutions of the Cauchy problem*, Comm. Pure Appl. Math. **13** (1960), 473–508. MR **22** #5816.

23. ———, *Uniform asymptotic expansions at a caustic*, Comm. Pure Appl. Math. **19** (1966), 215–250. MR **33** #4446.

24. B. Malgrange, *Intégrales asymptotiques et monodromie*, Grenoble, 1974 (preprint).

25. A. Melin and J. Sjöstrand, *Fourier integral operators with complex-valued phase functions*, 1974 (preprint).

26. L. Nirenberg and F. Trèves, *On local solvability of linear partial differential equations*. I. *Necessary conditions*. II. *Sufficient conditions*, Comm. Pure Appl. Math. **23** (1970), 1–38, 459–509. MR **41** #9064a, b.

27. M. Sato, T. Kawai and M. Kashiwara, *Microfunctions and pseudodifferential equations*, Lecture Notes in Math., vol. 287, Springer-Verlag, Berlin and New York, 1973, pp. 265–529.

28. J. Sjöstrand, *Parametrices for pseudo-differential operators with multiple characteristics*, Arkiv für Matematik **12** (1974), 85–130.

29. A. Weinstein, *Fourier integral operators, quantization and the spectra of Riemannian mani-*

folds, Proc. CNRS Colloq. Géométrie Symplectique et Physique Mathématique, Aix-en-Provence, June 1974 (to appear).

30. R. T. Seeley, *Complex powers of an elliptic operator*, Proc. Sympos. Pure Math., vol. 10, Amer. Math. Soc., Providence, R. I., 1967, pp. 288–307. MR **38** #6220.

MATHEMATISCH INSTITUUT DER RIJKSUNIVERSITEIT
UTRECHT, THE NETHERLANDS

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Elliptic Variational Inequalities

David Kinderlehrer*

1. The theory of elliptic variational inequalities has as its model the variational theory of boundary value problems for elliptic equations, but differs from this theory in that the competing functions belong to a convex set, rather than an affine space of functions. We propose to discuss here the influence of the constraints which define this convex set on the solution to the variational inequality. For simplicity of exposition, we focus our attention on several special problems of obstacle type. An interesting feature of such a problem is its set of coincidences, that set where the solution coincides with the obstacle. Also, in distinction to the solution of a boundary value problem, the solution to the variational inequality has a maximal smoothness, regardless of the smoothness of the data.

To initiate our discussion, we suppose given $\Omega \subset R^n$, a convex domain with smooth boundary $\partial\Omega$ and $\psi \in C^2(\bar{\Omega})$ an "obstacle," that is, $\max_\Omega \psi > 0$ *and* $\psi \leqq 0$ *on* $\partial\Omega$. Denote by

$$K = K_\psi = \{v : v \text{ is Lipschitz in } \bar{\Omega},\ v \geqq \psi \text{ in } \Omega, \text{ and } v = 0 \text{ on } \partial\Omega\}, \tag{1}$$

a convex set. We consider now

PROBLEM 1. *To find* $u \in K$: $\int_\Omega D_j u D_j(v - u)\, dx \geqq 0$ *for all* $v \in K$.

$D_j = \partial/\partial x_j$, $j = 1, \cdots, n$; and

PROBLEM 2. *To find* $u \in K$:

$$\int_\Omega \frac{D_j u}{(1 + |Du|^2)^{1/2}} D_j(v - u)\, dx \geqq 0 \quad \textit{for all } v \in K.$$

The existence of a solution to Problem 1 was established in G. Stampacchia [**36**] and J. L. Lions and G. Stampacchia [**29**]. In H. Lewy and G. Stampacchia [**27**] it was shown that $u \in H^{2,s}(\Omega) \cap C^{1,\lambda}(\bar{\Omega})$, $1 \leqq s < \infty$ and $0 < \lambda < 1$. The existence

*The preparation of this report was partially supported by ASOFR 71-2098.

of a solution $u \in H^{2,s}(\Omega) \cap C^{1,\lambda}(\bar{\Omega})$, $1 \leqq s < \infty$ and $0 < \lambda < 1$, to Problem 2 was proven by H. Lewy and G. Stampacchia [28] and also by M. Giaquinta and L. Pepe [14].

The interpretation of these problems is evident: The first refers to the function of minimum energy in K and the second to the function of minimum area in K. From the abstract point of view, each of the distributions $-\Delta u$ and Au defined via the pairings

$$(-\Delta u, \zeta) = \int_\Omega D_j u D_j \zeta \, dx, \qquad u, \zeta \in H^1(\Omega),$$

$$(Au, \zeta) = \int_\Omega \frac{D_j u}{(1 + |Du|^2)^{1/2}} D_j \zeta \, dx, \qquad u, \zeta \in H^{1,\infty}(\Omega),$$

is monotone and has some continuity property. Recall that an operator T which maps a Banach space V into its dual V^* is called monotone if $(Tv - Tv', v - v') \geqq 0$ for all $v, v' \in V$. The Laplace operator has the additional property of coerciveness meaning that $(-\Delta v, v) \geqq \alpha \|v\|^2$, $v \in H_0^1(\Omega)$, for an $\alpha > 0$, which the area functional does not possess. As a result the existence theory for Problem 2 involves finding an *a priori* estimate for $\sup |Du|$.

Let $I = \{x \in \Omega: u(x) = \phi(x)\}$ be the coincidence set of the solution and Γ its boundary, the "curve" of separation.

2. In this section we suppose that $n = 2$ and that ϕ is strictly concave, Ω strictly convex. Assuming that ϕ is also real analytic, H. Lewy and G. Stampacchia [27] proved that the Γ of Problem 1 has an analytic parametrization. Remarkably, it is not known beforehand in the proof of this theorem that Γ is even a curve. The proof relies on the holomorphic nature of $D_1 u - iD_2 u$ in $\Omega - I$ to extend, by means of an implicit function theorem, an appropriate conformal mapping.

This method did not generalize to Problem 2, where a different sequence of events transpired. There, it was first shown that Γ is a Jordan curve [16]. The analyticity of Γ was then achieved by the resolution of a system of differential equations which connects a conformal representation of the minimal surface $S = \{(x_1, x_2, x_3): x_3 = u(x_1, x_2), (x_1, x_2) \in \Omega - I\}$ with its harmonic, or analytic, extension. This idea is due to Hans Lewy, who used it to study minimal surfaces with prescribed or partly free boundaries ([23], [24]). Boundedness of the second derivatives $D_{jk}u$ is of vital importance to execute the program of finding, solving, and identifying the solution of the differential equations to which we have just referred. This is because of its use in the proof that Γ is a C^1 curve, a conclusion valid in both Problem 1 and Problem 2 assuming that $\phi \in C^3(\Omega)$ [18]. The limitation was established in [17] for this case. That Γ is analytic when ϕ is concave and analytic in Problem 2 answers a question of H. Lewy [25].

3. J. Frehse [8] gave an elegant proof that $D_{jk}u \in L^\infty(\Omega)$ for a variation of Problem 1. A general result was proved in [6] and a similar one by C. Gerhardt [13]. There is also other work by Frehse [9]. To explain this theorem, let Ω be as in § 1 and $a(p) = (a_1(p), \cdots, a_n(p))$, $p \in R^n$, be a C^2 vector field which satisfies the condition:

For each compact $C \subset R^n$, *there exists a* $\nu = \nu(C) > 0$ *such that*

$$(a(p) - a(q), p - q) \geqq \nu |p - q|^2 \quad \textit{for all } p, q \in C.$$

Now let $\psi \in C^2(\bar{\Omega})$ satisfy $\psi \leqq 0$ on $\partial\Omega$ and I be defined by (1). Let $f \in C^1(\bar{\Omega})$ and consider

PROBLEM 3. *To find* $u \in K$: $\int_\Omega a_j(Du)D_j(v - u)\,dx \geqq \int_\Omega f(v - u)\,dx$ *for all* $v \in K$.

In [6] it is shown that if u is a solution to Problem 3, then $D_{jk}u \in L^\infty_{\text{loc}}(\Omega)$. The proof depends on a deep result of G. Stampacchia [35]. Also in [6], using an idea of [4], it is proven that $Au = -D_j a_j(Du)$ is of bounded variation in Ω. Since the characteristic function of I, φ_I, may be written $\varphi_I = (Au - f)/(A\psi - f)$ a.e. in Ω, it follows that if $A\psi - f \neq 0$ in Ω and $\psi \in C^3(\Omega)$, then φ_I is also of bounded variation. Hence, in the language of DeGiorgi-Miranda (cf., e.g., [32]), I is a Caccioppoli set or a set of finite perimeter.

These results have led to generalizations of [18]; cf. [19], [20].

4. The existence of a solution to Problem 3 was shown in the case Ω convex and $f = 0$ in [28]. In general a relationship between f and $\partial\Omega$ must be fulfilled, for given $a(p)$, just as in the discussion of the Dirichlet problem for

$$-D_j a_j(Du) = f \text{ in } \Omega, \qquad u = g \text{ on } \partial\Omega \tag{2}$$

(cf. J. Serrin [34]). S. Mazzone has proved that if the vector field $a(p)$ is sufficiently coercive, then the solution to Problem 3 exists and is in $H^{2,s}(\Omega)$ [30]. She further shows that otherwise Problem 3 may lack solutions. The coerciveness conditions fail in particular when $a_j(p) = p_j/(1 + p^2)^{1/2}$, which is the obstacle problem for surfaces of prescribed mean curvature. For this case, S. Mazzone [31] and C. Gerhardt [12] have demonstrated the existence of a solution to Problem 3 under the same hypotheses necessary to obtain the solution to the Dirichlet problem (2).

Two noteworthy features of these papers are the construction of special global barriers for the solution (cf. especially [31]) and the use of a continuity method. In line with a discussion of variational inequalities having geometrical interpretations, G. Vergara-Caffarelli ([38], [39]) has considered an interesting problem about two surfaces of given mean curvature forced against each other.

5. At times, the curve of separation in Problem 3 arises as the graph of a function. The most celebrated example of this is perhaps due to C. Baiocchi ([1], cf. also these PROCEEDINGS [2]) who studied a problem of stationary fluid flow. In this connection there is also work of V. Benci [3] and a generalization to higher dimensions by G. Stampacchia [37]. A different generalization to higher dimensions has been given by A. Friedman [10].

An analogous situation is encountered in [22] where a free boundary problem in the plane is resolved. Given is a smooth f satisfying $\inf f < 0$ and $(\rho^2 f)_\rho \leqq 0$ in R^2. Here $z = x_1 + ix_2 = \rho e^{i\theta}$, $0 \leqq \theta < 2\pi$. Now consider

PROBLEM 4. *To find a bounded domain* Ω *and a function* u *such that*

$$\Delta u = \rho^{-1}(\rho^{-2} f)_\rho \quad \text{in } \Omega,$$

$$u = 0, \qquad \frac{\partial u}{\partial \nu} = \rho^2 f \frac{d\theta}{ds} \quad \text{on } \Gamma, \; u(0) = 1,$$

where $\Gamma = \partial\Omega$, ν is the outward directed normal vector, and s is the arclength on Γ.

The conditions prevailing on Γ are "natural". It turns out that u, Ω exist, that Ω is star-shaped with respect to $z = 0$, and Γ has a $C^{1,\lambda}$ parameterization (cf. also [21]). When f is analytic this result also follows from a work of H. Lewy [26]. The existence of u, Ω is secured by transforming Problem 4 into a form of Problem 3 with unbounded obstacle for a function $w(z)$ satisfying $u(z) = 1 - \rho w_\rho(z)$.

H. Brezis [5] has considered a variational inequality whose solution has compact support. This problem has a resemblance to the variational inequality for w. The possibility that the free boundary in a given situation was star-shaped motivated the study of a Stefan problem [11] after a formulation of G. Duvaut [7].

6. There are many important areas of variational inequalities omitted from this discussion. As two examples, we mention the work of M. K. V. Murthy and G. Stampacchia [33] and the paper of E. Giusti [15]. Both of these papers contain extensive biliographies.

References

1. C. Baiocchi, *Su un problema di frontiera libera connesso a questioni di idraulica*, Ann. Mat. Pura Appl. **92** (1972), 107–127.

2. ———, these PROCEEDINGS.

3. V. Benci, *On a filtration problem through a porous medium.*

4. H. R. Brezis, *Nouveaus théorèmes de regularité pour les problèmes unilateraux.* II, Rencoutre Phys. Math., Strasbourg, 1970.

5. ———, *Solutions with compact support of variational inequalities*, Uspehi Mat. Nauk **29** (1974), no. 2 (**176**), 103–108.

6. H. Brezis and D. Kinderlehrer, *The smoothness of solutions to nonlinear variational inequalities*, Indiana J. Math. **23** (1974), 831–844.

7. G. Duvaut, *Resolution d'un problème de Stefan (Fusion d'un bloc de glace à zero degre)*, C. R. Acad. Sci. Paris **276** (1973), 1461–1463.

8. J. Frehse, *On the regularity of the solution of a second order variational inequality*, Bol. Un. Mat. Ital. (4) **6** (1972), 312–315. MR **47** #7197.

9. ———, *On the regularity of solutions of linear elliptic variational inequalities.*

10. A. Friedman, *The shape and smoothness of the free boundary of the solutions for some elliptic variational inequalities.*

11. A. Friedman and D. Kinderlehrer, *A one phase Stefan problem*, Indiana J. Math.

12. C. Gerhardt, *Hypersurfaces of prescribed mean curvature over obstacles*, Math. Z. **133** (1973), 169–185.

13. ———, *Regularity of solutions of nonlinear variational inequalities*, Arch. Rational Mech. Anal. **52** (1973), 389–393.

14. M. Giaquinta and L. Pepe, *Esistenza e regolarità per il problema dell'area minima con ostacoli in n variabili*, Ann. Scuola Norm. Sup. Pisa (3) **25** (1971), 481–507. MR **46** #4335.

15. E. Giusti, *Minimal surfaces with obstacles*, CIME course on Geometric Measure Theory and Minimal Surfaces, Edizioni Cremonese, Rome, 1973, pp. 119–153.

16. D. Kinderlehrer, *The coincidence set of solutions of certain variational inequalities*, Arch. Rational Mech. Anal. **40** (1970/71), 231–250. MR **42** #6680.

17. ———, *The regularity of the solution to a certain variational inequality*, Proc. Sympos. Pure Math., vol. 23, Amer. Math. Soc., Providence, R. I., 1971, pp. 353–363.

18. ———, *How a minimal surface leaves an obstacle*, Acta Math. **130** (1970), 221–242.

19. ———, *Some questions related to the coincidence set in variational inequalities*, Sympos. Math.

20. ———, *Remarks about the free boundaries occurring in variational inequalities*, CIME course on New Variational Techniques in Mathematical Physics, Edizioni Cremonese, Rome, 1973, pp. 104–119.

21. ———, *The free boundary determined by the solution to a differential equation*, Indiana J. Math.

22. D. Kinderlehrer and G. Stampacchia, *A free boundary problem in the plane* (to appear).

23. H. Lewy, *On the boundary behavior of minimal surfaces*, Proc. Nat. Acad. Sci. U.S.A. **37** (1951), 103–110. MR **14**, 168.

24. ———, *On minimal surfaces with partially free boundary*, Comm. Pure Appl. Math. **4** (1952), 1–13. MR **14**, 662.

25. ———, *On a minimum problem for superharmonic functions*, Proc. Internat. Conf. on Functional Analysis, Tokyo, 1969.

26. ———, *On the nature of the boundary separating two domains with different regimes* (to appear).

27. H. Lewy and G. Stampacchia, *On the regularity of the solution of a variational inequality*, Comm. Pure Appl. Math. **22** (1969), 153–188. MR **40** #816.

28. ———, *On existence and smoothness of solutions of some noncoercive variational inequalities*, Arch. Rational Mech. Anal. **41** (1971), 241–253.

29. J.-L. Lions and G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math. **20** (1967), 493–519. MR **35** #7178.

30. S. Mazzone, *Existence and regularity of the solution of certain nonlinear variational inequalities with an obstacle*, Arch. Rational Mech. Anal.

31. ———, *Un problema di disequazioni variazionali per superficie di curvatura media assegnata*, Boll. Un. Mat. Ital. (4) **7** (1973), 318–329. MR **47** #4109.

32. M. Miranda, *Distribuzioni aventi derivate misure insiemi di perimetro localmente finito*, Ann. Scuola Norm. Sup. Pisa (3) **18** (1964), 27–54. MR **29** #2364.

33. M. K. V. Murthy and G. Stampacchia, *A variational inequality with mixed boundary conditions*, Israel J. Math. **13** (1972), 188–224.

34. J. Serrin, *The problem of Dirichlet for quasilinear elliptic differential equations with many independent variables*, Philos. Trans. Roy. Soc. London Ser. A **264** (1969), 413–496. MR **43** #7772.

35. G. Stampacchia, *Le problème de Dirichlet pour les équations elliptique du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble) **15** (1965), fasc. 1, 189–258. MR **33** #404.

36. ———, *Formes bilinéares coercitives sur les ensembles convexes*, C. R. Acad. Sci. Paris **258** (1964), 4413–4416. MR **29** #3864.

37. ———, *On the filtration of a fluid through a porous medium with variable cross-section* (to appear).

38. G. Vergara-Caffarelli, *Variational inequalities for two surfaces of constant mean curvature*, Arch. Rational Mech. Anal.

39. ———, *Surfaces de coubure moyenne donnée dans L^p. Application à un problème d'inéquations variationnelles*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A1043-A1045. MR **47** #7548.

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Monge-Ampère Equations and Some Associated Problems in Geometry

L. Nirenberg*

1. Given some geometric structure, it is usually extremely useful to find analytic structures, such as differential equations, which are associated in some natural way with the geometry. As we well know, geometry has been an extremely rich source of interesting problems in partial differential equations. The equations may be complicated and some have resisted efforts of analysts for many years.

This article is devoted to recent joint work with E. Calabi [4] on some analytic problems coming from geometric questions which involve nonlinear elliptic equations of Monge-Ampère type (to be explained).

Monge-Ampère equations occur in various geometric problems, for example, isometric embeddings; perhaps the most familiar one is the classical Minkowski problem. Let $\psi(\nu)$ be a given positive smooth function defined on $S^n \subset R^{n+1}$ satisfying

$$\int_{S^n} \nu\psi(\nu)\, d\omega = 0$$

where $d\omega$ is the usual measure on S^n (thus ψ satisfies $n+1$ conditions). Find a compact convex hypersurface $M \subset R^{n+1}$ with the property that for each $\nu \in S^n$, at the point where the exterior normal to M is ν, the curvature $K =$ product of principal curvatures is equal to $\psi(\nu)$.

Minkowski proved the existence of a "generalized solution" of the problem. After a number of years and the efforts of a number of mathematicians, it was shown, in case $n = 2$, that if ψ is smooth, so is the generalized solution. For $n > 2$ several authors have contributed to the problem of smoothness of the solution,

*This work was supported by an ONR grant N00014-67A-0467-0024.

but it was finally proved only in 1971 by A. V. Pogorelov [6]. Analytically, the problem may be expressed as a nonlinear partial differential equation for the support function, which turns out to be of Monge-Ampère type. Pogorelov has seen how to obtain the crucial a priori estimates for the second and third derivatives of the solutions. For the latter he makes use of some ideas used by Calabi [1].

In his papers [7], [8] in 1971 Pogorelov has considered the Dirichlet problem for a convex function $u(x)$ in a bounded convex domain Ω in R^n:

$$\begin{aligned} \det(u_{x^i x^j}) &= \phi(x) > 0 \quad \text{in } \Omega, \\ u &= g \qquad\quad \text{on } \partial\Omega, \end{aligned} \tag{1}$$

and has shown how to obtain smooth strictly convex solutions in the interior, under some smoothness conditions on ϕ and g. Equation (1) is an equation of Monge-Ampère type, and this usually means a partial differential equation of second order for a function u in which the principal terms of highest order are nonlinear and of the form $\det(u_{x^i x^j})$.

I am going to describe two geometric-analytic problems that Calabi and I have solved, and then discuss the related, purely technical, results on the Dirichlet problem for Monge-Ampère equations. In proving these we have used and extended the techniques of Pogorelov in [6] — [8].

2. First problem. Some years ago Charles Loewner formulated the problem of assigning to each bounded convex domain Ω in R^n a Riemannian metric which is invariant under projective transformations between such domains. The classical Hilbert metric defines a distance function between any two points $x, y \in \Omega$ which is invariant under projective transformations between convex domains. In general this metric does not correspond to a Riemannian metric—though it does in the case of a ball: For $|x| < 1$ in R^n,

$$ds^2 = (1 - |x|^2)^{-1} \sum_i dx^{i^2} + (1 - |x|^2)^{-2} \Big(\sum_i x^i dx^i \Big)^2. \tag{2}$$

Loewner proposed an invariant *Riemannian metric* which agrees with Hilbert's for the ball. It is defined by means of the solution of a boundary value problem for a negative strictly convex function u defined in Ω; u is to satisfy the Monge-Ampère equation

$$\det(u_{x^i x^j}) = \frac{1}{|u|^{n+2}} \quad \text{in } \Omega, \qquad u = 0 \text{ on } \partial\Omega. \tag{3}$$

Since $u = 0$ on the boundary, this is highly singular as well as nonlinear. The associated metric is

$$ds^2 = -\frac{1}{u} \sum u_{x^i x^j}\, dx^i dx^j. \tag{4}$$

The solution if it exists is unique, and for the ball $|x| < 1$ it is $u = -(1 - |x|^2)^{1/2}$ leading to the Hilbert metric (2) in that case.

Under a projective transformation of Ω to some other convex domain, the problem (3) and the metric (4) are invariant in a suitable sense—the function u

does not transform as a scalar but is multiplied by a suitable function depending on the projective transformation (see [5]).

Some remarks about the complicated equation (3) and the metric (4) are in order. It turns out that they are quite classical. For $n = 2$ the equation, and its invariance under projective transformations, were discussed in 1908 by G. Tzitzèica in his study of affine hyperspheres (these correspond to solutions of the equation after a Legendre transformation). Blaschke and others, most recently Calabi, have made further studies of affine hyperspheres; in particular the metric (4) was proposed by Berwald and Blaschke. See [2] for details and references.

The boundary value problem for $n = 2$, and Ω having smooth strictly convex boundary, was solved in Loewner, Nirenberg [5], where it was proved that there is a unique strictly convex solution $u \in C^\infty(\Omega) \cap C(\bar{\Omega})$, and the metric (4) is complete, i.e., C^1 curves tending to the boundary have infinite length when measured in the metric (4).

THEOREM A [4]. *For Ω bounded and convex in R^n, $n \geqq 2$ (with $\partial\Omega$ not necessarily smooth), there exists a unique, strictly convex, negative solution u belonging to $C^\infty(\Omega) \cap C(\bar{\Omega})$ of* (3), *and the corresponding metric* (4) *is complete.*

It follows from Theorem 5.4 of [2] that the Ricci curvature tensor (as a quadratic form) of the metric (4) is nonpositive.

REMARK. To prove completeness of the metric we establish the following estimate for the solution:

$$\frac{1}{|u|} \sum u^{ij} u_{x^i x^j} \leqq \text{constant in } \Omega, \tag{5}$$

where u^{ij} is the inverse matrix of the Hessian matrix $u_{x^i x^j}$.

Second problem. In [3] Calabi proposed a Kähler-Einstein metric associated with certain tubular domains in $\boldsymbol{C}^n$: $M = \Omega \oplus iR^n$ where Ω is a convex domain in R_n. The metric

$$ds^2 = \sum g_{i\bar{j}}\, dz^i d\bar{z}^j$$

is to be Kähler, i.e., $d(g_{i\bar{j}}\, dz^i \wedge d\bar{z}^j) = 0$, invariant under the group of transformations along iR^n, and is to be Einstein, i.e., the complex Ricci tensor

$$R_{i\bar{j}} = \frac{\partial^2}{\partial z^i \partial \bar{z}^j} \log g, \qquad g = \det(g_{i\bar{j}}),$$

is to satisfy

$$R_{i\bar{j}} = K g_{i\bar{j}}, \qquad K = \text{constant}. \tag{6}$$

For a negative constant (it suffices to consider $K = -1$) the metric is defined via a solution of the following boundary value problem: Find a real strictly convex function u in Ω satisfying

$$\det(u_{x^i x^j}) = e^u \quad \text{in } \Omega, \qquad u(x) \to +\infty \text{ as } x \to \partial\Omega. \tag{7}$$

The metric is then defined as (here $z^j = x^j + iy^j$)

$$ds^2 = \sum u_{x^i\bar{x}^j}(x)\, dz^i\, d\bar{z}^j. \tag{8}$$

With ds so defined, the equation (7) simply expresses condition (6). The condition $u = +\infty$ on $\partial\Omega$ is to make the metric as large as possible in order that it be complete.

THEOREM B [4]. *Let Ω be a bounded convex domain in R^n, $n \geqq 2$. There exists a unique, strictly convex, C^∞ solution u of* (7). *Furthermore, in $\Omega \oplus iR^n$ the Einstein metric* (8) *is complete.*

REMARK. Completeness is proved with the aid of the estimate

$$\sum u^{ij} u_{x^i} u_{\bar{x}^j} \leqq \text{constant in } \Omega,$$

where u^{ij} is the inverse matrix of the Hessian matrix $u_{x^i\bar{x}^j}$.

3. We now turn to the technical results concerning the Dirichlet problem for Monge-Ampère equations of the form

$$\begin{aligned} \det(u_{x^i\bar{x}^j}) &= \psi(x, u) > 0 \quad \text{in } \Omega, \\ u &= g \qquad\qquad\quad \text{on } \partial\Omega. \end{aligned} \tag{9}$$

Here ψ is a positive C^∞ function, g is continuous (usually smooth) and Ω is a bounded convex domain in R^n, $n \geqq 2$. We seek smooth and strictly convex solutions. First a result giving smooth solutions in $\bar{\Omega}$.

THEOREM 1 [4]. *Assume that $\psi \in C^\infty(\bar{\Omega} \times R^1)$ and $\partial\psi/\partial u \geqq 0$. Assume that $\partial\Omega$ is C^∞ and strictly convex, and that g belongs to $C^\infty(\partial\Omega)$. Then there exists a unique strictly convex solution u of* (9) *belonging to $C^\infty(\bar{\Omega})$ if and only if there exists a strictly convex function $u^0 \in C^\infty(\bar{\Omega})$ which equals g on $\partial\Omega$ and satisfies*

$$\det(u^0_{x^i\bar{x}^j}) \leqq \psi(x, u^0). \tag{10}$$

(It suffices that $u^0 \in C^3(\Omega)$.)

The proof is based on a priori estimates and makes use of the continuity method: For each t in $0 \leqq t \leqq 1$ we find a smooth solution of

$$\det(u_{x^i\bar{x}^j}) = t\psi(x, u) + (1 - t)\det(u^0_{x^i\bar{x}^j}) \text{ in } \Omega, \qquad u = g \text{ on } \partial\Omega.$$

The a priori estimates for the second and third derivatives[1] are extensions of those of Pogorelov for (1). He derived strictly interior estimates and regularity, assuming $g \in C^k$, $k \geqq 2$. Having interior estimates one might be tempted to believe that it should be possible to obtain solutions which are in $C^\infty(\Omega) \cap C(\bar{\Omega})$ in case g is merely continuous. However, this is not always the case. In [8] Pogorelov presents an interesting example with $g \in C^{1+\alpha}$, $0 < \alpha < 1$, $g \notin C^2$, for which the unique "generalizes solution" of (1) is not smooth in Ω. Furthermore he proves in [7] that if the generalized solution is strictly convex then it *is* smooth in Ω. It is possible to give reasonable conditions on continuous boundary values g to guarantee the existence of smooth solutions inside Ω:

THEOREM 1′ [4]. *Let Ω be a bounded convex domain in R^n. Assume $\psi \in C^\infty(\bar{\Omega} \times R^1)$*

[1] ADDED IN PROOF. We recently observed that our a priori estimates of the third derivatives at the boundary are incomplete. The results presented in this paper are, therefore at this point, not fully established.

(this may be weakened) and $\phi_u \geqq 0$. *Suppose* $g \in C(\partial\Omega)$ *and there exist two strictly convex functions* $u^0, u^1 \in C^\infty(\Omega) \cap C(\partial\Omega)$ *which equal* g *on* $\partial\Omega$ *and satisfy*

$$\det(u^0_{x^i x^j}) \leqq \phi(x, u^0), \qquad \det(u^1_{x^i x^j}) \geqq \phi(x, u^1) \tag{11}$$

in Ω. *Then there exists a unique, strictly convex, solution* $u \in C^\infty(\Omega) \cap C(\bar{\Omega})$ *of* (9).

Clearly the existence of such functions u^0, u^1 is also necessary.

Among other things the proof uses the maximum principle—guaranteed by the condition $\phi_u \geqq 0$. If we drop this condition we may still obtain solutions by a familiar (though in this case more tedious) argument.

THEOREM 2 [4]. *Let* Ω *be a bounded convex domain in* R^n *with* $\partial\Omega$ *strictly convex and* C^∞. *Assume* $\phi \in C^\infty(\bar{\Omega} \times R^1)$ *and* $g \in C^\infty(\partial\Omega)$. *Assume that there are two strictly convex functions* u^0, u^1 *in* $C^\infty(\bar{\Omega})$ *which equal* g *on* $\partial\Omega$ *and satisfy* (11) *and* $u^1 \leqq u^0$ *in* Ω. *Then there exists a strictly convex solution* $u \in C^\infty(\bar{\Omega})$ *of* (9) *satisfying* $u^1 \leqq u \leqq u^0$ *in* Ω. *In fact there are minimal and maximal such solutions.*

Again, the existence of such functions u^0, u^1 is also necessary.

To prove Theorems A and B, both of which involve singular boundary value problems, we first approximate them by regular ones: For positive constant ε we solve (3) and (7) with the respective modified boundary conditions $u = \varepsilon$, $u = 1/\varepsilon$ on $\partial\Omega$, and then let $\varepsilon \to 0$. In carrying out the limit process we use a priori estimates for the solutions and their derivatives in the interior of Ω which are independent of ε.

Bibliography

1. E. Calabi, *Improper affine hyperspheres of convex type and a generalization of a theorem by K. Jörgens*, Michigan Math. J. **5** (1958), 105–126. MR **21** #5219.

2. ———, *Complete affine hyperspheres.* I, Ist. Naz. Alta Mat., Symposia Mathematica 10, 1972, pp. 19–38.

3. ———, *A construction of nonhomogeneous Einstein metrics*, Proc. Sympos. Pure Math., vol. 27, Part 2, Amer. Math. Soc., Providence, R. I., 1975, pp. 17–24.

4. E. Calabi and L. Nirenberg (to appear).

5. C. Loewner and L. Nirenberg, *Partial differential equations invariant under conformal or projective transformations*, Contributions to Analysis, Academic Press, New York, 1974, pp. 245–272.

6. A. V. Pogorelov, *A regular solution of the n-dimensional Minkowski problem*, Dokl. Akad. Nauk SSSR **199** (1971), 785–788 = Soviet Math. Dokl. **12** (1971), 1192–1196. MR **44** #2180.

7. ———, *On the regularity of generalized solutions of the equation* $\det(\partial^2 u/\partial x^i \partial x^j) = \phi(x^1, x^2, \cdots, x^n) > 0$, Dokl. Akad. Nauk SSSR **200** (1971), 534–537 = Soviet Math. Dokl. **12** (1971), 1436–1440. MR **45** #2304.

8. ———, *The Dirichlet problem for the multidimensional analogue of the Monge-Ampère equation*, Dokl. Akad. Nauk SSSR **201** (1971), 790–793 = Soviet Math. Dokl. **12** (1971), 1727–1731. MR **45** #2305.

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY
NEW YORK, NEW YORK 10012, U.S.A.

Труды Международного Конгресса Математиков
Ванкувер, 1974

Аналитические Решения Уравнений с Вариационными Производными и их Приложения*

М. И. Вишик

В настоящем докладе излагаются результаты работ А. В. Фурсикова и М. И. Вишика [1]—[4], относящихся к аналитическим первым интегралам нелинейных параболических систем дифференциальных уравнений, а также уравнений Навье—Стокса и их приложениям.

Для краткости изложение ведется в случае периодических граничных условий. Случай всего пространства рассмотрен в [2]. Доклад разбит на пять пунктов: (1) Аналитические первые интегралы параболических систем; (2) Функционально-аналитическая зависимость решений параболических систем от начальных условий; (3) Уравнения Навье—Стокса; (4) Построение моментных функций и характеристического функционала статистических решений; (5) Асимптотическое разложение моментных функций статистических решений.

1. Аналитические первые интегралы параболических систем. Пусть $A(\xi) = \|A_{ij}(\xi)\|_{i,j=1,\cdots,p}$—матрица размера $p \times p$, элементы которой $A_{ij}(\xi)$ являются многочленами порядка $\leqq m$ от $(\xi^1, \cdots, \xi^n) = \xi$:

$$A_{ij}(\xi) = \sum_{|\alpha| \leqq m} a_\alpha^{ij} \xi^\alpha, \qquad \alpha = (\alpha_1, \cdots, \alpha_n), \ |\alpha| = \alpha_1 + \cdots + \alpha_n.$$

Предполагается, что система дифференциальных операторов

$$\frac{\partial u}{\partial t} + A(D)u, \qquad D = (D_1, \cdots, D_n), \qquad D_k = \frac{1}{i}\frac{\partial}{\partial x_k}, \tag{1.1}$$

$$u = (u^1(t, x), \cdots, u^p(t, x)), \qquad x = (x^1, \cdots, x^n) \in \boldsymbol{R}^n,$$

*Not presented in person.

—параболическая по И.Г. Петровскому, т.е. при любом $\xi \in \boldsymbol{R}^n$ все собственные значения $\lambda_j^0(\xi)$ $(j = 1, \cdots, p)$ матрицы $A^0(\xi) = \left\| \sum_{|\alpha|=m} a_\alpha^{ij} \xi^\alpha \right\|$ имеют положительную действительную часть, т.е. $\operatorname{Re} \lambda_j^0(\xi) > 0$, следовательно, $\operatorname{Re} \lambda_j^0(\xi) \geqq c_0^2 |\xi|^m$. Более того, будем предполагать, что собственные значения $\lambda_j(\xi)$ матрицы $A(\xi)$ удовлетворяют условию $\operatorname{Re} \lambda_j(\xi) > 0$, $\xi \in \boldsymbol{R}^n$. Отметим, что в этом случае справедлива оценка

$$\left\| e^{-A(\xi)t} \right\| \leqq Ce^{-\alpha(1+|\xi|)^m t} \qquad [1]. \tag{1.2}$$

Пусть

$$c_k(D) = (c_k^1(D)(u), \cdots, c_k^p(D)(u)) \tag{1.3}$$

—векторнозначный k-линейный дифференциальный оператор порядка $\sigma, \sigma \leqq m$, причем

$$c_k^l(D)(u) = \sum_{j_1, \cdots, j_k=1}^{p} \sum_{|\alpha_{j_1}|, \cdots, |\alpha_{j_k}| \leqq \sigma} c_k^l(j_1, \cdots, j_k; \alpha_{j_1}, \cdots, \alpha_{j_k}) D^{\alpha_{j_1}} u^{j_1} \cdots D^{\alpha_{j_k}} u^{j_k},$$

где $c_k^l(\cdots, \cdots)$—постоянные коэффициенты, α_{j_i}—мультииндексы. Через $f(D^\gamma u)$, $|\gamma| \leqq \sigma$, обозначается следующая, вообще говоря, бесконечная сумма k-линейных операторов:

$$f(D^\gamma u) = \sum_{k=2}^{\infty} c_k(D)(u). \tag{1.4}$$

Сопоставим функции $D^{\alpha_l} u^j$ числа $z_{\alpha_l} \in \boldsymbol{C}$. Тогда оператору (1.4) сопоставляется функция $f(z) = (f^1(z), \cdots, f^p(z))$ со значениями в $\boldsymbol{C}^p$, где $z = (\cdots, z_{\alpha_l}, \cdots)$ —точка многомерного комплексного пространства. Предполагается, что выполнено

Условие 1.1. Функции $f^l(z)$ $(l = 1, \cdots, p)$ голоморфны в некотором полидиске $|z_{\alpha_l}^j| < R$, $R > 0$.

Изучается нелинейная система дифференциальных уравнений вида

$$\partial u(t, x)/\partial t = -A(D)u + f(D^\gamma u) \tag{1.5}$$

при периодических граничных условиях

$$u(t, x^1, \cdots, x^k + 2\pi, \cdots, x^n) = u(t, x^1, \cdots, x^k, \cdots, x^n), \tag{1.6}$$

$t > 0$, или во всем пространстве $\boldsymbol{R}^n : x \in \boldsymbol{R}^n$, $t > 0$. Через

$$v(\eta) = (2\pi)^{-n} \int_0^{2\pi} \cdots \int_0^{2\pi} e^{-i\eta \cdot x} u(x)\, dx = F_{x \to \eta} u(x), \qquad \eta \in \boldsymbol{Z}^n, \tag{1.7}$$

обозначаются коэффициенты Фурье функции $u(x)$.

Легко видеть, что

$$F_{x \to \xi}(c_k(D)(u)) = \sum_{\eta_1, \cdots, \eta_k} \delta(\xi - \eta_1 - \cdots - \eta_k)\, c_k(\eta_1, \cdots, \eta_k)\, (v(\eta_1), \cdots, v(\eta_k)), \tag{1.8}$$

где $c_k(\eta_1, \cdots, \eta_k)$—$k$-линейный оператор, соответствующий оператору (1.3), $\delta(\zeta) = 0$ при $\zeta \neq 0$, $\delta(0) = 1$, $\zeta \in \boldsymbol{Z}^n$ (см. [1]).

Для краткости в дальнейшем используется обозначение

$$L_k(\eta_1, \cdots, \eta_k)(v(\eta_1), \cdots, v(\eta_k)) = L_k(\eta)v(\eta), \tag{1.9}$$

причем индекс k у L_k указывает на то, что $L_k(\eta) = L_k(\eta_1, \cdots, \eta_k)$ и что этот оператор действует на набор из k векторов $v(\eta) = (v(\eta_1), \cdots, v(\eta_k))$. Из (1.4), (1.8), (1.9) следует, что

$$F_{x\to\xi}(f(D^r u)) = H(\xi, v) = \sum_{k=2}^{\infty} \delta(\xi - \eta_1 - \cdots - \eta_k)c_k(\eta)v(\eta). \tag{1.10}$$

Переходя к коэффициентам Фурье в обеих частях (1.5), получим следующую систему уравнений:

$$\partial v(t, \xi)/\partial t = -A(\xi)v(t, \xi) + H(\xi, v(t, \cdot)), \qquad \xi \in \mathbf{Z}^n, \tag{1.11}$$

в которой $H(\xi, v)$ задается формулой (1.10).

Пусть V_s—банахово пространство, состоящее из всех вектор-функций $v(\xi) = (v^1(\xi), \cdots, v^p(\xi))$, $\xi \in \mathbf{Z}^n$, для которых конечна норма

$$\|v\|_s = \sum_{\xi\in Z^n} (1 + |\xi|)^s |v(\xi)|, \quad \text{где } |v(\xi)| = \Big(\sum_{j=1}^{p} |v^j(\xi)|^2\Big)^{1/2}. \tag{1.12}$$

Через B_a^s обозначается шар радиуса a в пространстве V_s:

$$B_a^s = \{v : \|v\|_s < a\}.$$

Определение 1.1. Функционал $\Psi(v(\cdot))$, заданный в шаре B_a^s, $a > 0$, называется аналитическим и принадлежащим классу $\mathfrak{A}(B_a^s)$, если внутри этого шара он разлагается в сходящийся ряд вида

$$\Psi(v) = \sum_{r=0}^{\infty} \Psi_r(v), \tag{1.13}$$

где $\Psi_0(v) = \Psi_0 = \text{const}$ и при $r \geqq 1$,

$$\Psi_r(v) = \sum_{\eta_1, \cdots \eta_r} \Psi_r(\eta_1, \cdots, \eta_r)(v(\eta_1), \cdots, v(\eta_r)) = \sum_{\eta} \Psi_r(\eta)v(\eta),$$

где $\Psi_r(\eta) = \Psi_r(\eta_1, \cdots, \eta_r)$—$r$-линейная форма относительно $v(\eta_1), \cdots, v(\eta_r)$, зависящая от $\eta_1, \cdots, \eta_r$ как от параметров. При этом предполагается, что все функционалы $\Psi_r(\eta)$ для любого набора $\eta_1, \cdots, \eta_r$ удовлетворяют следующему условию симметрии:

$$\begin{aligned} \Psi_r(\eta_{j_1}, \cdots, \eta_{j_r})(v(\eta_{j_1}), \cdots, v(\eta_{j_r})) &= \Psi(\eta_1, \cdots, \eta_r)\,(v(\eta_1), \cdots, v(\eta_r)) \\ &= \Psi_r(\eta)v(\eta) = \sum_{k_1, \cdots, k_r=1}^{p} \Psi_r(\eta; k_1, \cdots, k_r)v^{k_1}(\eta_1), \cdots, v^{k_r}(\eta_r), \end{aligned} \tag{1.14}$$

где $(j_1, \cdots, j_r)$—любая перестановка $(1, \cdots, r)$. Введем следующие обозначения:

$$|\Psi_r(\eta)| = \Big(\sum_{k_1, \cdots, k_r=1}^{p} |\Psi_r(\eta; k_1, \cdots, k_r)|^2\Big)^{1/2},$$

$$[\Psi_r]_s = \sup_{\eta_1, \cdots, \eta_r} |\Psi_r(\eta)| \Big/ \prod_{j=1}^{r} (1 + |\eta_j|)^s, \qquad \eta_j \in \mathbf{Z}^n.$$

Если функционал $\Psi(v)$ аналитичен и ограничен в шаре $B_{\gamma_1}^s$ и $\gamma_1 > \gamma_2$, то найдется такая константа D, что

$$[\Psi_r]_s \leqq D(\gamma_1 \cdot e \cdot \sqrt{p})^r = C\gamma_0^r, \quad \forall r. \tag{1.15}$$

Обратно, из оценок (1.15) следует, очевидно, что функционал $\Psi(v)$ аналитичен в шаре $B^s_{\gamma_0^{-1}}$ (см. [1]).

Определение 1.2. Функционал $\Phi(t, v)$, определенный в некоторой области $\Omega \subset \boldsymbol{R}^1_t \times V_s$, называется первым интегралом системы (1.11), если на любом решении $v(t, \xi)$ системы (1.11), таком, что $(t, v(t, \cdot)) \in \Omega$, он принимает постоянное значение, т.е. $\Phi(t, v(t, \cdot)) = \text{const}$. Аналогично определяется первый интеграл системы (1.5).

Пусть $\Phi(t, v)$—первый интеграл системы (1.11), дифференцируемый по Фреше в некоторой области Ω пространства $\boldsymbol{R}^1_t \times V_s$. Тогда, как показано в [1], этот функционал удовлетворяет следующему дифференциальному уравнению в частных производных первого порядка от бесконечного числа переменных:

$$\frac{\partial \Phi(t, v)}{\partial t} - \sum_{\xi} \left\langle \frac{\partial \Phi(t, v)}{\partial v(\xi)}, Av(\xi) \right\rangle + \sum_{\xi} \left\langle \frac{\partial \Phi(t, v)}{\partial v(\xi)}, H(\xi, v) \right\rangle = 0, \tag{1.16}$$

где $\partial\Phi/\partial v(\xi) = (\partial\Phi/\partial v^1(\xi), \cdots, \partial\Phi/\partial v^p(\xi))$. Это уравнение для первых интегралов.

В том случае, когда решения уравнения (1.5) рассматриваются во всем пространстве: $x \in \boldsymbol{R}^n$, параметр ξ в (1.11) пробегает все $\boldsymbol{R}^n_\xi$, а в уравнении (1.16) знак суммы по $\xi \in \boldsymbol{R}^n$ заменяется интегралом по $\xi \in \boldsymbol{R}^n$, частная производная $\partial\Phi/\partial v(\xi)$ заменяется вариационной производной $\delta\Phi/\delta v(\xi)$ (см. [2]). Для краткости ниже мы ограничиваемся случаем периодических граничных условий.

Для уравнений первых интегралов (1.16) имеет место следующий аналог теоремы Коши—Ковалевской:

Теорема 1.1. *Пусть параболическая матрица $A(\xi)$ удовлетворяет условию* (1.2) *с константой $C = 1$, а оператор $f(D^r u)$ удовлетворяет условию* 1.1. *Пусть при $t = 0$ задано начальное условие*

$$\Phi\big|_{t=0} = \Psi(v), \tag{1.17}$$

где функционал $\Psi(v)$ принадлежит классу $\mathfrak{A}(B^s_{\rho_0})$, причем имеют место оценки (1.15). *Тогда для $s \geqq \sigma$ (где σ—максимальный порядок производных, входящих в $f(D^r u)$) существует функционал $\Phi(t, v)$, удовлетворяющий следующим условиям*:

(а) *Для любого $t \leqq 0$ существует такое $\rho > 0$, не зависящее от t, что $\Phi(t, v) \in \mathfrak{A}(B^s_\rho)$, следовательно,*

$$\begin{aligned} \Phi(t, v) &= \sum_{r=0}^{\infty} \Phi_r(t, v), \qquad \Phi_0(t, v) = \Phi_0(t), \\ \Phi_r(t, v) &= \sum_{r} \Phi_r(t, \eta)\, v(\eta) = \sum_{\eta_1, \cdots, \eta_r} \Phi(t, \eta_1, \cdots, \eta_r)(v(\eta_1), \cdots, v(\eta_r)), \end{aligned} \tag{1.18}$$

причем r-линейные формы $\Phi_r(t, \eta)$ удовлетворяют условию симметрии (1.14).

(б) *Существуют такие константы $C > 0, \gamma > 0$, не зависящие от t, что*

$$[\Phi_r(t, \cdot)]_s \leqq C\gamma^r, \qquad \rho < \gamma^{-1}, \ -\infty < t < 0. \tag{1.18'}$$

(в) *Функционал* $\Phi(t, v)$ *удовлетворяет начальному условию* (1.17).

(г) *При* $t < 0$ *функционал* $\Phi(t, v)$ *на шаре* $B_\rho^{s+m} \subset V_{s+m}$ *удовлетворяет уравнению* (1.16).

Решение $\Phi(t, v)$ *задачи Коши* (1.16), (1.17), *удовлетворяющее условиям* (а)—(г), *единственно.*

При доказательстве этой теоремы (см. [1]) выводится система рекуррентных дифференциальных уравнений для коэффициентов $\Phi_r(t, \eta_1, \cdots, \eta_r)$, из которых далее выводятся оценки (1.18′). При этом существенным оказывается выбор пространства V_s с нормой (1.12).

Теорема 1.1 остается справедливой и для систем вида (1.11), двойственных к системам (1.1), коэффициенты которых зависят от t и от x [2].

Аналогичная теорема доказана в [2] для уравнений с вариационными производными, отвечающих случаю всего пространства $\boldsymbol{R}^n$.

2. Функционально-аналитическая зависимость решений параболических систем от начальных условий. (а) Рассматривается система уравнений (1.11) при начальных условиях

$$v\big|_{t=0} = v_0(\xi), \qquad \xi \in \boldsymbol{Z}^n. \tag{2.1}$$

Теорема 2.1. *Пусть выполнено условие* (1.2) *с любой константой* C *и выполнено условие* 1.1. *Тогда при* $\|v_0\|_s < a$, *где* a—*достаточно малое число, существует решение* $v(t, \xi)$ *задачи* (1.11), (2.1), *функционально-аналитически зависящее от* $v_0(\eta)$, *т.е.*

$$v(t, \xi) = v(t, \xi; v_0(\cdot)) = \sum_{k=1}^{\infty} \sum_{\zeta_1, \cdots, \zeta_k} B_k(t, \xi; \zeta_1, \cdots, \zeta_k)(v_0(\zeta_1), \cdots, v_0(\zeta_k)). \tag{2.2}$$

Решение $v(t, \xi)$ *удовлетворяет неравенству*

$$\sup_{t \geqq 0} \|v(t, \cdot)\|_s \leqq Ca. \tag{2.3}$$

В доказательстве этой теоремы [1] выводятся рекуррентные соотношения для коэффициентов B_k разложения (2.2) и с помощью метода производящих функций доказываются для B_k оценки, аналогичные (1.18′).

Для решения $u(t, x)$ задачи Коши для (1.1) при условиях периодичности (1.6) и при начальном условии $u|_{t=0} = u_0(x)$ имеет место разложение, двойственное к (2.2), точнее

$$\begin{aligned} u(t, x) &= u(t, x; u_0(\cdot)) \\ &= \sum_{k=1}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} G_k(t, x; y_1, \cdots, y_k)(u_0(y_1), \cdots, u_0(y_k)) dy_1 \cdots dy_k, \end{aligned} \tag{2.4}$$

где ядра G_k являются, вообще говоря, обобщенными функциями относительно $y_1, \cdots, y_k$ (см. [1]).

(б) Следующий простой пример показывает, что при больших $\|v_0\|_s$ ряд (2.2) может расходиться. Легко видеть (см. [4]), что функция $v(t, \xi) = 0$ при $\xi \neq 0$ и

$$v(t, 0) = -\kappa A(0) e^{-A(0)t} / (-A(0) + \beta\kappa(1 - e^{-A(0)t}))$$

удовлетворяет начальному условию $v|_{t=0} = v_0 = \kappa$ при $\xi = 0$ и $v|_{t=0} = v_0 = 0$ при $\xi \neq 0$ и системе уравнений (1.11), отвечающей скалярному уравнению

$$\partial u/\partial t = -A(D)u + \beta u^2, \qquad \beta > 0.$$

Функция $v(t, \xi)$ аналитична по κ для $|\kappa| < A(0)/\beta$ и имеет особенность в точке $\kappa = A(0)/\beta(1 - e^{-A(0)t})$.

Как показано в [4], уравнение Бюргерса

$$\partial u(t, x)/\partial t = \partial^2 u/\partial x^2 - \partial u^2/\partial x$$

при периодических граничных условиях и при начальном условии

$$u|_{t=0} = \eta_1 e^{ix} + \eta_{-1} e^{-ix}$$

имеет, и притом единственное, решение $u = u(t, x; \eta_1, \eta_{-1})$, аналитически зависящее от η_1 и η_{-1} при достаточно малой $|\eta_1| + |\eta_{-1}|$.

Однако, при любых фиксированных x, $0 \leqq x \leqq 2\pi$, и $t > 0$, существуют такие комплексные η_1 и η_{-1} (зависящие вообще говоря от t, x), при которых функция $u(t, x; \eta_1, \eta_{-1})$ имеет особенность (точнее, обращается в бесконечность).

(в) Пусть K_n—выпуклый конус, для которого существует такой вектор $e \subset K_n$, $|e| = 1$, что

$$(e, \eta) \geqq \gamma |\eta|, \quad \forall \eta \in K_n,$$

где $\gamma > 0$. Пусть $K = \boldsymbol{Z}^n \cap K_n \backslash \{0\}$. В том случае, когда функция $v_0(\xi)$ в начальном условии (2.1) удовлетворяет условиям

$$|v_0(\xi)| \leqq Ce^{M|\xi|}, \qquad v_0(\xi) = 0 \text{ при } \xi \notin K, \tag{2.5}$$

решение $v(t, \xi)$ задачи (1.11), (2.1) выражается рядом (2.2), в котором лишь конечное число слагаемых отлично от нуля. При этом, как показано в [3], при любых $v_0(\xi)$, удовлетворяющих (2.5), существует решение $v(t, \xi)$ задачи (1.11), (2.1), причём

$$|v(t, \xi)| \leqq C_1 e^{M_1|\xi| - \alpha t|\xi|}, \qquad \alpha > 0, \; v(t, \xi) = 0 \text{ при } \xi \notin K. \tag{2.6}$$

Пространство функций, у которых коэффициенты Фурье удовлетворяют условиям (2.5), обозначим через V'_K. Таким образом в [3] показано, что если начальное условие задачи (1.11), (2.1) принадлежит V'_K, то и её решение также принадлежит V'_K при каждом $t > 0$. Пространство V'_K может быть описано и в терминах функций в x-представлении, а именно: функция $u(z)$, $z = x + iy$, принадлежит V'_K, если она: (а) аналитична при $z = x + iy \in \boldsymbol{R}^n + iK^* + iy_0$, где K^*—конус, двойственный к K, $y_0 \in K^*$—некоторый вектор; (б) периодична по x с периодом 2π; (в) при $z \in \boldsymbol{R}^n + iK^* + iy_0$ удовлетворяет оценке

$$|u(z)| \leqq C_1 \frac{e^{-|y-y_0|\cos(y-y_0, \partial K)}}{(1 - e^{-|y-y_0|\cos(y-y_0, \partial K)})^n},$$

где $\cos(y, \partial K) = \inf_{\xi \in \partial K}(y/|y|, \xi/|\xi|)$. Отметим, что при достаточно больших t, решение $u(t, z) = F_{\xi \to z} v(t, \xi)$ определено и аналитично при действительных z (см. (2.6)).

3. Уравнения Навье—Стокса. Как показано в [1], Теоремы 1.1 и 2.1 обобщаются на случай системы уравнений Навье—Стокса. В этом случае появляются два новых обстоятельства. Во-первых, в результате исключения давления из системы Навье—Стокса получается система уравнений, содержащая нелинейные интегро-дифференциальные слагаемые. Во-вторых, матрица $A(\xi)$ у последней системы совпадает с $\xi^2 I$, где I—единичная матрица, следовательно, не выполнена при $\xi = 0$ оценка (1.2). Оба эти факта не препятствуют (см. [1]) справедливости Теорем 1.1, 2.1, в которых можно даже взять $s = 0$ (см. [1]).

4. Построение моментных функций и характеристического функционала статистических решений. (а) Пусть $\mu_0(dv)$—вероятностная мера, заданная на борелевской алгебре множеств пространства V_s. Под статистическим решением уравнений (1.11) понимается, как известно ([5], [6]), вероятностная мера $\mu_t(dv)$, зависящая от t, задаваемая формулой $\mu_t(\omega) = \mu_0(S_t^{-1}\omega)$, $\omega \subset V_s$, где S_t—оператор сдвига вдоль решений задачи (1.11), (2.1): $S_t v_0(\xi) = v(t, \xi)$. В том случае, когда оператор S_t задан лишь на шаре $B_a^s \subset V_s$, где a —такое же, как в п.2, а мера $\mu_0(\omega)$ имеет носитель в B_a^s, статистическое решение задается формулой:

$$\mu_t(\omega) = \mu_0(S_t^{-1}(\omega \cap Q_t)) \quad \text{где } Q_t = S_t B_a^s.$$

Здесь под $S_t^{-1}(\omega \cap Q_t)$ понимается множество таких $v \in B_a^s$, для которых $S_t v \in \omega \cap Q_t$.

Пусть $\mu_0(\omega)$—вероятностная мера с носителем в шаре B_a^s. Пусть $\Phi(-t, v)$ $(t > 0)$—аналитическое решение задачи Коши (1.16), (1.17), которое задается рядом (1.18), сходящимся на B_a^s, а начальное условие $\Psi(v) = \Phi(0, v)$ аналитично в шаре B_ρ^s, где $\rho = Ca$, C—такое же как в (2.3). Тогда справедливо следующее интегральное тождество:

$$\int \Psi(v)\mu_t(dv) = \int \Phi(-t, v)\mu_0(dv). \tag{4.1}$$

Здесь функционалы $\Psi(v)$ и $\Phi(-t, v)$ считаются как-нибудь непрерывно продолженными вне шаров B_ρ^s $(\rho = Ca)$ и B_a^s.

Доказательство (4.1) основано на том, что

$$\Phi(-t, v) = \Psi(S_t v), \qquad v \in B_a^s.$$

Для упрощения записи будем считать ниже, что $v(t, \xi)$—скалярная функция (т.е. $p = 1$). (Общий случай см. в [1].) Моментной функцией статистического решения $\mu_t(dv)$ называется функция

$$M_t(\xi_1, \cdots, \xi_k) = \int v(\xi_1) \cdots v(\xi_k)\mu_t(dv). \tag{4.2}$$

Теорема 4.1. *Пусть $\mu_0(\omega)$—вероятностная мера, сосредоточенная в шаре $B_a^s \in V_s$, где a—такое же, как в Теореме* 2.1, *а $\mu_t(\omega)$—соответствующее статистическое решение. Пусть $\Phi(-t, \xi_1, \cdots, \xi_k; v)$—аналитическое решение уравнения* (1.16) *при начальном условии*

$$\Phi(0, \xi_1, \cdots, \xi_k; v) = v(\xi_1) \cdots v(\xi_k).$$

Тогда для любого $t > 0$ имеет место разложение:

$$M_t(\xi_1, \cdots, \xi_k) = \sum_{r=k}^{\infty} \sum_{\eta_1, \cdots, \eta_r} \Phi_r(-t, \xi_1, \cdots, \xi_k; \eta_1, \cdots, \eta_r)\, M_0(\eta_1, \cdots, \eta_r), \tag{4.3}$$

где $\Phi_r(-t, \cdots)$—коэффициенты разложения в ряд по $v(\eta_i)$ функционала $\Phi(-t, \xi_1, \cdots, \xi_k; v(\cdot))$, а $M_0(\eta_1, \cdots, \eta_r)$—моментная функция меры $\mu_0(d\omega)$.

Аналогичная теорема справедлива также для моментных функций $\mathcal{M}_t(x_1, \cdots, x_k)$ в x-представлении:

$$\mathcal{M}_t(x_1, \cdots, x_k) = \int u(x_1) \cdots u(x_k)\, \hat{\mu}_t(du),$$

где $\hat{\mu}(du) = \mu_t(dv)$, если элементы du и dv соответствуют друг другу при отображении $F_{x \to \xi}(du) = dv$. В этом случае формула, аналогичная (4.3), имеет следующий вид:

$$\mathcal{M}_t(x_1, \cdots, x_k) = \sum_{r=k}^{\infty} \int_0^{2\pi} \cdots \int_0^{2\pi} K_r(-t, x_1, \cdots, x_k; y_1, \cdots, y_r)\, \mathcal{M}_0(y_1, \cdots, y_r)\, dy_1 \cdots dy_r,$$

где K_r—вообще говоря, обобщенные функции относительно $y_1, \cdots y_r$.

(б) Как известно, характеристический функционал меры $\mu_t(dv)$ задается формулой

$$\chi(t, w(\cdot)) = \int \exp\left(i \sum_{\xi} v(\xi) \cdot w(\xi)\right) \mu_t(dv),$$

где $v(\xi) \in V_s$, $w(\xi)$—любая функция с конечной нормой

$$[w]_s = \sup_{\eta} |w(\eta)|/(1 + |\eta|)^s. \tag{4.4}$$

Теорема 4.2. *Пусть $\mu_t(\omega)$ удовлетворяет условиям Теоремы* 4.1 *и $\chi(t, w)$ —её характеристический функционал. Тогда функция $\chi(t, w)$ при любом w с конечной нормой* (4.4) *разлагается в абсолютно сходящийся ряд*

$$\chi(t, w) = \Phi_{0,w}(-t) + \sum_{r=1}^{\infty} \sum_{\eta_1, \cdots, \eta_r} \Phi_{r,w}(-t, \eta_1, \cdots, \eta_r)\, M_0(\eta_1, \cdots, \eta_r),$$

где $\Phi_{r,w}$—коэффициенты разложения в степенной ряд по $v(\eta)$ функционала $\Phi_w(-t, v)$, который является в B_a^s аналитическим решением уравнения (1.16) *при следующем начальном условии:*

$$\Phi(0, w) = \exp\left(i \sum_{\xi} v(\xi) w(\xi)\right) = \sum_{n=0}^{\infty} \frac{i^n}{n!} \sum_{\xi_1, \cdots, \xi_n} v(\xi_1)\, w(\xi_1) \cdots v(\xi_n)\, w(\xi_n).$$

5. Асимптотическое разложение моментных функций статистических решений. Через RV_s обозначается подпространство V_s, состоящее из тех элементов $v(\eta) \in V_s$, которые удовлетворяют условию $v(-\eta) = \overline{v(\eta)}$. Отметим, что элементы $v_0(\eta) \in RV_s$ являются коэффициентами Фурье вещественных функций $u_0(x) = F^{-1}v_0(\eta)$.

Допустим, что уравнение (1.1)—вещественное и выполнены следующие условия: (1) Для любых $v_0 = v_0(\eta)$, принадлежащих некоторому банахову пространству B_0, задача (1.1), (1.6) при начальном условии

(5.1) $$u\big|_{t=0} = u_0(x) = \sum_{\eta} v_0(\eta) e^{i\eta x}, \qquad v_0(\eta) \in B_0,$$

имеет, и притом единственное, решение $u(t, x)$, принадлежащее банахову пространству B_1; (2) имеет место вложение $B_0 \Subset RV_s$; (3) для решения $u(t, x)$ задачи (1.5), (1.6), (5.1) при всех $t \in [0, T)$ и $\xi \in \boldsymbol{Z}^n$ определены коэффициенты Фурье $v(t, \xi; v_0) = F_{x\to\xi}\, u(t, x)$, причем

$$\big|v(t, \xi; v_0)\big| \leqq C_2 \|v_0\|_{B_0},$$

где C_2 не зависит от $v_0 \in B_0$, $t \in [0, T)$ и $\xi \in \boldsymbol{Z}^n$; (4) Коэффициенты Фурье $v(t, \xi; v_0)$ непрерывно зависят от $v_0 \in B_0$:

$$\big|v(t, \xi; v_0) - v(t, \xi; v_0')\big| \to 0 \quad \text{при } \|v_0 - v_0'\|_{B_0} \to 0.$$

Отметим, что все эти условия выполнены для случая монотонных параболических уравнений [7], [8], в частности, например, для уравнения вида

$$\frac{\partial u(t, x)}{\partial t} = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[\left(1 + \left(\frac{\partial u}{\partial x_i}\right)^{p-2}\right)\frac{\partial u}{\partial x_i}\right] \qquad \text{(см. [7], [8])}.$$

При этом в качестве B_0 можно взять пространство Соболева $H_\rho \subset RV_s$, $\rho > s + n/2$, всех функций $v(\xi)$, для которых конечна норма

$$\|v\|_{H_\rho} = \left(\sum_{\xi} (1 + |\xi|)^{2\rho} |v(\xi)|^2\right)^{1/2}.$$

Условия (1)—(4) выполнены также для двумерной системы уравнений Навье—Стокса.

Пусть $\mu(dv)$—вполне аддитивная вероятностная борелевская мера на пространстве B_0, для которой при любом $N \geqq 1$

(5.2) $$\int \|v_0\|_{B_0}^N \mu(dv_0) < +\infty.$$

Заметим, что $\mu(dv_0)$ может иметь неограниченный носитель в B_0.

Обозначим через $\mu_\sigma(dv)$ меру, определенную равенством $\mu_\sigma(\omega) = \mu(G_\sigma^{-1}\omega)$, где G_σ—оператор растяжения: $G_\sigma v = \sigma v$ для $v \in B_0$. Рассмотрим моментные функции относительно мер μ_σ

$$M_t(\xi_1, \cdots, \xi_k, \sigma) = \int v(t, \xi_1; v_0) \cdots v(t, \xi_k; v_0)\, \mu_\sigma(dv_0).$$

Теорема 5.1. *Пусть для задачи* (1.5), (1.6), (5.1) *выполнены условия* (1)—(4). *Тогда моментные функции* $M_t(\xi_1, \cdots, \xi_k, \sigma)$ *разлагаются при* $\sigma \to 0$ *в асимптотический ряд*

(5.3) $$M_t(\xi_1, \cdots, \xi_k, \sigma) \sim \sum_{r=k}^{\infty} A_r(t, \xi_1, \cdots, \xi_k)\, \sigma^r,$$

где $A_r(t, \xi_1, \cdots, \xi_k)$ *выражаются через коэффициенты* $B(t, \xi_i; \eta_1, \cdots, \eta_r)$ $(i = 1, \cdots, k)$ *разложения* (2.2).

Точнее, для любого $N \geqq k$

$$\left|M_t(\xi_1, \cdots, \xi_k, \sigma) - \sum_{r=k}^{N} A_r(t, \xi_1, \cdots, \xi_k)\sigma^r\right| \leqq C_{N+1}\, \sigma^{N+1},$$

где константы C_{N+1} не зависят от t, $\xi_1, \cdots, \xi_k$, σ, но, вообще говоря, неограниченно растут при $N \to \infty$ (см. [4]).

Пусть $\boldsymbol{Z}_1^n \subset \boldsymbol{Z}^n$—подмножество целочисленной решетки, такое, что: (а) если $\eta \neq 0$ и $\eta \in \boldsymbol{Z}_1^n$, то—$\eta \notin \boldsymbol{Z}_1^n$; (б) $\boldsymbol{Z}_1^n \cup (-\boldsymbol{Z}_1^n) = \boldsymbol{Z}^n$.

В случае $v(\eta) = a(\eta) + ib(\eta) \in B_0 = H_\rho$ в качестве меры $\mu(dv)$, удовлетворяющей перечисленным выше условиям, можно взять гауссовскую меру, считая компоненты $a(\eta)$, $b(\eta)$ ($\eta \in \boldsymbol{Z}_1^n$) независимыми случайными величинами, распределенными по гауссовскому закону с математическим ожиданием, равным нулю, и дисперсией, равной $\sigma^2(\eta) = (1 + |\eta|)^{-2\lambda}$, $\lambda > \rho + n/2$.

Цитированная Литература

1. М.И. Вишик и А.В. Фурсиков, *Аналитические первые интегралы нелинейных параболических в смысле И.Г. Петровского систем дифференциальных уравнений и их приложения*, Успехи матем. наук **29** (1974), 123–153.

2. ———, *Аналитические первые интегралы нелинейных параболических уравнений и их приложения*, Мат. сборник **92** (1973), 347–377.

3. ———, *Некоторые вопросы теории нелинейных эллиптических и параболических уравнений*, Мат. сборник **94** (1974), 324–357.

4. ———, *Асимптотическое разложение моментных функций решений нелинейных параболических уравнений*, Мат. сборник **95** (1974), 588–605.

5. E. Hopf, *Statistical hydromechanics and functional calculus*, J. Rational Mech. Anal. **1** (1952), 87–123.

6. Ч. Фояш, *Статистические решения нелинейных эволюционных уравнений*, Математика **17**:3 (1973), 90–113.

7. М.И. Вишик, *О решении краевых задач для квазилинейных параболических систем любого порядка*, Мат. сборник **59** (1962), 289–325.

8. J.L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969.

Московский Государственный Университет
Москва, СССР

Section 14

Ordinary Differential Equations and Dynamic Systems

Труды Международного Конгресса Математиков
Ванкувер, 1974

Геодезические в Финслеровой Геометрии

Д. В. Аносов

1. В финслеровой геометрии, как и в римановой, рассматривается гладкое многообразие M, для касательных векторов которого определено понятие длины (так что можно говорить о длине параметризованной кривой; последняя длина равна интегралу от длины вектора скорости). Отличие от римановой геометрии состоит в том, что выражение для длины может быть более общим. Именно, длина вектора $v \in T_xM$ даётся функцией $L(x, v)$, которая обращается в нуль лишь при $v = 0$, положительна при $v \neq 0$ и является положительно-однородной первой степени по v. На L налагаются два общих условия:

Условие достаточной гладкости вне нулевого сечения касательного расслоения;

Условие выпуклости "единичных сфер" $L(x,v) = 1$ во всех касательных пространствах T_xM, усиленное ещё дополнительным требованием, чтобы кривизна "единичной сферы" (вычисленная по отношению к произвольной эвклидовой метрике в T_xM) нигде не обращалась в нуль.

Геодезическими линиями в финслеровой геометрии называют экстремали функционала длины. Мы имеем здесь дело с вариационной задачей в параметрической форме. В вариационном исчислении "единичные сферы" называют индикатрисами, а условие их выпуклости вместе с усиливающим его требованием о кривизне эквивалентно так называемому усиленному условию Лежандра, при выполнении которого говорят о положительной регулярности вариационной задачи.

Иногда к сформулированным двум общим условиям добавляют еще третье: индикатриса центрально-симметрична, т.е. $L(x,v) = L(x, -v)$. Эквивалентная формулировка: длина каждой параметризованной кривой совпадает с длиной той же самой кривой, проходимой в обратном направлении. При выполнении

этого условия говорят о "симметричности" или "обратимости" (последний термин предпочтительнее, ибо первый имеет в римановой геометрии другой смысл).

Помимо этих общих условий, в различных задачах на L могут налагаться различные другие ограничения, связанные со спецификой того или иного вопроса.

2. Большинство (если не все) содержательных результатов финслеровой геометрии связано с геодезическими линиями. Цель настоящего доклада состоит не в обзоре всех этих результатов, а скорее в изложении моей точки зрения в связи с некоторыми из них.

Рассматривая движение по геодезической с постоянной скоростью, обычным образом приходим к геодезическому потоку. (Этот термин, в зависимости от контекста, относится как к динамической системе, фазовым пространством которой служит всё касательное расслоение TM, так и к её ограничению на инвариантную гиперповерхность, образованную векторами единичной длины.) Геодезический поток сохраняет естественную инвариантную меру—объём в фазовом пространстве—и является гамильтоновой системой (при надлежащем понимании этого термина).

Мой интерес к финслеровой геометрии связан отчасти с тем, что таким путём она приводит к довольно широкому классу динамических систем, который допускает применение геометрических понятий и соображений при формулировке задач и при их исследовании. Отчасти же этот интерес вызван желанием выделить в задачах римановой геометрии то, что связано с одной их вариационной природой, а от остальных специфических особенностей римановой геометрии не зависит. Имея в виду эти два мотива, я остановлюсь на двух вопросах, которые в римановой геометрии принадлежат к наиболее известным и восходят, соответственно, к Адамару и Пуанкаре: свойства геодезических потоков на многообразиях отрицательной кривизны и замкнутые геодезические на многообразиях, гомеоморфных сфере. Оказывается, что в ряде случаев является более или менее существенной обратимость метрики, специфическое же выражение для римановой метрики не играет никакой роли, по крайней мере, в большинстве случаев (но в одном интересном случае это пока не выяснено, см. ниже).

3. Различные авторы предлагали неэквивалентные определения параллельного перенесения и ковариантного дифференцирования в финслеровой геометрии (см. подробное изложение в [1]; с тех пор появились ещё новые варианты). Это показывает, что на самом деле в финслеровой геометрии нет "хорошего" аналога названных понятий. Однако всё, что нам требуется—это приведение уравнений в вариациях для геодезического потока к столь же удобному виду, какой они имеют в римановой геометрии при использовании параллельного перенесения ортонормированного базиса вдоль рассматриваемой геодезической. Это проще всего сделать, рассматривая вспомогательную риманову метрику, "соприкасающуюся" (в некотором смысле) с финслеровой вдоль рассматриваемой геодезической. В уравнения в вариациях войдёт кривизна

соприкасающейся метрики по двумерному направлению, содержащему касательную к рассматриваемой геодезической. Эта кривизна не зависит от случайностей выбора соприкасающейся римановой метрики, т.е. является некоторой внутренней характеристикой финслерова многообразия. Она зависит не только от точки и двумерной касательной плоскости в этой точке, но и от некоторого одномерного направления в этой точке (направления исходной геодезической); поэтому в финслеровой геометрии следует говорить о кривизне в направлении двумерного флага (т.е. фигуры, состоящей из плоскости с выделенной на ней прямой или даже с выделенным лучом,—последнее в необратимом случае).

После сказанного нетрудно понять, что для геодезических потоков на замкнутых финслеровых многообразиях, кривизна которых в каждой точке и в направлении любого двумерного флага отрицательна, выполняются те же самые "условия У", что и в римановом случае (см. [2], [3]). В частности, эти потоки эргодичны. Чтобы сделать более сильное заключение об их эргодических свойствах, нужно доказать, что у них нет непрерывных собственных функций. В римановом случае это делается либо на основании теоремы Арнольда о числах вращения (обсуждение и литературу см. напр., в [3]; там имеются некоторые усовершенствования по сравнению с первоначальной формулировкой Арнольда, указанные Маргулисом, но само понятие "числа вращения" формально не используется, так что в этом отношении [3] надо использовать вместе с цитированной там литературой), либо на основании теоремы Риба [9] о невозможности трансверсальных сечений для геодезических потоков. Последняя теорема в равной мере годится и для всех финслеровых метрик, включая необратимые, для которых первая теорема не верна.

В связи со сказанным о теореме Риба любопытно заметить, что методы Катка, о работе которого [5] еще будет речь дальше, позволяют построить пример такой необратимой финслеровой метрики, для которой геодезический поток эргодичен, но имеет собственную функцию. Согласно Катку, можно даже добиться, чтобы последняя имела сравнительно мало разрывов—все они расположены на двух окружностях, причем у функции существуют пределы по направлениям при стремлении к этим окружностям; имеется трансверсальное к потоку многообразие с краем, ограниченное этими самыми окружностями.

Вернемся к "условиям У". Клингенберг [10] исследовал вопрос о том, что можно сказать о свойствах риманова многообразия, если известно, что соответствующий поток удовлетворяет "условиям У". Одно из основных рассуждений в [10] состоит в доказательстве того, что соответствующий геодезический поток удовлетворяет тогда и другому условию, хорошо известному в дифференциальной геометрии—у него нет сопряженных точек. Для необратимых финслеровых метрик рассуждения Клингенберга в этом месте не проходят и должны быть заменены другими. Можно предложить доказательство отсутствия сопряженных точек, основанное в двумерном случае на упомянутой выше теореме Риба, а в более высоких размерностях—на некотором ее обобщении. Рассмотрим те точки фазового пространства, для

которых проекция "сжимающегося" касательного подпространства (подпространства X_w^k в обозначениях [2]) на исходное финслерово многообразие вырождается. Оказывается, что в двумерном случае эти точки образуют в фазовом пространстве трансверсальную к потоку поверхность, что противоречит теореме Риба. В более высоких размерностях из-за возможных изменений кратности вырождения многообразия может и не получиться, но оказывается, что все же получается нечто вроде трансверсально иммерсированного многообразия коразмерности один. Вся неточность этой формулировки связана только с необходимостью более детального описания этого понятия "иммерсии, трансверсальной к потоку",—из-за обычной в подобных вопросах негладкости тут необходима некоторая осторожность, не представляющая, однако, ничего принципиально нового по сравнению с [2].

4. В задаче о замкнутых геодезических на сфере наиболее законченные результаты относятся в настоящее время к двумерной сфере S^2. Они были получены Люстерником и Шнирельманом в 1929 г. и подробно изложены в [4]. Это изложение ограничивается римановыми метриками; правда, в одном месте мимоходом сказано (часть II, §6), что можно говорить и о замкнутых экстремалях для положительно регулярной вариационной задачи, однако дальнейшие рассуждения (в §8) непосредственно относятся к римановому случаю. Как бы то ни было, соответствующие рассуждения можно модифицировать таким образом, чтобы распространить их на финслеровы метрики. Для обратимых финслеровых метрик окончательный результат—тот же, что в [4]. Для необратимых метрик можно гарантировать только существование двух замкнутых геодезических.

Последний результат представляет тот интерес, что он является окончательным. Соответствующий пример указан Катком ([5, §6, п. 5]). В [5] этот пример, сам по себе простой, является исходным этапом сложной конструкции, приводящей в пределе к необратимой финслеровой метрике класса C^∞ на S^2, для которой геодезический поток эргодичен, причём можно обеспечить, чтобы предельная метрика попрежнему имела только две замкнутые геодезические. Вместе с тем упомянутый пример наводит на мысль: хотя число замкнутых геодезических, которое можно гарантировать или хотя бы ожидать на основании современного вариационного исчисления в целом, обычно весьма невелико, быть может, оно является неулучшаемым?

Для n-мерной сферы S^n пример Катка даёт необратимую финслерову метрику, сколь угодно близкую к "стандартной" (к метрике постоянной кривизны) и имеющую $2[(n-1)/2]$ замкнутых геодезических. Это число совпадает с оценкой снизу, которую естественно ожидать для необратимых финслеровых метрик на S^n и которую можно доказать для метрик, достаточно близких к "стандартной".

5. Выше я дважды упомянул о числе замкнутых геодезических, "ожидаемом" на основании современного вариационного исчисления в целом. Я имею в виду число различных точек в функциональном пространстве замкнутых

кривых, на которых "висят" циклы, представляющие гомологии этого пространства (точнее—некоторую часть гомологий, их "начальную серию"). В обзоре [6] об этом числе говорится как о числе "геометрически различных" решений вариационной задачи. Эта исторически сложившаяся терминология неудачна, ибо не ясно как раз то, соответствуют ли эти различные точки в функциональном пространстве различным с геометрической точки зрения кривым в исходном многообразии (ведь две различные точки могли бы соответствовать одной и той же замкнутой кривой, обходимой различное число раз). В настоящее время в этом состоит основная трудность при $n > 2$, которую удалось преодолеть лишь для римановых метрик с определёнными ограничениями на кривизну (см. [6]). При этом используется, что при этих ограничениях имеется некоторая оценка длины замкнутых геодезических. Неизвестно, верен ли аналог этой оценки для финслеровых метрик. Это представляется актуальной задачей.

6. Вернёмся к двумерной сфере, которой занимался ещё Пуанкаре [7]. Его рассуждения бездоказательны, но дают некоторые эвристические доводы в пользу следующего утверждения: если длины наших трёх геодезических различны, то две из них будут устойчивы в линейном приближении, а третья —неустойчива. В [8] сказано (§7), что это утверждение можно строго доказать посредством рассуждений типа имеющихся в [7], в чём я сомневаюсь; во всяком случае, подробной публикации не появилось. Тем не менее само это утверждение верно.

Литература

1. H. Rund, *The differential geometry of Finsler spaces*, Die Grundlehren der math. Wissenschaften, Band 101, Berlin-Göttingen-Heidelberg, Springer-Verlag, 1959.

2. Д.В. Аносов, *Геодезические потоки на замкнутых римановых многообразиях отрицательной кривизны*, Труды МИАН, т. 90, "Наука", Москва, 1967.

3. Д.В. Аносов и Я.Г. Синай, *Некоторые гладкие эргодические системы*, Успехи матем. наук **22** (1967), 107–172.

4. Л. Люстерник и Л. Шнирельман, *Топологические методы в вариационных задачах и их приложения к дифференциальной геометрии поверхностей*, Успехи матем. наук **2** (1947), 166–217.

5. А.Б. Каток, *Эргодические возмущения вырожденных интегрируемых гамильтоновых систем*, Известия АН СССР, сер. матем. **37** (1973), 539–576.

6. С.И. Альбер, *Топология функциональных многообразий и вариационное исчисление в целом*, Успехи матем. наук **25** (1970), 57–122.

7. А. Пуанкаре, *О геодезических линиях на выпуклых поверхностях*, Избранные труды, т. II, "Наука", Москва, 1972, стр. 735–774.

8. Л.А. Люстерник, *Топология и вариационное исчисление*, Успехи матем. наук **1** (1946), 30–56.

9. G. Reeb, *Sur certaines propriétés topologiques des trajectoires des systèmes dynamiques*, Acad. Roy. Belg., Cl. Sci. Mém. Coll. in-8° **27** (1952).

10. W. Klingenberg, *Riemannian manifolds with geodesic flow of Anosov type*, Ann. of Math. (2) **99** (1974), 1–13.

Математический Институт им. В.А. Стеклова АН СССР
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Symbolic Dynamics for Hyperbolic Flows

Rufus Bowen

Let f_t ($t \in R$) be a differentiable flow on a compact manifold M. A compact invariant set Λ containing no fixed points is called *hyperbolic* if the tangent bundle restricted to Λ can be written as the Whitney sum of three Df_t-invariant continuous subbundles $T_\Lambda M = E + E^s + E^u$, where E is the one-dimensional bundle tangent to the flow, and there are constants $c, \lambda > 0$ so that

(a) $\|Df_t(v)\| \leqq ce^{-\lambda t}\|v\|$ for $v \in E^s$, $t \geqq 0$,
(b) $\|Df_{-t}(v)\| \leqq ce^{-\lambda t}\|v\|$ for $v \in E^u$, $t \geqq 0$.

A hyperbolic set Λ is called *basic* if

(a) the periodic orbits of $f_t|\Lambda$ are dense in Λ,
(b) $f_t|\Lambda$ is a topologically transitive flow,
(c) there is an open set $U \supset \Lambda$ with $\Lambda = \bigcap_{t\in R} f_t\, U$.

Basic hyperbolic sets occur in Smale's Axiom A flows [11], a class containing all known structurally stable flows. An important special case is an Anosov flow; here M itself is a hyperbolic set.

We will outline a method for studying the structure of basic sets, namely symbolic dynamics. The space $\Sigma_n = \prod_z\{1, \cdots, n\}$ is compact when given the product topology (and $\{1, \cdots, n\}$ the discrete topology). One writes $\boldsymbol{x} = (x_i)_{i=-\infty}^{\infty}$ for a point in Σ_n and $x_i = (\boldsymbol{x})_i$. The *shift* homeomorphism $\sigma:\Sigma_n \to \Sigma_n$ is defined by $\sigma(\boldsymbol{x})_i = x_{i+1}$. For A an $n \times n$ matrix of 0's and 1's, the set $\Sigma_A = \{\boldsymbol{x} \in \Sigma_n: A_{x_i, x_{i+1}} = 1 \text{ for all } i\}$ is compact and σ-invariant. A basic hyperbolic set Λ will be closely related to a certain symbolic space Σ_A.

1. The model. For $g:\Sigma_A \to R$ a positive continuous function one considers

$$Y = \{(x, s): s \in [0, g(x)]\} \subset \Sigma_A \times R$$

and then identifies $(x, g(x))$ with $(\sigma(x), 0)$ for all $x \in \Sigma_A$ to get a new space $\Lambda(A, g)$. This is a compact metric space, and one can define a flow ϕ_t on it by

$$\phi_t(x, s) = (x, s + t) \quad \text{for } s + t \in [0, g(x)]$$

and remembering identifications. This flow is the *suspension* of $\sigma | \Sigma_A$ under the function g. There is a restriction on g which will be important to us. Let

$$\operatorname{var}_n(g) = \sup \{|g(x) - g(y)| : x, y \in \Sigma_A \text{ and } x_i = y_i \ \forall\, i \in [-n, n]\}$$

and say $g \in \mathscr{F}_A$ if $\operatorname{var}_n(g) \leqq c\alpha^n$ for some constants c and $\alpha \in (0, 1)$.

THEOREM [1]. *The following two classes of flows coincide:*

(a) $f_t | \Lambda$ *with* Λ *a one-dimensional basic hyperbolic set for a differentiable flow,*

(b) $\phi_t | \Lambda(A, g)$ *with* A *such that* $\sigma | \Sigma_A$ *is transitive and* $g: \Sigma_A \to R$ *strictly positive with* $g \in \mathscr{F}_A$.

THEOREM [2]. *Let* $f_t | \Lambda$ *be a hyperbolic basic set. Then one can find* A, $g \in \mathscr{F}_A$, *and a continuous surjection* $\pi: \Lambda(A, g) \to \Lambda$ *so that* $f_t \pi = \pi \phi_t$.

The first theorem identifies the simplest basic sets and the second shows how to relate any basic set to one of these simple ones. One then studies basic sets in general by first considering the simple models and then "pushing down" results via π. Below we will see several examples of this.

For the Anosov case, π was constructed independently by M. Ratner [7] and anticipated by Ja. Sinaĭ [10]. For diffeomorphisms (instead of flows) the analogous results are due to Sinai and the author.

2. Recurrence properties.

THEOREM [2]. *The point* $x = \pi(x')$ *is periodic, transitive, recurrent or almost periodic w.r.t.* f_t *if and only if* x' *is w.r.t.* ϕ_t. *The minimal sets of* $f_t | \Lambda$ *are one dimensional.*

One should note that $x' = (x, s)$ has one of these properties iff x does (w.r.t. $\sigma | \Sigma_A$). These results are proved by finding an N so that card $\pi^{-1}(x) \leqq N$ for all $x \in \Lambda$ and showing that π is a local homeomorphism over various subsets of Λ. The above theorem generalizes results of M. Morse [6].

An interesting problem is to calculate the number $N_\tau(f_t)$ of closed orbits γ in Λ with some period equal to τ. Using a method of Manning [4], one can find matrices $A_1, \cdots, A_m$ and positive functions $g_i \in \mathscr{F}_{A_i}$ so that [2]

$$N_\tau(f_t) = \sum_{i=1}^{m} \varepsilon_i N_\tau(A_i, g_i).$$

Here $\varepsilon_i = \pm 1$ and $N_\tau(A_i, g_i) = N_\tau(\phi_t)$ on $\Lambda(A_i, g_i)$. This reduces the counting problem to the case of the model $\Lambda(A, g)$. The model is complicated and it is unknown for instance whether its zeta function (see Smale [11]) is meromorphic in the whole plane. This is true (Manning and the author) if g is locally constant.

The growth rate of $N_\tau(f_t)$ as $\tau \to \infty$ equals the topological entropy $h(f_1)$, the closed orbits are equidistributed as $\tau \to \infty$ w.r.t. a measure μ, and μ is the unique

invariant probability measure on Λ maximizing entropy. These facts can be derived using symbolic dynamics (i.e., via π) though they were first proved in other ways.

3. Ergodic theory. Let m be some smooth measure on M and define

$$W^s(\Lambda) = \{y \in M: d(f_t y, \Lambda) \to 0 \text{ as } t \to \infty\}.$$

The basic set Λ is called an *attractor* if $W^s(\Lambda)$ is a neighborhood of Λ in M.

THEOREM. *Assume f_t is C^2. Then $m(W^s(\Lambda)) > 0$ iff Λ is an attractor. If Λ is an attractor, then there is an invariant probability measure μ^+ on Λ so that*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T w(f_t y)\, dt = \int w\, d\mu^+$$

for m-almost every $y \in W^s(\Lambda)$ and every $w \in C(M)$.

This theorem is due to Ruelle and the author [3]. The convergence part or something similar was proved earlier for diffeomorphisms by Ruelle [8], for Anosov diffeomorphisms by Sinaĭ [9], and for Anosov flows by Margulis [5].

Let $\lambda_t(x)$ be the Jacobian of the linear map $Df_t: E^u_x \to E^u_{f_t x}$ and define $\phi: \Lambda \to R$ by

$$\phi(x) = -\left.\frac{d \ln \lambda_t(x)}{dt}\right|_{t=0}.$$

Then μ^+ is the unique invariant measure ν on Λ which maximizes the expression [3] $h_\nu(f_1) + \int \phi\, d\nu$. Analogously to statistical mechanics, μ^+ is called the *equilibrium state* of ϕ. This concept is useful in symbolic dynamics because it reacts well to π. We remark that Sinaĭ [10] used the slightly different formalism of Gibbs state.

If $f_t|\Lambda$ is topologically mixing, then the measure μ^+ is Bernoulli. This follows from the corresponding theorem on the model flow $\Lambda(A, g)$ due to Ratner and Bunimovic (independently). Finally we remark that all this theory depends on f_t being C^2; one can find a C^1 horseshoe for instance with positive Lebesgue measure.

4. Other properties. There are two directions in which symbolic dynamics has been applied to diffeomorphisms, though not yet for flows. The first is homology theory with papers by Ruelle and Sullivan, Shub and Williams, and the author. The second is the dynamics of the unstable foliation. If Λ is an attractor, dim $E^u = 1$ and E^u is orientable, then one can define a W^u-flow on Λ whose orbits are the unstable manifolds $W^u(x)$. For Λ a mixing hyperbolic attractor *for a diffeomorphism* B. Marcus has shown that this flow is uniquely ergodic. For flows this is still open; if it is true, it would generalize Furstenberg's theorem on the horocycle flow.

ADDED IN PROOF. Marcus has now proved this result for Anosov flows, and Marcus and the author have done this for Axiom A flows.

References

1. R. Bowen, J. Differential Equations **12** (1972), 173–179.
2. ———, Amer. J. Math. **95** (1973), 429–460.
3. R. Bowen and D. Ruelle, *The ergodic theory of Axiom A flows*, Invent. Math. (to appear).

4. A. Manning, *Axiom* A *diffeomorphisms have rational zeta functions*, Bull. London Math. Soc. **3** (1971), 215–220. MR **44** #5982.

5. G. A. Margulis, *Certain measures that are connected with Y-flows on compact manifolds*, Funkcional. Anal. i Priložen. **4** (1970), no. 1, 62–76 = Functional Anal. Appl. **4** (1970), 55–67. MR **42** #7865.

6. M. Morse, *Recurrent geodesics on a surface of negative curvature*, Trans. Amer. Math. Soc. **22** (1921), 84–110.

7. M. Ratner, *Markov partitions for Anosov flows on n-dimensional manifolds*, Israel J. Math. **15** (1973), 92–114.

8. D. Ruelle, *A measure associated with Axiom* A *attractors*, Amer. J. Math. (to appear).

9. Ja. G. Sinaĭ, *Markov partitions and Y-diffeomorphisms*, Funkcional. Anal. i Priložen. **2** (1968), no. 1, 64–89 = Functional Anal. Appl. **2** (1968), 61–82. MR **38** #1361.

10. ———, *Gibbs measures in ergodic theory*, Uspehi Mat. Nauk **27** (1972), no. 4 (**166**), 21–64 = Russian Math. Surveys **27** (1972), no. 4, 21–70.

11. S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817. MR **37** #3598; erratum, **39**, p. 1593.

UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

On Generators in Ergodic Theory

W. Krieger

We are concerned with measure preserving transformations T of a Lebesgue measure space $(X, \mathscr{B}, \mu)$ where $\mu(X) = 1$. Consider a finite partition $\mathscr{P} = (P_1, \cdots, P_n)$ or a countably infinite partition $\mathscr{P} = (P_1, \cdots)$ of X. Such a partition $\mathscr{P}$ is said to be a generator of T if

$$\mathscr{B} = \bigvee_{i \in \mathbf{Z}} T^i \mathscr{P} \pmod{\mu}.$$

More generally a sub-σ-algebra $\mathscr{A}$ of $\mathscr{B}$ is called a generator for T if

$$\mathscr{B} = \bigvee_{i \in \mathbf{Z}} T^i \mathscr{A} \pmod{\mu}.$$

For a survey of the theory of generators see U. Krengel's 1971 Prague conference address [23].

Generators have been of use in entropy theory. The entropy of the partition $\mathscr{P}$ is defined by

$$h(\mathscr{P}) = -\sum_m \mu(P_m) \log \mu(P_m).$$

The mean entropy of the partition $\mathscr{P}$ with respect to the measure preserving transformation T is defined by

$$h(\mathscr{P}, T) = \lim_{k \to \infty} \frac{1}{k} h\left(\bigvee_{1 \le i \le k} T^{-i} \mathscr{P} \right),$$

and the entropy of T is then defined by $h(T) = \sup h(\mathscr{P}, T)$ where the supremum is taken over all finite partitions. The Kolmogoroff-Sinai theorem states that $h(T) = h(\mathscr{P}, T)$ if $\mathscr{P}$ is a finite generator for T. As a consequence one has that $h(T) \leqq n$ if T has a generator of size n. In the converse direction it is known that every ergodic measure preserving transformation T with entropy $h(T) < \infty$ has a generator with

no more than $e^{h(T)} + 1$ elements [24], [25]. A simpler proof of this result has been given by M. Denker [4]. Denker observed further that finite generators are dense in the set Π of all partitions with mean entropy $h(T) < \infty$ if Π is given the entropy metric $|\mathscr{P}, \mathscr{Q}| = 2h(\mathscr{P} \vee \mathscr{Q}) - h(\mathscr{P}) - h(\mathscr{Q})$, $\mathscr{P}, \mathscr{Q} \in \Pi$ [5].

A partition $\mathscr{P}$ is said to be a strong generator for T if $\mathscr{B} = \bigvee_{i \in N} T^{-i}\mathscr{P} \pmod{\mu}$. Finite strong generators (in fact strong generators of size 2) exist if and only if $h(T) = 0$. Recall that by a theorem of Parry and Rohlin every aperiodic transformation T has a countable strong generator. $\mathscr{P}$ is a strong generator for T if and only if its remote past equals $\mathscr{B}$,

$$\mathscr{B} = \bigcap_{j \in N} \bigvee_{i \geqq j} T^{-i}\mathscr{P} \pmod{\mu}.$$

Partitions with such a property have been called deterministic. Define $\mathscr{P}$ as bilaterally deterministic if

$$\mathscr{B} = \bigcap_{j \in N} \bigvee_{|i| \geqq j} T^{-i}\mathscr{P} \pmod{\mu}.$$

D. Ornstein and B. Weiss have shown that every ergodic measure preserving transformation with finite entropy has a bilaterally deterministic finite partition [33]. This answers also a question that was raised by O. E. Lanford and D. Ruelle in connection with K-automorphisms [31].

Consider the shift space $\Omega_n = \prod_{i \in \mathbf{Z}} \{1, \cdots, n\}$ over the alphabet $\{1, \cdots, n\}$. Ω_n is given the product topology of the discrete topologies. The n-shift S on Ω_n, $(Sx)_i = x_{i-1}$, $i \in \mathbf{Z}$, $x = (x_i)_{i \in \mathbf{Z}} \in \Omega_n$ is a homeomorphism of Ω_n. Let now $\mathscr{P} = (P_1, \cdots, P_n)$ be a generator for the measure preserving transformation T. One obtains a mapping of X into Ω_n if one sets

$$(Ux)_i = m, \quad \text{if } T^i x \in P_m,\ 1 \leqq m \leqq n,\ i \in \mathbf{Z},\ \text{f, } \mu\text{-a.a. } x \in X.$$

Setting for a Borel set $A \subset \Omega_n$, $U\mu(A) = \mu(U^{-1}A)$, one transports the measure μ from X to Ω_n, and in this way one has produced a shift-invariant measure $\nu = U\mu$ on Ω_n. As a consequence of the generation property of $\mathscr{P}$ the systems (X, μ, T) and (Ω_n, ν, S) are isomorphic in the sense of measure theory. For the case of the shift the finite generator theorem provides therefore an answer to the following question: Does a given homeomorphism Φ of a compact metric space E have an invariant probability measure ν such that the system (E, ν, Φ) is isomorphic to a given system (X, μ, T)? The notion of topological entropy furnishes a necessary condition for such an imbedding. Topological entropy is defined as follows [1]: For an open cover $\mathscr{C}$ of E let $h(\mathscr{C})$ be the logarithm of the minimal cardinality of a subcover of $\mathscr{C}$. Then set, for a homeomorphism Φ of E,

$$h(\mathscr{C}, \Phi) = \lim_{k \to \infty} \frac{1}{k} \left(\bigvee_{1 \leq i \leq k} \Phi^{-i}\mathscr{C} \right)$$

and define the topological entropy of Φ by $h(\Phi) = \sup h(\mathscr{C}, \Phi)$ where the supremum is taken over all open covers. The n-shift has topological entropy $\log n$. One has for all Φ-invariant probability measures μ that $h(\mu, \Phi) \leqq h(\Phi)$ [12]. In fact, one has $h(\Phi) = \sup h(\mu, \Phi)$ where the supremum is taken over all Φ-invariant probability

measures μ [10]. This notion of topological entropy is an invitation to define the topological analog of a generator. This was done by H. Keynes and J. Robertson [20] as follows: An open cover $\mathscr{C}$ is a topological generator of the homeomorphism Φ if for all $(C(i))_{i\in\mathbf{Z}} \in \mathscr{C}^{\mathbf{Z}}$ the set $\bigcap_{i\in\mathbf{Z}} \overline{C(i)}$ contains at most one point. This definition stands the test: If $\mathscr{C}$ is a topological generator for Φ then $h(\Phi) = h(\mathscr{C}, \Phi)$ [20]. Φ is called expansive if there exists a $\delta > 0$ such that, for all $x, y \in E$, $x \neq y$, there is an $i \in \mathbf{Z}$ such that $d(\Phi^i x, \Phi^i y) > \delta$ where d is a metric of E. Exactly the expansive homeomorphisms have topological generators [20]. The expansive homeomorphisms of the Cantor discontinuum are given by the subshifts; these are the closed shift-invariant subsets of the Ω_n with the shift acting on them. Topological Markoff chains are subshifts of what is called finite type. A topological Markoff chain can, e.g., be described as the shift acting on a set M that is given by a transition matrix $\eta(m, l) \in \{0, 1\}$, $1 \leqq m, l \leqq n$, where $M = \bigcap_{i\in\mathbf{Z}} \{x \in \Omega_n : \eta(x_i, x_{i+1}) = 1\}$. One can extend the finite generator theorem from the n-shift to aperiodic topological Markoff chains [28]: For all aperiodic topological Markoff chains M and for all ergodic measure preserving transformations T such that $h(M) > h(T)$ there exists a shift-invariant probability measure ν on M such that the systems (X, μ, T) and (M, ν, S) are isomorphic in the sense of measure theory. In this context one can consider minimal expansive homeomorphisms instead of ergodic measure preserving transformations. Recall that a homeomorphism is called minimal if all of its orbits are dense. One has [28]: If C is a minimal subshift and if M is a topological Markoff chain such that $h(C) < h(M)$ then C is topologically conjugate to the shift acting on a closed invariant subset of M.

Consider the n-shift (Ω_n, S) and let $\mathscr{M}$ be the set of ergodic S-invariant probability Borel measures on Ω_n with support Ω_n. Denote by $\mathscr{H}$ the group of homeomorphisms of Ω_n that commute with S. $\mathscr{H}$ acts on $\mathscr{M}$ by $\mu \to U\mu$ $(\mu \in \mathscr{M})$, $U \in \mathscr{H}$. The following theorem was proved by A. Kuntz [30]: Let $\mu, \nu \in \mathscr{M}$, $h(\nu) \geqq h(\mu)$, and let $I \in \mathbf{N}$, $\varepsilon > 0$. Then there exists a $U \in \mathscr{H}$ such that for all cylinder sets $Z(a) = \{x \in \Omega_n : x_i = a_i,\ 1 \leqq i \leqq I\}$, $a \in \{1, \cdots, n\}^I$, one has $|\mu(Z(a)) - U\nu(Z(a))| < \varepsilon$, $a \in \{1, \cdots, n\}^I$. This is also a statement about generators since every ordered generator of size n yields a shift-invariant measure on Ω_n. An approximation result of this kind was first obtained in [24, §3]. The present proofs of the isomorphism theorem for Bernoulli systems [32], [35], [36] apply in their initial stage such an approximation (for partitions, not necessarily for generators). In their later stage they use an approximation with respect to a different topology that is given by the $\bar{d}$-metric.

A homeomorphism is said to be strictly ergodic if it is minimal and if it has a unique invariant probability measure. Representing an ergodic measure preserving transformation T as a strictly ergodic homeomorphism amounts to finding a generator of T with the appropriate properties. It was proved by R. Jewett that every weakly mixing measure preserving transformation has a representation as a strictly ergodic homeomorphism of the Cantor discontinuum [16]. The possibility of such a representation for all ergodic measure preserving transformations was shown in [26]. This proof used the finite generator theorem. A proof that did not use the finite generator theorem was then given by G. Hansel and J. P. Raoult [14].

For the case at an ergodic measure preserving transformation T with finite entropy $h(T)$ one has that T can be represented as a strictly ergodic expansive homemorphism of the Cantor discontinuum, more precisely, as a strictly ergodic subshift in a shift space over $[e^h + 1]$ symbols [26]. Denker has given a simpler proof of this [3]. This line of investigation was continued by G. Hansel who considered measure preserving transformations that are not necessarily ergodic [13]. He proved that every measure preserving transformation can be imbedded in a homeomorphism all of whose orbit closures are strictly ergodic.

There exist minimal homeomorphisms that are not strictly ergodic (see [34]). Indeed, E. Effros and F. Hahn [9] have constructed a minimal homeomorphism with more than countably many ergodic invariant probability measures. Their example is distal and hence has entropy zero [19]. I. P. Kornfel'd [21] has shown that one can have this situation also with positive entropy, and T. N. T. Goodman [11] has produced such examples in the shift space. These examples suggest the possibility of imbedding (not necessarily ergodic) measure preserving transformations into minimal homeomorphisms. Ch. Grillenberger has recently shown that such an imbedding is always possible.

Attempts have been made to develop this theory for groups other than Z [2], [18], [28]. Let us consider the case of a countably infinite group $\mathcal{G}$ that acts ergodically and freely on $(X, \mathcal{B}, \mu)$ by measure preserving transformations T_g, $g \in \mathcal{G}$. The action is called free if $\mu\{x \in X: T_g x = x\} = 0$ for all $g \in \mathcal{G}$ that are not equal to the unit e of $\mathcal{G}$. Call a sequence $\mathcal{F}(k) \subset \mathcal{G}$, $k \in \boldsymbol{N}$, of finite sets a summing sequence if

$$\lim_{k\to\infty} |\mathcal{F}(k)|^{-1} |\mathcal{F}(k) \Delta g \mathcal{F}(k)| = 0, \qquad g \in \mathcal{G}.$$

Using a summing sequence $\mathcal{F}(k), k \in \boldsymbol{N}$, one attempts for a partition $\mathcal{P}$ the definition

$$h(\mathcal{P}, \mathcal{G}) = \lim_{k\to\infty} |\mathcal{F}(k)|^{-1} h\Big(\bigvee_{g\in\mathcal{F}(k)} T_g \mathcal{P}\Big)$$

and one sets then $h(\mathcal{G}) = \sup h(\mathcal{P}, \mathcal{G})$ where the supremum is taken over all finite partitions. We know that we obtain in this way an entropy theory with the familiar features including, e.g., a finite generator theorem, for a class of groups (e.g., for solvable groups), provided that we can prove an analog of the tower theorem [27]. The tower theorem asserts for an ergodic measure preserving transformation T that for all $I \in \boldsymbol{N}$ and $\varepsilon > 0$ there exists an $F \subset X$ such that $I\mu(F) > 1 - \varepsilon$ and $F \cap T^i F = \emptyset$, $0 < i < I$. An analog statement for the action T_g, $g \in \mathcal{G}$, would be as follows: For some summing sequence $\mathcal{F}(k)$, each $\mathcal{F}(k)$ containing e, one can find for all $k \in \boldsymbol{N}$ and all $\varepsilon > 0$ an $F \subset X$ such that $k\mu(F) > 1 - \varepsilon$ and $F \cap T_g F = \emptyset$, $g \in \mathcal{F}(k)$, $g \neq e$. In most cases proving such a statement is equivalent to proving the hyperfiniteness of $\{T_g: g \in \mathcal{G}\}$. Let us recall here the notion of hyperfiniteness that is due to H. A. Dye [7]: The action $\{T_g: g \in \mathcal{G}\}$ is called hyperfinite if there exists a measure preserving transformation T whose orbits are a.e. the same as the orbits of the action:

$$\{T_g x: g \in \mathcal{G}\} = \{T^i x: i \in \boldsymbol{Z}\} \qquad \text{f.a.a. } x \in X.$$

There are results available on hyperfiniteness in the papers of H. A. Dye [7], [8].

E.g., abelian groups are hyperfinite. On grounds of these results we have the tower theorem for a significant class of groups, and hence we have entropy theory for a significant class of groups. The opinion that hyperfiniteness would enter into the entropy theory of groups was expressed by A. Stepin [37] and A. Veršik, who has announced further results on hyperfiniteness [38].

Investigations on generators for one-parameter flows of measure preserving transformations have also been carried out, e.g., the representation of an ergodic one-parameter flow of measure preserving transformations as a strictly ergodic flow was achieved by K. Jacobs [15] in the weakly mixing case and by M. Denker and E. Eberlein [6] in the general case. For more information on the results for flows we refer to U. Krengel's survey [23].

The situation changes completely if one drops the assumption that a finite measure is preserved. As a matter of fact, as was shown by A. Kuntz [29], if one has a group $\mathcal{G}$ of nonsingular transformations acting on $(X, \mathcal{B}, \mu)$ such that no probability measure that is absolutely continuous with respect to μ is preserved by $\mathcal{G}$, then there exists a set $A \subset X$ such that $\{TA: T \in \mathcal{G}\}$ is dense in $\mathcal{B}$. For previous results on single nonsingular transformations see [22] and [17]. For more information on this topic we refer again to U. Krengel's survey [23].

References

1. R. L. Adler, A. G. Konheim and M. H. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319. MR **30** #5291.

2. J. P. Conze, *Entropie d'un groupe abélien de transformations*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **25** (1972), 11–30.

3. M. Denker, *On strict ergodicity*, Math. Z. **134** (1973), 231–253.

4. ———, *Finite generators for ergodic, measure preserving transformations*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **29** (1974), 45–55.

5. ———, *Einige Bemerkungen zu Erzeugersätzen*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **29** (1974), 57–64.

6. M. Denker and E. Eberlein, *Ergodic flows are strictly ergodic*, Advances in Math. **13** (1974), 437–473.

7. H. A. Dye, *On groups of measure preserving transformations*. I, Amer. J. Math. **81** (1959), 119–159. MR **24** #A1366.

8. ———, *On groups of measure preserving transformations*. II, Amer. J. Math. **85** (1963), 551–576. MR **28** #1275.

9. E. G. Effros and F. Hahn, *Locally compact transformation groups and C*-algebras*, Mem. Amer. Math. Soc. No. 75 (1967). MR **37** #2895.

10. T. N. T. Goodman, *Relating topological entropy and measure entropy*, Bull. London Math. Soc. **3** (1971), 176–180. MR **44** #6934.

11. ———, *Maximal measures for expansive homeomorphisms*, J. London Math. Soc. (2) **5** (1972), 439–444. MR **47** #2031.

12. L. W. Goodwyn, *Topological entropy bounds measure-theoretic entropy*, Proc. Amer. Math. Soc. **23** (1969), 679–688. MR **40** #299.

13. G. Hansel, *Strict uniformity in ergodic theory*, Math. Z. **135** (1974), 221–248.

14. G. Hansel and J. P. Raoult, *Ergodicity, uniformity and unique ergodicity*, Indiana Univ. Math. J. **23** (1973/74), 221–237. MR **47** #6996.

15. K. Jacobs, *Lipschitz functions and the prevalence of strict ergodicity for continuous-time flows*, Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State Univ., Columbus,

Ohio, 1970), Lecture Notes in Math., vol. 160, Springer-Verlag, Berlin and New York, 1970, pp. 87–124. MR 43 #471.

16. R. I. Jewett, *The prevalence of uniquely ergodic systems*, J. Math. Mech. 19 (1969/70), 717–729. MR 40 #5824.

17. L. K. Jones and U. Krengel, *On transformations without finite invariant measure*, Advances in Math. 12 (1974), 275–295.

18. Y. Katznelson and B. Weiss, *Commuting measure-preserving transformations*, Israel J. Math. 12 (1972), 161–173.

19. H. B. Keynes, *Lifting of topological entropy*, Proc. Amer. Math. Soc. 24 (1970), 440–445. MR 43 #472.

20. H. B. Keynes and J. B. Robertson, *Generators for topological entropy and expansiveness*, Math. Systems Theory 3 (1969), 51–59. MR 40 #300.

21. I. P. Kornfel'd, *Invariant measures of minimal dynamical systems*, Dokl. Akad. Nauk SSSR 202 (1972), 280–283 = Soviet Math. Dokl. 13 (1972), 87–90. MR 46 #5576.

22. U. Krengel, *Transformations without finite invariant measure have finite strong generators*, Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State Univ., Columbus, Ohio, 1970), Lecture Notes in Math., vol. 160, Springer-Verlag, Berlin and New York, 1970, pp. 133–157. MR 42 #4703.

23. ———, *Recent results on generators in ergodic theory*, Trans. Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Academia, Prague, 1973.

24. W. Krieger, *On entropy and generators of measure-preserving transformations*, Trans. Amer. Math. Soc. 149 (1970), 354–464; Erratum 168 (1972), 519. MR 41 #3710; 45 #3673.

25. ———, *On generators in exhaustive σ-algebras of ergodic measure-preserving transformations*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 20 (1971), 75–82.

26. ———, *On unique ergodicity*, Proc. Sixth Berkeley Sympos. on Math. Statist. and Probability, vol. II, Univ. of California Press, Berkeley and Los Angeles, Calif., 1972, pp. 327–346.

27. ———, *On the entropy of groups of measure preserving transformations*, Ann. Probability (to appear).

28. ———, (to appear).

29. A. J. Kuntz, *Groups of transformations without finite invariant measures have strong generators of size* 2, Ann. Probability 2 (1974), 143–146.

30. ———, *Homeomorphisms and shift invariant ergodic measures on product spaces* (to appear).

31. O. E. Lanford and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. 13 (1969), 194–215. MR 41 #1343.

32. D. Ornstein, *Ergodic theory, randomness and dynamical systems*, Yale Univ. Press, New Haven, Conn., 1974.

33. D. Ornstein and B. Weiss, *Every transformation is bilaterally deterministic*, Israel J. Math. (to appear).

34. J. C. Oxtoby, *Ergodic sets*, Bull. Amer. Math. Soc. 58 (1952), 116–136. MR 13, #850.

35. P. Shields, *The theory of Bernoulli shifts*, Univ. of Chicago Press, Chicago and London, 1973.

36. M. Smorodinsky, *Ergodic theory, entropy*, Lecture Notes in Math., vol. 214, Springer-Verlag, Berlin and New York, 1971.

37. A. M. Stepin, *The entropy invariant of decreasing sequences of measurable partitions*, Funkcional. Anal. i Priložen. 5 (1971), no. 3, 80–84 = Functional Anal. Appl. 5 (1971), 237–240. MR 44 #1793.

38. A. M. Veršik, *Nonmeasurable decompositions, orbit theory, algebras of operators*, Dokl. Akad. Nauk SSSR 199 (1971), 1004–1007 = Soviet Math. Dokl. 12 (1971), 1218–1222. MR 44 #4537.

INSTITUT FÜR ANGEWANDTE MATHEMATIK DER UNIVERSITÄT HEIDELBERG
HEIDELBERG, FEDERAL REPUBLIC OF GERMANY

Труды Международного Конгресса Математиков
Ванкувер, 1974

О Поведении Гамильтоновых Систем, Близких к Интегрируемым

Н. Н. Нехорошев

Доклад посвящен исследованию поведения переменных действия I в системе канонических уравнений Гамильтона

$$\dot{I} = -\partial H/\partial\varphi, \qquad \dot{\varphi} = \partial H/\partial I,$$

с гамильтонианом H

$$H = H_0(I) + \varepsilon H_1(I, \varphi), \qquad \varepsilon \ll 1, \tag{1}$$

в течение большого по сравнению с $1/\varepsilon$ в любой степени отрезка времени; здесь возмущение εH_1—периода 2π по угловым переменным $\varphi_1, \cdots, \varphi_n$, а $I = I_1, \cdots, I_n$—n-мерный вектор.

1. Колмогоровские торы и диффузия Арнольда. А.Н. Колмогоров, В.И. Арнольд и Ю. Мозер показали (см. [1], [2], [3]), что если невозмущенный гамильтониан H_0—функция "общего положения", то в фазовом пространстве системы с гамильтонианом (1) существует массивное множество, состоящее из n-мерных инвариантных торов, близких к торам, выделяемым уравнениями $I = \text{const}$. Массивное в том смысле, что мера дополнения к нему мала вместе с ε. Это множество, называемое колмогоровским, замкнуто и нигде не плотно.

Если число степеней свободы n равно 2, то 2-мерные колмогоровские торы "разделяют" 3-мерные поверхности уровня функции Гамильтона H. Поэтому при всех начальных условиях $I(0)$, $\varphi(0)$ решения $I(t)$, $\varphi(t)$ будут такими, что точка $I(t)$ будет вечно устойчивой, то есть величина $\sup_{t\in R} \|I(t) - I(0)\|$ мала вместе с ε. Если $n > 2$, то, как показывают примеры, построенные В.И. Арнольдом в [4], в дополнительном к колмогоровскому множестве точка $I(t)$ может уходить, хотя и очень медленно, от начального положения.

2. Экспоненциальная оценка. В сформулированной в п. 4 теореме оценивается снизу "время удержания" точки $I(t)$ вблизи $I(0)$ сразу для всех начальных условий $I(0)$, $\varphi(0)$. Оказывается, что если функция H_0 удовлетворяет некоторым условиям, названным условиями крутизны, то за большой промежуток времени порядка $\exp(1/\varepsilon^a)$ точка $I(t)$ не уйдет от $I(0)$ дальше, чем на малое расстояние ε^b, где $0 < a < 1$, $0 < b < 1$.

3. Переменные действия—адиабатические инварианты. Экспоненциальная оценка справедлива и для таких систем, у которых возмущение εH_1 зависит также и от медленного времени εt, то есть для систем с гамильтонианом

$$H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon t).$$

Таким образом, переменные I_i являются адиабатическими инвариантами. Из этого факта, в частности, следует устойчивость в течение экспоненциально большого промежутка времени планетной системы, то есть системы типа нашей Солнечной системы.

4. Точная формулировка экспоненциальной оценки для системы (1). Пусть функция $H = H_0(I) + \varepsilon H_1(I, \varphi)$ аналитична в комплексной области F вида

$$F: \operatorname{Re} I \in G, \quad |\operatorname{Im} I| < \rho, \quad |\operatorname{Im} \varphi| < \rho, \quad \text{где } \rho > 0, \text{ а } G \subset \boldsymbol{E}^n,$$

$\mathbf{E}^n$—евклидово пространство. Будем предполагать, что

$$\inf_{I \in G} \|\operatorname{grad} H_0|_I\| = 1, \qquad \sup_{I, \varphi \in F} |H_1(I, \varphi)| = 1, \qquad \text{а} \sup_{I, \varphi \in F} \left\| \left(\frac{\partial^2 H_0(I)}{\partial I_i \partial I_j} \right)\Big|_I \right\| < \infty,$$

где $\|(\partial^2 H_0(I)/\partial I_i \partial I_j)|_I\|$—норма линейного оператора в $\boldsymbol{C}^n$, матрица которого имеет вид $(\partial^2 H_0(I)/\partial I_i \partial I_j)$.

Теорема [5]. *Пусть функции H_0 и H_1 удовлетворяют описанным выше условиям и пусть функция H_0 крутая в области G (определение крутизны дано в следующем пункте). Тогда существует константа $\varepsilon_0 = \varepsilon_0(H_0, \rho) > 0$, обладающая следующим свойством. Обозначим через $G - d$ множество точек, содержащихся в G вместе со своей d-окрестностью. Тогда если $0 < \varepsilon < \varepsilon_0$, то для любого решения $I(t)$, $\varphi(t)$ системы (1), такого что $I(0) \in G - d$*

$$\|I(t) - I(0)\| < d \quad \textit{при всех } t \in [0, T],$$

где

$$d = \varepsilon^b, \qquad T = \exp(1/\varepsilon^a). \tag{2}$$

Константы $a > 0$ и $b > 0$, зависящие лишь от H_0, будут определены ниже.

5. Условия крутизны. Пусть λ—произвольное (аффинное) подпространство пространства $\boldsymbol{E}^n$, содержащее точку $I' \in G$, $\dim \lambda \neq 0$. Обозначим градиент сужения H_0 на плоскость λ через $\operatorname{grad}(H_0|_\lambda)$. Обозначим минимальное значение длины вектора $\operatorname{grad}(H_0|_\lambda)$ на сфере радиуса η с центром в I' через $m_{I', \lambda}(\eta)$,

$$m_{I',\lambda}(\eta) = \min_{\{I \in \lambda : \|I - I'\| = \eta\}} \|\operatorname{grad}(H_0|_\lambda)|_I\|.$$

Определение. Функцию H_0 назовем крутой в точке I на плоскости λ, если найдутся числа $C > 0$, $\delta > 0$ и $\alpha \geqq 1$, такие что

$$\max_{0 \leqq \eta \leqq \xi} m_{I,\lambda}(\eta) > C\xi^\alpha$$

при всех $\xi \in (0, \delta]$. Константы C и δ назовем коэффициентами, а α—показателем крутизны в точке I на λ.

Обозначим через $\Lambda^r(I)$ множество всех r-мерных плоскостей, проходящих через точку $I \in E^n$.

Определение. Функцию H_0 назовем крутой в точке I с коэффициентами C_r и δ_r и показателями α_r, $r = 1, \cdots, n-1$, если, во-первых, $\operatorname{grad} H_0|_I \neq 0$, во-вторых, для каждого $r = 1, \cdots, n-1$, функция H_0 будет крутой в точке I на каждой плоскости $\lambda \in \Lambda^r(I)$ перпендикулярной вектору $\operatorname{grad} H_0|_I$, с коэффициентами C_r и δ_r и показателем α_r.

Функцию H_0 назовем крутой в области G с коэффициентами C_r и δ_r и показателями α_r, $r = 1, \cdots, n-1$, если она крутая в каждой точке $I \in G$ с теми же коэффициентами и показателями.

6. Показатели крутизны и степень "устойчивости" системы. Константы a и b, определяющие оценки (2) Теоремы 4, являются функциями показателей крутизны $\alpha_1, \cdots, \alpha_{n-1}$:

$$a = 2/(12\zeta + 3n + 14), \qquad b = a/\alpha_{n-1},$$

где $\zeta = [(\alpha_1(\alpha_2 \cdots (n\alpha_{n-2} + n - 2) + \cdots + 2) + 1)] - 1$ при $n > 2$ и $\zeta = 1$ при $n = 2$.

Для константы a можно получить и лучшее значение. Но вид ее зависимости от показателей крутизны указывает на то, что чем больше эти показатели, то есть чем менее крутая функция H_0, тем меньше время удержания точек $I(t)$ вблизи начального положения.

7. Бесконечновырожденность некрутых функций. Рассмотрим пространство J_r r-струй функций в произвольной точке I, то есть пространство младших коэффициентов Тейлора до порядка r включительно разложения функций в этой точке. В работе [6] доказано, что для каждого r в J_r существует полуалгебраическое множество Σ_r, обладающее следующими свойствами. Во-первых, каждый представитель H_0 каждой струи, лежащей вне Σ_r, либо является крутой функцией в некоторой окрестности точки I, либо $\operatorname{grad} H_0|_I = 0$. Во-вторых, коразмерность Σ_r в J_r стремится в бесконечности при $r \to \infty$.

Таким образом, некрутые в окрестности некритической точки функции бесконечновырожденные: коэффициенты Тейлора разложения этих функций в такой точке удовлетворяют бесконечному числу независимых полиномиальных уравнений.

8. Алгебраические критерии крутизны. Для каждого r явно выписываются системы полиномиальных уравнений относительно коэффициентов Тейлора и

некоторых других переменных (параметров) такие, что если исключить параметры, то получатся полуалгебраические условия, задающие Σ_r. Тем самым в некотором смысле решается вопрос об эффективной проверке условий крутизны.

9. Квазивыпуклые функции. Примерами крутых функций являются функции H_0, обладающие следующим свойством: для каждой точки I из ее области определения $\operatorname{grad} H_0\big|_I \neq 0$ и система

$$\sum_{i=1}^{n} \frac{\partial H_0(I)}{\partial I_i}\eta_i = 0; \qquad \sum_{i,j=1}^{n} \frac{\partial^2 H_0(I)}{\partial I_i \partial I_j}\eta_i\eta_j = 0$$

не имеет решений $\eta = \eta_1, \cdots, \eta_n$, кроме тривиального $\eta = 0$. Такие функции будем называть квазивыпуклыми.

Рассмотрим определитель симметричной матрицы порядка $n + 1$

$$\Delta = \det\begin{pmatrix} \partial^2 H_0/\partial I^2 & \partial H_0/\partial I \\ \partial H_0/\partial I & 0 \end{pmatrix}.$$

Оказывается, что при $n = 2$ квазивыпуклость эквивалентна не обращению в ноль Δ, а при $n = 3$ эквивалентна условию, что при всех $I \in G$, Δ меньше нуля.

Для квазивыпуклых функций и только для них все показатели крутизны α_r минимальны, то есть равны единице. Если справедлива гипотеза из пункта 6, то при $n = 3$ системы, для которых Δ меньше нуля, должны быть гораздо устойчивее, чем те, для которых Δ не меньше нуля.

10. О доказательстве экспоненциальной оценки. Доказательство опирается на технику, развитую в [2]. Общий ход доказательства близок к доказательству формальной устойчивости, проведенному Дж. Глиммом в [7]. Отметим, что и условия крутизны близки к условиям, наложенным Глиммом на H_0, хотя он и не выяснил, удовлетворяют ли функции общего вида его условиям.

11. Почти интегралы. Основная идея доказательства состоит в том, что область G, $I \in G \subset \boldsymbol{E}^n$ пронизана почти интегралами. То есть существуют функции, определенные на $G \times T^n$, где $\varphi \in T^n$, каждая из которых, во-первых, почти не отличается от функции, зависящей только от I (эту функцию и будем называть почти интегралом), во-вторых, экспоненциально медленно меняется на решениях системы (1). Почти интегралы линейно зависят от переменных I.

Как число линейно независимых почти интегралов, так и расположение в $\boldsymbol{E}^n$ пересечений их поверхностей уровня, вообще говоря, меняются от точки к точке. В частности, существуют области, в которых есть полный набор, то есть n линейно независимых почти интегралов. Следовательно, в этих областях средняя скорость точки $I(t)$ будет экспоненциально малой, здесь $I(t)$, $\varphi(t)$—произвольное решение.

12. Ловушки. Там, где число почти интегралов меньше n, точка $I(t)$ может с "большой", то есть порядка ε, скоростью двигаться вдоль пересечений

поверхностей уровня этих почти интегралов. Но если функция H_0—крутая, то и в этом случае почти интегралы из соседних областей "запирают" точку $I(t)$ в множестве малого вместе с ε диаметра: точка $I(t)$ как бы находится в ловушке малых размеров, по которой может бегать с большой скоростью.

Литература

1. А.Н. Колмогоров, Доклады АН СССР **98** (1954).
2. В.И. Арнольд, Успехи матем. наук **18** (1963).
3. J. Moser, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II (1962).
4. В.И. Арнольд, Доклады АН СССР **156** (1964).
5. Н.Н. Нехорошев, Функц. анализ и его приложения **5** (1971).
6. ——— Мат. сборник **90** (132) (1972).
7. J. Glimm, Comm. Pure Appl. Math. **17** (1964).

Московский Государственный Университет
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

On Bifurcations of Dynamical Systems

M. M. Peixoto

1. Introduction. We survey here a certain line of developments of bifurcation theory of dynamical systems on a compact differentiable manifold M^n.

The word bifurcation was introduced by Poincaré in connection with differential equations of celestial mechanics depending on a real parameter λ. He focused on a certain feature of the equation, say an isolated closed orbit, present for $\lambda = \lambda_0$. If when λ varies past λ_0 this closed orbit disappears or undergoes a sudden topological change, λ_0 is said to be a bifurcation value; otherwise it is an ordinary value of the parameter. Traditionally bifurcation theory has been developed from this local point of view.

After the emergence of the generic theory of dynamical systems via the coming of age of the concept of structural stability it became natural to take a more general and global approach to bifurcation theory, as follows. Let $\mathfrak{X}$ be the space of all C^r-flows (diffeomorphisms) on M^n with the C^r-topology, $1 < r < \infty$, and $\Sigma \subset \mathfrak{X}$ be the structurally stable ones. Bifurcation theory is then the study of maps of a compact manifold Λ^k (parameter space) into M^n or more precisely the study of the intersections of the image of Λ^k with the connected components of Σ.

The first to adopt this point of view was J. Sotomayor [15], [16] who considered for flows the case where $n = 2$ and $k = 1$. He describes how, generically, an arc in $\mathfrak{X}$ intersects Σ. Later on he generalized part of this work for $n > 2$ but keeping $k = 1$ [18], this generalization being some kind of analog of the Kupka-Smale theorem.

A natural approach to the much more difficult case $n > 2$ is to substitute for Σ the best known open subset of Σ, namely the subset Δ of all Morse-Smale diffeomorphisms (flows) on M^n.

This approach was adopted by Newhouse and Palis [8], [9] who give a very

delicate analysis indicating what happens, generically, when an arc chosen within a certain open set of arcs and starting at an interior point $X \in \Delta$ meets the boundary $\partial\Delta$ for the first time.

Another natural approach making even more drastic assumptions on the data in order to get some simplicity in high dimension bifurcation is the simple arc approach of Newhouse and the author. Here one starts with the endpoints X, Y on Δ (or Σ) and one tries to find an arc γ on $\mathfrak{X}$ connecting X and Y and having only a finite number of "simple" bifurcations and this in a stable way.

On the historical side it should be mentioned here the important role played in bifurcation theory by the paper of Andronov and Leontovich [1] where they characterize first order structural stability (i.e., structural stability in $\mathfrak{X} - \Sigma$) on the two-dimensional disc. It is a role similar to the one played in dynamical system theory by the original paper of Andronov and Pontrjagin [2] on structural stability. The fact that the important concepts developed in these papers remained dormant for decades can probably be explained by the circumstance that their authors refrained from making a metric space out of the totality of all differential equations that they considered.

We now pass to describe in more detail the three approaches to bifurcation theory mentioned above.

2. The generic approach of Sotomayor. Consider first flows on M^2, $r \geqq 4$, $\mathfrak{X}$ being a Banach space.

(2.1) THEOREM. *There exists a dense subset Σ_1 of $\mathfrak{X} - \Sigma$ which is a codimension* 1 *immersed submanifold of class C^{r-1} of $\mathfrak{X}$ and such that any $X \in \Sigma_1$ has a neighborhood in the intrinsic topology of Σ_1 made up of flows topologically equivalent to X.*

In $\mathfrak{X} - \Sigma$ there is then a codimension 1 differentiable part. Whether or not it is maximal remains an enticing but very difficult open problem. Σ_1 was constructed so as to contain the flows structurally stable of the first order. The subset $\tilde{\Sigma}_1$ corresponding to them constitutes an imbedded submanifold of $\mathfrak{X}$. Along these lines, a program for higher order structural stability was suggested in [17].

Now use Σ_1 to get information about the generic arc in $\mathfrak{X}$. Let Φ be the Banach manifold of all C^1-maps of $J = [0, 1]$ into $\mathfrak{X}$ and $\mathcal{W} \subset \mathfrak{X}$ be the set of all Kupka-Smale flows.

(2.2) THEOREM. *There exists a Baire set $\Gamma \subset \Phi$ of maps such that $\gamma \in \Gamma$ implies: γ is transversal to Σ_1, $\gamma(J) \subset \mathcal{W} \cup \Sigma_1$ and besides the set of ordinary values of γ is open and dense in J and coincides with $\gamma^{-1}(\Sigma)$.*

The above situation suggests the consideration of structural stability of arcs in $\mathfrak{X}$, i.e., of elements of Φ and in particular the question: Are they dense in Φ? In [5], J. Guckenheimer constructed, for T^2, an open set of nonstructurally stable elements of Φ and gave a sufficient condition for structural stability of arcs in terms of $\tilde{\Sigma}_1$.

We now pass to $n > 2$ and discuss Sotomayor's analog to the Kupka-Smale theorem.

Define the quasi-Kupka-Smale flows $Q \subset \mathfrak{X}$ [11] as the flow obtained by weakening in the least possible way exactly one of the conditions that characterize the Kupka-Smale flows and maintaining the others. So a flow in Q either presents a single quasi-hyperbolic singular point or closed orbit and all the appropriate stable and unstable manifolds are transversal or else all the singular points and closed orbits are hyperbolic and all stable and unstable manifolds are transversal except exactly one pair of these which are quasi-transversal.

We now explain the meaning of quasi-hyperbolic and quasi-transversal.

A fixed point of a diffeomorphism is said to be quasi-hyperbolic if all the eigenvalues are simple and lie off the unit circle with a single exception which is either 1, -1 or a pair of complex eigenvalues. In each case one has to impose nondegeneracy conditions involving derivatives up to order 4.

Now a singular point of a flow is said to be quasi-hyperbolic if the corresponding time 1 diffeomorphism has this point as a quasi-hyperbolic fixed point. Then the corresponding stable and unstable manifolds are immersed euclidean spaces or immersed half-euclidean spaces so that these are manifolds with boundary. A closed orbit is quasi-hyperbolic if the corresponding Poincaré map is quasi-hyperbolic.

The above picture of quasi-hyperbolicity results from the work of Hopf [**20**], Sacker [**14**], Brunovsky [**3**] and Ruelle and Takens [**13**] in the case of complex eigenvalues; and from that of Sotomayor [**16**], [**18**] in the real eigenvalue case. We say that the differentiable manifolds $V, W \subset M^n$ are quasi-transversal at a point p of their intersection if $T_pV + T_pW$ is an $(n-1)$-dimensional subspace of T_pM^n and besides at p the contact of V and W is "as small as possible". This concept is due to Sotomayor who communicated it to several persons. See [**8**] and [**11**].

We can now state the analog to the Kupka-Smale theorem. Let $Q \subset \mathfrak{X}$ be the set of all quasi-Kupka-Smale flows and, as before, Φ the space of all C^1-arcs in $\mathfrak{X}$.

(2.3) THEOREM. (a) *The set Q is the union of an expanding sequence of embedded codimension one submanifolds Q_i, $i = 1,2, \cdots$, of $\mathfrak{X}$.*

(b) *There is a Baire set Γ in Φ such that every arc $\gamma \in \Gamma$ is such that $\gamma(t)$ is Kupka-Smale for all $t \in J = [0, 1]$ with the eventual exception of a countable set in J and besides γ is transversal to Q.*

In [**18**] there is a version of this theorem which is weaker because it does not deal with quasi-transversal intersections but on the other hand gives further information about stratifications associated to the stable and unstable manifolds.

3. The approach of Newhouse and Palis. Since the set Δ of Morse-Smale diffeomorphisms constitutes the best known open set of the set of structurally stable diffeomorphisms, it is natural to use Δ to get a hold on more complicated diffeomorphisms and more precisely the ones that lie on the boundary $\partial\Delta$ and beyond. Along this line Newhouse and Palis [**8**], [**9**] consider an arc $\xi(t)$, $0 \leqq t \leqq 1$, such that $\xi(0) \in \Delta$ and call b the first value of t for which $\xi(t) \notin \Delta$. Their problem is then to describe, for a Baire set of the family Φ of all arcs ξ with $\xi(0) \in \Delta$ and

$\xi(b) \in \partial\Delta$, the phase portrait of $\xi(b)$ and that of $\xi(t)$ for $b < t < t + \varepsilon$ and some $\varepsilon > 0$. In the above papers they give a partial answer to this very difficult problem.

Their results can be summarized briefly as follows. Let $L^-(f)$ and $L^+(f)$ stand respectively for the closure of the set of α-limit points of f and the closure of the ω-limit points of f. Call $L(f) = L^-(f) \cup L^+(f)$.

(3.1) THEOREM. *The set of arcs $\Gamma \subset \Phi$ such that $\xi \in \Gamma$ implies that $L(\xi(b))$ is finite is open.*

The following theorem describes the behaviour of a generic arc in Γ.

(3.2) THEOREM. *There is a Baire set of arcs $\Gamma_1 \subset \Gamma$ such that for $\xi \in \Gamma_1$ we have:*

(a) *If $L^-(\xi(b))$ has no cycles, then, for some $\varepsilon > 0$, $\xi(t)$ is Morse-Smale for t in a dense open set in $[b, b + \varepsilon]$.*

(b) *If $L^-(\xi(b))$ is hyperbolic and has a cycle, then for some $\varepsilon > 0$ there are infinitely many topologically distinct structurally stable $\xi(t)$'s with $L(\xi(t))$ infinite, $b < t < b + \varepsilon$.*

An outstanding conjecture of this theory is whether or not the set of arcs ξ for which either $L^+(\xi(b))$ or $L^-(\xi(b))$ is finite is dense in Φ. Another is whether or not there is a Baire set in Φ such that for an arc ξ on it there is an $\varepsilon > 0$ such that $\xi(t)$ is structurally stable for an open dense set of values of t in $[b, b + \varepsilon]$.

4. The simple arc approach. Analogously to what was done for Kupka-Smale flows on M^n we say that a flow is quasi-Morse-Smale if it differs from a Morse-Smale flow by the fact that either a unique singularity is quasi-hyperbolic or else a unique pair of intersecting stable and unstable manifolds is quasi-transversal. Say that an arc $\gamma = \gamma(t) \subset \mathfrak{X}$ is simple if $\gamma(t) \subset \Delta$ for $t = 0$ and $t = 1$ and all other values of t except a finite number of them for which the corresponding flow is quasi-Morse-Smale and this in a stable way, i.e., any neighboring arc has the same number of exceptional points. In [**10**] Newhouse and the author proved the following

(4.1) THEOREM. *There is a simple arc joining any two Morse-Smale flows.*

When $M^n = S^2$ this theorem is an immediate consequence of an argument used in [**12**] for another purpose, and in this simple case we have an actual construction of the arc in which the exceptional points correspond to flows devoid of saddle connections, i.e., of quasi-transversal intersections.

To find a constructive proof of (4.1) is an open problem which seems to be far off, involving too much knowledge of the topology of M^n. Even when $n = 2$ this is open but M. Barone (unpublished) found such a proof for T^2. Another interesting open problem is whether or not (4.1) can be strengthened so that the exceptional points exhibit no quasi-transversal intersections.

There are some interesting relations between (4.1) and other topics. Fleitas [**4**] indicated that (4.1) for gradient-like flows on M^3 is essentially equivalent to an old theorem by Singer [**19**] about "moves" on Heegard diagrams. Also Fleitas remarked that the proof of the h-cobordism theorem in [**6**] contains as a special case the construction of a simple arc between any two gradient flows on S^n, $n \geqq 6$. Simple

arcs are then a nice differentiable way to express some surgeries used in topology.

Simple examples show that there is no analog of (4.1) for diffeomorphisms.

We finally remark that in [7] Newhouse extended (4.1) for flows substituting for Δ the wider class of flows satisfying Axiom A and the strong transversality condition, in case $n < 4$. If $n \geqq 4$ the extension still goes but a further condition related to the nonwandering set has to be assumed.

Bibliography

1. A. A. Andronov and E. A. Leontovic, *Sur la théorie de la variation de la structure qualitative de la division du plan en trajectoires*, Dokl. Akad. Nauk. **21** (1938), 427–430.

2. A. A. Andronov and L. S. Pontrjagin, *Systèmes grossiers*, Dokl. Akad. Nauk **14** (1937), 247–251.

3. P. Brunovksý, *On one-parameter families of diffeomorphisms*. I, II, Comment. Math. Univ. Carolinae **11** (1970), 559–582; ibid. **12** (1971), 765–784. MR **43** #5548; **46** #8241.

4. G. Fleitas, *On the classification of flows and manifolds in dimension two and three*, Thesis, IMPA, 1972.

5. J. Guckenheimer, *One parameter families of vector fields on two-manifolds: another non-density theorem*, Dynamical Systems (Proc Sympos. Univ. of Bahia, Salvador, Brazil, 1971), Academic Press, New York, 1973, p. 111. MR **48** #1255.

6. J. Milnor, *Lectures on the h-cobordism theorem*, Princeton Univ. Press, Princeton, N. J., 1965. MR **32** #8352.

7. S. Newhouse, *On simple arcs between structurally stable flows* (to appear).

8. S. Newhouse and J. Palis, *Bifurcation of Morse-Smale dynamical systems*, Dynamical Systems (Proc. Sympos Univ. of Bahia, Salvador, Brazil, 1971), Academic Press, New York, 1973, p. 303.

9. ———, *Cycles and bifurcation theory* (to appear).

10. S. Newhouse and M. M. Peixoto, *There is a simple arc joining any two Morse-Smale flows* (to appear).

11. J. Palis, *Arcs of dynamical systems: bifurcations and stability* (to appear).

12. M. M. Peixoto, *On the classification of flows on 2-manifolds*, Dynamical Systems (Proc. Sympos. Univ. of Bahia, Salvador, Brazil, 1971), Academic Press, New York, 1973, p. 303.

13. D. Ruelle and F. Takens, *On the nature of turbulence*, Comm. Math. Phys. **20** (1971), 167.

14. R. J. Sacker, *Invariant surfaces and bifurcation of periodic solutions of ordinary differential equations*, Report IMM-NYU, 333, New York, 1964.

15. J. Sotomayor, *Estabilidade estrutural de primeira ordem e variedades de Banach*, Thesis, IMPA, 1964.

16. ———, *Generic one-parameter families of vector fields on two-dimensional manifolds*, Inst. Hautes Études Sci. Publ. Math. No. 43 (1974).

17. ———, *Structural stability and bifurcation theory*, Dynamical Systems (Proc. Sympos. Univ. of Bahia, Salvador, Brazil, 1971), Academic Press, New York, 1973, p. 549.

18. ———, *Generic bifurcations of dynamical systems*, Dynamical Systems (Proc. Sympos. Univ. of Bahia, Salvador, Brazil, 1971), Academic Press, New York, 1973, p. 561.

19. J. Singer, *Three-dimensional manifolds and their Heegard diagrams*, Trans. Amer. Math. Soc. **35** (1933), 88.

20. E. Hopf, *Abzweigung einer periodischen Lösung einer stätionären Lösung eines Differentialsystems*, Ber. Math.-Phys. Kl. Sächs. Akad. Weiss. Leipzig **94** (1942), 1.

UNIVERSIDADE DE SÃO PAULO
SÃO PAULO, BRAZIL

IMPA
RIO DE JANEIRO, BRAZIL

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Structure of Bernoulli Systems

Benjamin Weiss

In the wake of D. Ornstein's proof five years ago that Bernoulli shifts with the same entropy are isomorphic came a flood of new results concerned in one way or another with Bernoulli shifts. Perhaps most surprising has been the discovery that so many of the classical examples of ergodic theory, as well as many "physical" systems, are isomorphic to Bernoulli shifts. We shall not discuss this aspect in any detail and merely mention some of the systems that have been shown to be Bernoullian: (a) ergodic automorphisms of T^n [**4**], T^∞ [**7**]; (b) the geodesic flow on surfaces of negative curvature [**13**]; (c) Anosov flows with smooth measures [**16**]; (d) two-dimensional billiards with a convex scatterer [**3**]. Even though many innocent looking questions still remain unanswered, such as [**20**, Problem 7.2], it seems that the time is ripe to focus on other aspects of Bernoulli systems. In particular, having shown that some physical system is Bernoullian, what does that allow one to say about the system itself? To answer such questions one must dig deeper and gain a better understanding of a Bernoulli system, and here the work has in some sense just begun. While some progress has been made in understanding properties that are universally true for *all* partitions of a Bernoulli shift, there are many questions, such as the central limit theorem, whose answers depend upon the particular partition chosen. To concretize this last point let us look at the property of weak Bernoulli (WB) (as general references I shall use [**12**], [**17**] where the terms not defined here, as well as further references, may be found). It is known that not every partition of a Bernoulli shift is WB, and there are some properties that hold for them that do not hold in general (for example the generators α mentioned at the end of §2(a) cannot be WB, since for WB β, $\bigcap_n \bigvee_{|i|>n} \varphi^i\beta$ is trivial). The problem is on the one hand—what else is true specifically for WB partitions and on the other hand—how are we to recognize them? For example if an Anosov flow

is Bernoulli are all smooth partitions WB? (See Note 1 at the end of this paper.)

My article will divide into two parts. In the first I will try to give a general framework for Bernoulli systems. A host of open problems is implicit in the gaps in the picture that will be drawn. In the second part I shall discuss in outline what is known about the deeper structure of Bernoulli systems—and there too, as will be seen, we stand really at the threshold.

1. Generalized Bernoulli systems. Let $(X, \mathscr{B}, \mu)$ be a measure space; for most of the discussion $\mu(X)$ is assumed finite and normalized $\mu(X) = 1$. If G is any group then an *action* of G on X is a representation of G as a group of measure preserving transformations (m.p.t.) $\varphi_g : X \to X$ and will be denoted by $\Phi = \{\varphi_g\}_{g \in G}$. Most of ergodic theory has been developed for $G = Z$ when the action is defined by φ_1, a single invertible m.p.t. We shall also refer to Φ as a *system* or *dynamical system.* A system Φ is said to be *Bernoulli,* in case G is countable, if there is defined on X a random variable (= measurable function) $f_0 : X \to R$ such that the random variables $\{f_g(x) = f(\varphi_g(x))\}_{g \in G}$ are mutually independent and $\mathscr{B}$ is the smallest α-algebra with respect to which all the f_g are measurable. In case f_0 can be chosen to be finite or countably valued the partition α of X into sets of constancy for f_0 is a finite or countable generator, while the independence of the f_g means, of course, that the partitions $\{\varphi_g \alpha\}_{g \in G}$ are independent. By the standard construction of product measure spaces Bernoulli systems always exist in this case, and provide, in some sense, the simplest kind of actions.

The theory of Bernoulli systems for $G = Z^d$, $d \geqq 2$, was developed by several authors [2], [5], [19]. It turns out that the entire Ornstein-isomorphism theory carries over to Z^d-Bernoulli systems. The main examples for Z^d-actions, so far, have been statistical mechanical systems where the action of Z^d arises from a group of symmetries preserving the physical system. Several years ago I conjectured that most such Gibbs states, or Markov random fields, would turn out to be Bernoulli systems, and this indeed has turned out to be the case. I should like to emphasize, however, that it is not at all clear what is the "physical" meaning of such results. It seems that too many diverse systems are isomorphic with the usual concept of measure theoretic isomorphism for this latter notion to be useful. This point will be taken up again in § 2(c).

W. Krieger [6] has further extended the theory to a wide class of countable groups including many noncommutative ones, but the case of nonamenable groups such as the free group on two generators still seems rather intractable. Further fruitful development here will probably have to await the discussion of good examples to which the theory may be applied. The most likely source of such examples is in the statistical mechanics of physical systems which have G as a group of symmetries.

Turning to noncountable groups we encounter quite a different situation. It is natural to suppose here that G is a topological group and that $\varphi_g(x)$ is jointly measurable in (g, x). We cannot require now the existence of a random variable f_0 with $\{f_0(\varphi_g x)\}$ mutually independent since that would force the measure space $(X, \mathscr{B}, \mu)$ to be nonseparable. We will say instead that Φ is a *Bernoulli system* if

for every discrete subgroup $G_0 \subset G$, $\{\varphi_g\}_{g\in G_0}$ is a Bernoulli system. It is now a non-trivial task to establish that Bernoulli systems exist at all. This was done for $G = R^1$ by D. Ornstein [11], who also extended the isomorphism theory to these Bernoulli flows. The key tool required to extend this construction and isomorphism theory to R^d, namely an R^d version of the Kakutani-Rohlin tower theorem, has been established [8], but the isomorphism theory has not yet been carried out in detail. Here too statistical mechanics provides us with natural candidates for B-systems such as infinite systems of hard spheres in R^d. Beyond R^d nothing seems to be known, although it would appear possible to extend the theory at least to something like nilpotent Lie groups. (See Note 2 at the end of this paper.)

2. The fine structure of Bernoulli systems.

(a) A system can be said to be well understood only if one also understands its substructures, and thus one is led to the study of factors of Bernoulli systems. We restrict now to $G = Z^1$ and represent Φ by $\varphi = \varphi_1$. The first result here was that of D. Ornstein [10] who characterized partitions of a B-shift by the property of being finitely determined and thus showed that any factor of a Bernoulli shift was a Bernoulli shift. If one asks what kind of independence is inherited by any partition of a B-shift one is led to the notion of very weakly Bernoulli which we proceed to define. Given two finite sequences of labeled partitions $\{\alpha_i\}_1^n$, $\{\beta_i\}_1^n$ (possibly on different spaces) we define

$$\bar{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) = \inf \frac{1}{n}\sum_1^n d(\bar{\alpha}_i\, \bar{\beta}_i)$$

where the infimum is taken over all partitions $\{\bar{\alpha}_i, \bar{\beta}_i\}_1^n$ defined on a normalized measure space that satisfy

$$\operatorname{dist} \bigvee_1^n \alpha_i = \operatorname{dist} \bigvee_1^n \bar{\alpha}_i, \qquad \operatorname{dist} \bigvee_1^n \beta_i = \operatorname{dist} \bigvee_1^n \bar{\beta}_i$$

and $d(\alpha, \beta) = \sum \mu(A_i \Delta B_i)$ when $\alpha = \{A_1, \cdots, A_k\}$, $\beta = \{B_1, \cdots, B_k\}$.

A process (φ, α) is said to be *very weakly Bernoulli* (VWB) if for any $\varepsilon > 0$ there is an n_0 such that, for all m and $n > n_0$,

$$\bar{d}(\{\varphi^i\alpha\}_1^n, \{\varphi^i\alpha \,|\, B\}_1^n) < \varepsilon$$

holds for a set of atoms $B \in \bigvee_1^m \varphi^{-i}\alpha$ of total measure at least $1 - \varepsilon$. Naturally if one can satisfy these conditions with $\varepsilon = 0$ we have independence, and this is one way of weakening absolute independence. This turns out to be the correct notion in the sense that any process defined on a B-shift is VWB [10], [11], while on the other hand any process that is VWB is isomorphic to a B-shift [14]. In passing we may mention that many other features of independence do not carry over to arbitrary partitions of Bernoulli shifts. For example one might think that any partition α of a B-shift satisfies the following strengthening of the condition for K-automorphisms: $\bigcap_n \bigvee_{|i|>n} \varphi^i\alpha$ is trivial. However it follows from [15] that there are generators α for any Bernoulli shift such that $\bigvee_{|i|>n} \varphi^i\alpha = \bigvee_{-\infty}^{\infty} \varphi^i\alpha$ for all n.

(b) Having established what the factors of a B-shift look like when viewed on

their own we turn to the question of how they are placed or situated. We shall say that two φ-invariant subalgebras $\mathscr{A}_1, \mathscr{A}_2$ of a B-shift $(X, \mathscr{B}, \mu, \varphi)$ are *similarly placed* if there is an invertible measure preserving transformation ψ of X into itself that commutes with φ and maps $\mathscr{A}_1$ onto $\mathscr{A}_2$. Clearly $h(\varphi, \mathscr{A}_1) = h(\varphi, \mathscr{A}_2)$ is a necessary condition, and that it is not sufficient, so that there is some interest in the question, can be seen already from trivial examples such as the following. Let ϑ be an involution that commutes with φ and let $\mathscr{A} = \{A : \vartheta A = A\}$. Then $h(\varphi, \mathscr{A}) = h(\varphi, \mathscr{B})$ but $\mathscr{A}$ and $\mathscr{B}$ are *not* similarly placed, since one can add sets to A without increasing the entropy. J.-P. Thouvenot has initiated in [18] a program of relativizing ergodic theory to deal with such questions. The simplest way that a subalgebra $\mathscr{A}_1$ can be placed is for there to exist an independent algebra $\mathscr{A}_2$, φ-invariant, such that $\mathscr{A}_1 \vee \mathscr{A}_2$ equals $\mathscr{B}$. In that case we say that $\mathscr{A}_1$ has an independent complement. Clearly a necessary condition for $\mathscr{A}_1$ to have an independent complement is that it be maximal in its entropy class, i.e., $\bar{\mathscr{A}}_1 \supset \mathscr{A}_1$ with $h(\varphi, \bar{\mathscr{A}}_1) = h(\varphi, \mathscr{A}_1)$ implies $\mathscr{A}_1 = \bar{\mathscr{A}}_1$. D. Ornstein has constructed an example to show that the condition is not sufficient. (See Note 3 at the end of this paper.) It is likely that his construction can be modified to produce uncountably many "maximal" subalgebras of the same entropy that are pairwise not similarly placed. In some sense it seems that all m.p.t.'s are reflected in the relative structure of factors of B-shifts.

(c) Now that so many different systems have turned out to be Bernoulli one is tempted to refine the equivalence relation and study more restrictive kinds of isomorphisms. We shall mention here only one of the possibilities. The first examples of isomorphisms between Bernoulli systems [9] or Markov shifts [1] were "finitistic" in the sense that the coding from one sequence to another terminated after a finite number of steps, with probability one. To be precise let A_i, $i = 1, 2$, be two finite alphabets and μ_i shift invariant measures on $A_i^Z = X_i$, $i = 1, 2$. A shift invariant mapping $\vartheta : X_1 \to X_2$ that maps μ_1 onto μ_2 is said to be *finitistic* if, for each $a \in A_2$, the set $\vartheta^{-1}\{x_2 : x_2(0) = a\}$ differs from a union of finite cylinder sets by a set of measure zero. Naturally there may be countably many cylinder sets involved, so that $\vartheta_0(x)$ is determined after a finite, but unbounded in general, number of symbols in x have been examined. It is not difficult to give examples of partitions α in B-shifts such that (φ, α) cannot be finitistically coded by an independent process, so that the notion is indeed finer than that of isomorphism. A determination of possible invariants for this equivalence relation would shed much light on the variety of Bernoulli systems.

It seems that about a hundred years after the invention of "ergodic theory" by the physicist Boltzmann the circle is being completed and finally answers are being found to the problems that he raised. To be sure, this record does not hold forth much hope for those scientists seeking help from mathematics for the problems of today—but in any event with patience all turns out well.

ADDED IN PROOF. 1. R. Bowen has answered in the affirmative; see *Smooth-partitions of Anosov diffeomorphisms are weak Bernoulli* to appear in Israel J. Math.

2. Cf. *Bernoullian translations and minimal horospheres on homogenous spaces* by S. G. Dani, Tata Institute Preprint.

3. For details see *Factors of Bernoulli shifts*, to appear in Israel J. Math. **20**.

References

1. R. L. Adler and B. Weiss, *Similarity of automorphisms of the torus*, Mem. Amer. Math. Soc. No. 98 (1970). MR **41** #1966.

2. J.-P. Conze, *Entropie d'un groupe abélien de transformations*, Z. Warscheinlichkeitstheorie und Verw. Gebiete **25** (1972), 11–30.

3. G. Gallavotti and D. Ornstein, *Billiards and Bernoulli schemes*, Comm. Math. Phys. (to appear).

4. Y. Katznelson, *Ergodic automorphisms of T^n are Bernoulli shifts*, Israel J. Math. **10** (1971), 186–195. MR **45** #3672.

5. Y. Katznelson and B. Weiss, *Commuting measure-preserving transformations*, Israel J. Math. **12** (1972), 161–173. MR **47** #5227.

6. W. Krieger, *On entropy of groups of measure preserving transformations*, Ann. Probability (to appear).

7. D. Lind, *Ergodic automorphisms of the infinite torus are Bernoulli*, Israel J. Math. **17** (1974), 162–168.

8. ———, *Locally compact measure preserving flows*, Advances in Math. (to appear).

9. L. D. Mešalkin, *A case of isomorphism of Bernoulli schemes*, Dokl. Akad. Nauk SSSR **128** (1959), 41–44. (Russian) MR **22** #1650.

10. D. S. Ornstein, *Factors of Bernoulli shifts are Bernoulli shifts*, Advances in Math. **5** (1970), 349–364. MR **43** #479b.

11. ———, *Imbedding Bernoulli shifts in flows*, Contributions to Ergodic Theory and Probability (Proc. Conf., Ohio State Univ., Columbus, Ohio, 1970), Lecture Notes in Math., vol. 160, Springer-Verlag, Berlin and New York, 1970, pp. 178–218. MR **42** #7866.

12. ———, *Ergodic theory, randomness and dynamical systems*, Yale Univ. Press, New Haven, Conn., 1974.

13. D. Ornstein and B. Weiss, *Geodesic flows are Bernoullian*, Israel J. Math. **14** (1973), 184–198.

14. ———, *Finitely determined implies very weak Bernoulli*, Israel J. Math. **17** (1973), 94–104.

15. ———, *Every transformation is bilaterally deterministic*, Israel J. Math. **20** (to appear).

16. M. Ratner, *Anosov flows with Gibbs measures are also Bernoulli*, Israel J. Math. **17** (1974), 380–391.

17. P. Shields, *Theory of Bernoulli shifts*, Univ. of Chicago Press, Chicago, 1973.

18. J.-P. Thouvenot, *Quelques propriétés des systèmes dynamique qui se décomposant en un produit de deux systèmes dont l'un est un schéma de Bernoulli*, Israel J. Math. **20** (to appear).

19. ———, *Convergence en moyenne de l'information pour l'action de Z^2*, Z. Warscheinlichkeitstheorie und Verw. Gebiete **24** (1972), 135–137. MR **47** #10145.

20. B. Weiss, *The isomorphism problem in ergodic theory*, Bull. Amer. Math. Soc. **78** (1972), 668–684. MR **46** #3751.

HEBREW UNIVERSITY
JERUSALEM, ISRAEL

Section 15

Control Theory and Related Optimization Problems

Actes du Congrès International des Mathematiciens
Vancouver, 1974

Contrôle Impulsionnel et Inéquations Quasi Variationelles

A. Bensoussan

I. Introduction. On considère un système dont l'état à l'instant t est un vecteur de R^n donné soit par la solution de l'équation différentielle ordinaire

$$dy/dt = g(y), \qquad y(0) = x \in R^n, \tag{1.1}$$

soit par la solution de l'équation différentielle stochastique (du type Itô)

$$dy = g(y)dt + \sigma(y)dw(t), \qquad y(0) = x. \tag{1.2}$$

Les fonctions g et σ sont des fonctions Lipschitziennes et bornées. On exerce un contrôle sur l'évolution de la manière suivante: en une suite d'instants $0 \leqq \theta^1 \leqq \theta^2 \leqq \cdots$ l'état du système passe instantanément de $y(\theta^{i-})$ à $y(\theta^{i-}) + \xi^i$. Les instants θ^i sont appelés *instants d'impulsion* et les valeurs ξ^i *niveaux d'impulsion*. Entre deux instants d'impulsion successifs, l'évolution du système est décrite par (1.1) ou par (1.2). L'ensemble $v = \{\theta^1, \xi^1; \cdots; \theta^i, \xi^i; \cdots\}$ est appelé un *contrôle impulsionnel*. On se donne un critère

$$J_x(v) = E\left[\int_0^\infty e^{-\alpha t} f(y_x(t))\, dt + k \sum_i e^{-\alpha\theta^i}\right], \qquad k > 0, \alpha > 0, \tag{1.3}$$

et on s'intéresse à la fonction

$$u(x) = \operatorname*{Inf}_v J_x(v). \tag{1.4}$$

Notre objet est de caractériser la fonction u et d'en déduire le contrôle impulsionnel optimal, s'il existe. Cette approche est similaire à celle de Hamilton-Jacobi-Bellman pour le contrôle continu ou à celle de Bensoussan-Lions pour les problèmes de temps d'arrêt optimal. Nous renvoyons à l'exposé de J. L. Lions [1] à ce congrès

pour une présentation générale des différentes situations ainsi que pour les références bibliographiques.

Des situations du type contrôle impulsionnel sont fréquentes en économie ou en gestion (gestion des stocks, production, maintenance, etc., cf. Bensoussan et Lions [1], Breton et Leguay [1], Goursat, Leguay et Maarek [1]) ce qui explique en partie pourquoi les modèles en temps discret sont, dans ce domaine d'application, préférés aux modèles en temps continu (on peut alors appliquer la programmation dynamique en temps discret; toutefois on introduit des difficultés du type non convexité; cf. Veinott [1]).

II. Problème type de contrôle impulsionnel: conditions suffisantes d'optimalité. On précise tout d'abord les notations de l'introduction. On prend (pour simplifier) $\sigma(x) = \sigma I$ (I = matrice identité, σ réel > 0) et $f(x)$ une fonction continue bornée sur $R^n \geqq 0$ et $\in L^2(R^n)$. On se donne un espace de probabilité $(\Omega, \mathcal{A}, P)$ et une famille croissante de σ-algèbres $\mathcal{F}_t \subset \mathcal{A}$. Le processus $w(t)$ est un processus de Wiener (standard) par rapport à la suite $\mathcal{F}_t$ et à valeurs dans R^n. On se place systèmatiquement dans le cas stochastique (toutefois le cas deterministe s'obtient au moins formellement en faisant $\sigma = 0$; cela se justifie par ailleurs sous certaines conditions). Les θ^i forment une suite croissante de temps d'arrêt par rapport à $\mathcal{F}_t$. On suppose que ξ^i est $\mathcal{F}_{\theta^i}$ mesurable et $\xi^i \in K$ (sous ensemble compact de R^n). On se limite (sans perte de généralité) au cas où

$$\text{p.s.} \qquad \theta^i \to +\infty, \qquad i \to \infty.$$

L'évolution de l'état du système $y_x(t)$ est définie par

$$\begin{aligned} &dy_x = g(y_x)dt + \sigma dw(t), \qquad \theta^i \leqq t < \theta^{i+1}, \ \forall\, i \geqq 0, \\ &y_x(\theta^i) = y_x(\theta^{i-}) + \xi^i, \\ &y_x(0) = x, \end{aligned} \tag{2.1}$$

avec les notations

$$\theta^0 = 0, \qquad (\theta^{i+1})^- = \theta^i \quad \text{si } \theta^{i+1} = \theta^i \tag{2.2}$$

et le critère par (1.3).

On introduit alors le problème suivant: trouver $u(x)$ telle que

$$u \text{ est continue et bornée, } u \in H^1(R^n),\ \Delta u \in L^2(R^n), \tag{2.3}$$

$$-\frac{\sigma^2}{2}\Delta u - \sum_i g_i(x)\frac{\partial u}{\partial x_i} + \alpha u \leqq f \text{ p.p.}, \tag{2.4}$$

$$u \leqq Mu, \tag{2.5}$$

$$(u - Mu)\left[-\frac{\sigma^2}{2}\Delta u - \sum_i g_i(x)\frac{\partial u}{\partial x_i} + \alpha u - f\right] = 0 \text{ p.p.}, \tag{2.6}$$

où on posé

$$Mu(x) = k + \operatorname*{Inf}_{\xi \in K} u(x + \xi). \tag{2.7}$$

Si u est une fonction vérifiant (2.3)—(2.7) on pose

$$C_u = \{x \in R^n \,|\, u(x) < Mu(x)\}, \tag{2.8}$$

et on associe à u un contrôle impulsionnel $\hat{v} = v_u$ de la manière suivante: on considère l'équation

$$dy = g(y)dt + \sigma dw(t), \qquad y(0) = x,$$

et on pose

$$\hat{\theta}^1 = \operatorname*{Inf}_{t \geqq 0} \{y(t) \notin C_u\}, \tag{2.9}$$

$$\xi^1 = \xi(y(\hat{\theta}^{1-})), \tag{2.10}$$

où $\xi(x)$ est une fonction mesurable de $R^n \to K$ définie par la condition

$$\operatorname*{Inf}_{\xi \in K} u(x + \xi) = u(x + \xi(x)) \qquad \forall\, x. \tag{2.11}$$

D'une manière générale, connaissant $\hat{\theta}^i$ et $\hat{\xi}^i$ on considère l'équation

$$\begin{aligned} dy &= g(y)dt + \sigma dw(t), \qquad t \geqq \hat{\theta}^i, \\ y(\hat{\theta}^i) &= y(\hat{\theta}^{i-}) + \hat{\xi}^i, \end{aligned} \tag{2.12}$$

et on pose

$$\hat{\theta}^{i+1} = \operatorname*{Inf}_{t \geqq \hat{\theta}^i} \{y(t) \notin C_u\}, \tag{2.13}$$

$$\hat{\xi}^{i+1} = \xi(y(\hat{\theta}^{(i+1)-})). \tag{2.14}$$

On démontre (cf. Bensoussan et Lions [2]) les résultats suivants: p.s. $\hat{\theta}^i \to +\infty$, $i \to \infty$, de sorte que $\{\cdots \hat{\theta}^i, \hat{\xi}^i; \cdots\}$ est bien un contrôle impulsionnel $\hat{v}$ et

$$u(x) = J_x(\hat{v}) \leqq J_x(v) \qquad \forall\, v, \tag{2.15}$$

de sorte que l'existence d'une solution $u(x)$ du problème (2.3)—(2.6) implique bien *l'existence d'un contrôle impulsionnel optimal.* On remarquera que si $\mathscr{Y}_t \equiv \mathscr{Y}_t^{\hat{v}}$ désigne la famille de σ-algèbres engendrée par le processus $y(t) \equiv y(t; \hat{v})$ alors $\mathscr{Y}_t^{\hat{v}} \subset \mathscr{F}_t \; \forall\, t$, les instants d'impulsion sont des temps d'arrêt par rapport à cette famille et les niveaux d'impulsion sont obtenus à l'aide d'un *feedback* $\xi(x)$. Le contrôle impulsionnel optimal est *markovien.* On retrouve ici l'analogue d'un résultat fondamental de la théorie du contrôle stochastique, à savoir que lorsque l'on peut observer l'évolution de l'état, alors il existe en général un contrôle stochastique markovien optimal dans la classe des contrôles markoviens ou non (cf. Fleming [1], Fleming et Rishel [1]).

III. Inéquations quasi variationnelles. Le problème mathématique restant à résoudre est donc celui de l'existence d'une fonction u solution de (2.3)—(2.6). Pour cela J. L. Lions et l'A, ont introduit une technique nouvelle, susceptible d'applications nombreuses même en dehors du contrôle impulsionnel. Cette technique généralise celle des inéquations variationnelles (I.V.) de Lions et Stampacchia [1], ce qui justifie la terminologie d'inéquations quasi variationnelles (I.Q.V.). On introduit la forme bilinéaire $a(u, v)$ sur $H^1(R^n)$ définie par

$$a(u, v) = \frac{\sigma^2}{2} \int_{R^n} \operatorname{grad} u \cdot \operatorname{grad} v \, dx - \sum_i \int_{R^n} g_i(x) \frac{\partial u}{\partial x_i} v \, dx + \alpha \int_{R^n} uv \, dx. \tag{3.1}$$

On cherche u solution de

$$u \in L^{\infty}_{loc}(R^n) \cap H^1(R^n), \qquad u \leqq Mu \text{ p.p.}, \tag{3.2}$$

$$a(u, v-u) \geqq (f, v-u) \qquad \forall\, v \in H^1(R^n) \text{ avec } v \leqq Mu \text{ p.p.} \tag{3.3}$$

On a noté par $(\ ,\ )$ le produit scalaire dans $L^2(R^n)$ ($((\ ,\ ))$ désignera le produit scalaire dans $H^1(R^n)$.) La relation entre (3.2), (3.3) et (2.3)—(2.6) est la suivante: s'il existe u solution de (3.2), (3.3) vérifiant en outre les *propriétés de régularité* (2.3) alors u satisfait aussi à (2.4), (2.5), (2.6).

L'étude de (3.2), (3.3) repose sur la propriété fondamentale de monotonie de l'opérateur M, à savoir

$$u, v \in L^{\infty}_{loc}(R_n) \text{ et } u \leqq v \text{ p.p.} \quad \Rightarrow \quad Mu \leqq Mv \text{ p.p.} \tag{3.4}$$

Pour démontrer l'existence d'une solution de (3.2), (3.3) on peut utiliser la méthode *constructive* suivante : on part de u^0 solution de

$$a(u^0, v) = (f, v) \qquad \forall\, v \in H^1(R^n),\ u^0 \in H^1(R^n). \tag{3.5}$$

Ayant défini u^n, on définit u^{n+1} solution de

$$\begin{aligned} &a(u^{n+1}, v - u^{n+1}) + \lambda(u^{n+1}, v - u^{n+1}) \geqq (f, v - u^{n+1}) + \lambda(u^n, v - u^{n+1}), \\ &u^{n+1} \leqq Mu_u, \quad \forall\, v \leqq Mu^n,\ n \geqq 1. \end{aligned} \tag{3.6}$$

Le nombre $\lambda > 0$ est choisi de façon que $a(v, v) + \lambda |v|^2 \geqq \beta \|v\|^2\ \forall\ v \in H^1(R^n)$, $\beta > 0$. A l'aide du principe du maximum, on vérifie que la suite u^n est décroissante et positive. On démontre alors que $u^n \downarrow u$ solution de (3.2), (3.3). Une autre technique non constructive mais plus générale, due à Tartar repose sur l'étude de la famille T_λ d'opérateurs (non linéaires) de $L^2(R^n) \to H^1(R^n)$ définie par : $u = T_\lambda w$ est solution de l'inéquation variationnelle

$$a(u, v-u) + \lambda(u, v-u) \geqq (f, v-u) + \lambda(w, v-u) \qquad \forall\, v \leqq Mw,\ u \leqq Mw. \tag{3.7}$$

Il est clair que si u est un point fixe de T_λ alors u est solution de (3.2), (3.3). L'application T_λ est croissante pour tout λ. Utilisant le théorème de Zorn, on montre alors l'existence d'un tel point fixe. L'unicité de la solution de (3.2), (3.3) peut se montrer d'une manière directe (cf. L. Tartar [1], Th. Laetsch [1]).

Le problème de la régularité de u (telle que (2.3)) est pour l'instant ouvert. On remarquera que, grâce à (2.15), il ne peut exister au plus qu'une solution de (2.3), ⋯, (2.6).

IV. Autres situations et exemples.

4.1. *Le cas d'évolution.* On peut considérer un problème de contrôle impulsionnel sur un horizon fini avec des fonctions $f(x, t)$ dépendant du temps. On est alors conduit à une I.Q.V. d'évolution. Posant $Q = R^n \times]0, T[$ on démonter (cf. Bensoussan et Lions [2]) l'existence d'une solution $u(x, t)$ du problème suivant (f étant supposée $\geqq 0, \in L^2(Q)$, $\partial f/\partial t \in L^2(Q)$, $\partial f/\partial t \leqq 0$)

$$\begin{aligned} &u \in L^{\infty}(0, T; H^1(R^n)) \cap L^{\infty}(0, T; L^{\infty}(\mathcal{O})) \qquad \forall\, \mathcal{O} \text{ ouvert borné de } R^n, \\ &du/dt \in L^2(0, T; L^2(R^n)), \qquad du/dt \leqq 0 \text{ p.p.}, \ u \geqq 0 \text{ p.p.}, \end{aligned} \tag{4.1}$$

$$u \leqq Mu \text{ p.p.}, \tag{4.2}$$

$$-(du/dt, v-u) + a(u, v-u) \geqq (f, v-u) \qquad \forall\, v \in H^1(R^n),\ v \leqq Mu, \tag{4.3}$$

$$u(T) = 0. \tag{4.4}$$

4.2. *Cas de contraintes sur l'état.* Lorsque l'on veut tenir compte de contraintes sur *l'état* à chaque instant, représentées par $y(t) \in \bar{\mathcal{O}}\ \forall\, t$, $\mathcal{O}$ ouvert de R^n, on est conduit à remplacer le modèle d'évolution (2.1) par un modèle d'équation différentielle stochastique sur un ouvert. On démontre alors que les relations (2.3)—(2.6) restent valables sur $\mathcal{O}$ (au lieu de R^n). Il convient toutefois de rajouter des conditions aux limites dont un exemple possible est le suivant

$$\begin{aligned} u &\leqq Mu \quad \text{sur } \Gamma = \partial\mathcal{O}, \\ \partial u/\partial n\big|_\Gamma &\leqq 0 \qquad (u - Mu)\,\partial u/\partial n\big|_\Gamma = 0, \end{aligned} \tag{4.5}$$

où $\partial u/\partial n$ est la dérivée normale de u sur Γ (n-normale orientée ver l'extérieur). Utilisant la formulation I.Q.V., il suffit de remplacer R^n par $\mathcal{O}$ dans (3.1), (3.2), (3.3). Une extension similaire est valable pour le cas d'évolution.

4.3. *Contrôle impulsionnel et contrôle continu.* Dans les applications du contrôle impulsionnel, on rencontre souvent la situation où entre deux instants d'impulsions, on exerce sur le système un contrôle continu. Dans ce cas, les fonctions f, g dépendent d'une variable de décision $d \in D$, soit $f \equiv f(x, d)$, $g = g(x, d)$. On définit l'Hamiltonien

$$H(x, p) = \operatorname*{Inf}_{d \in D} [f(x, d) + (p, g(x, d))]. \tag{4.6}$$

Les relations (2.3)—(2.6) doivent être remplacées par

$$u \text{ continue et bornée}, \quad u \in H^1(R^n),\ \Delta u \in L^2(R^n), \tag{4.7}$$

$$-(\sigma^2/2)\Delta u + \alpha u \leqq H(x, \partial u/\partial x) \text{ p.p. } H(x, \partial u/\partial x), \tag{4.8}$$

$$u \leqq Mu \text{ p.p.}, \tag{4.9}$$

$$(u - Mu)[-\sigma^2/2 - \sigma^2 \Delta u/2 + \alpha u - H(x, \partial u/\partial x)] = 0 \text{ p.p.} \tag{4.10}$$

On se reportera à Bensoussan et Lions [3], [5].

4.4. *Système d'I.Q.V.* D'autres applications (modèles de production en particulier) conduisent à des systèmes d'I.Q.V. On cherche deux (ou plus généralement n) fonctions

$$u_1, u_2 \text{ continues bornées sur } R^n, \quad \Delta u_i \in L^2(R^n),\ u_i \in H^1(R^n), \tag{4.11}$$

$$\begin{aligned} &-\frac{\sigma^2}{2}\Delta u_1 - \left(g^1, \frac{\partial u_1}{\partial x}\right) + \alpha u_1 \leqq f_1, \qquad u_1 \leqq u_2 + k_2, \\ &(u_1 - u_2 - k_2)\left[-\frac{\sigma^2}{2}\Delta u_1 - \left(g_1, \frac{\partial u_1}{\partial x}\right) + \alpha u_1 - f_1\right] = 0, \end{aligned} \tag{4.12}$$

$$\begin{aligned} &-\frac{\sigma^2}{2}\Delta u_2 - \left(g_2, \frac{\partial u_2}{\partial x}\right) + \alpha u_2 \leqq f_2, \qquad u_2 \leqq u_1 + k_1, \\ &(u_2 - u_1 - k_1)\left[-\frac{\sigma^2}{2}\Delta u_2 - \left(g_2, \frac{\partial u_2}{\partial x}\right) + \alpha u_2 - f_2\right] = 0 \end{aligned} \tag{4.13}$$

(cf. Bensoussan et Lions [4]).

De très nombreuses autres situations liées à des applications concrètes sont possibles (I.Q.V. avec contraintes bilatérales provenant de situations de jeux différentiels avec instants d'impulsions, I.Q.V. avec retard, I.Q.V. correspondant à des opérateurs intégro-différentiels, etc.). Elles seront décrites dans un ouvrage général de J. L. Lions et l'A.

Références

A. Bensoussan and J.-L. Lions

[1] *Temps optimal, contrôle impulsionnel et applications* (livre à paraître).

[2] *Nouvelles méthodes en contrôle impulsionnel.* I, J. Appl. Math. Opt., no. 4, (1974).

[3] *Nouvelle formulation de problèmes de contrôle impulsionnel et applications*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A1189–A1192. MR **47** #5690.
Contrôle impulsionnel et inéquations quasi-variationnelles d'évolution, C. R. Acad. Sci. Paris Sér. A-B **274** (1973), A1333-A1338. MR **47** #5692.

[4] *Contrôle impulsionnel et contrôle continu. Méthods des inéquations quasi variationnelles non linéaires*, C. R. Acad. Sci. Paris Sér. A-B **278** (1974), A675–A679.
Contrôle impulsionnel et systèmes d'inéquations quasi variationnelles, C. R. Acad. Sci. Paris Sér. A-B **278** (1974), A747–A751.

[5] *Optimal impulse and continuous control: Method of nonlinear quasi variational inequalities*, Steklov Institute (à paraître).

A. Bensoussan, M. Goursat and J.-L. Lions

[1] *Contrôle impulsionnel et inéquations quasi-variationnelles stationnaires*, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A1279–A1284. MR **47** #5691.

A. Breton and C. Leguay

[1] *Application du contrôle stochastique à la gestion des centrales thermiques et hydrauliques*, Collog. IRIA Juin 1974, Lecture Notes in Math., Springer-Verlag, Berlin and New York.

W. H. Fleming

[1] *Optimal continuous-parameter stochastic control*, SIAM Rev. **11** (1969), 470–509. MR **41** #9633.

W. H. Fleming and R. Rishel

[1] (livre à paraître).

M. Goursat, C. Leguay and G. Maarek

[1] (rapport laboria à paraître).

T. Laetsch

[1] Communication personnelle, 1974.

J.-L. Lions

[1] *Sur la théorie du contrôle*, these PROCEEDINGS.

J.-L. Lions and G. Stampacchia

[1] *Variational inequalities*, Comm. Pure Appl. Math. **20** (1967), 493–519. MR **15** #7178.

L. Tartar

[1] C. R. Acad. Sci. Paris Sér. A-B **278** (1974).

A. F. Veinott, Jr.

[1] *On the optimality of (s, S) inventory policies: New conditions and a new proof*, SIAM J. Appl. Math. **14** (1966), 1067–1083. MR **34** #6623.

UNIVERSITÉ DE PARIS IX
75016 PARIS, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Some Minimax Problems in Optimization Theory*

V. F. Demyanov

1. Many problems arising in engineering, economics and mathematics are of the form: Minimize a function $\varphi(x)$ subject to $x \in \Omega$ where $\varphi(x)$ is one of the following functions:

$$\varphi(x) = \max_{y \in G} f(x, y), \tag{1}$$

$$\varphi(x) = \max_{y \in G(x)} f(x, y), \tag{2}$$

$$\varphi(x) = \max_{y \in G_1(x)} \min_{z \in G_2(x)} f(x, y, z), \tag{3}$$

$$\varphi(x) = \max_{y_1 \in G_{11}(x)} \min_{z_1 \in G_{21}(x)} \cdots \max_{y_k \in G_{1k}(x)} \min_{z_k \in G_{2k}(x)} f(x, y_1, \cdots, y_k, z_1, \cdots, z_k) \tag{4}$$

and sets $G(x)$, $G_{ij}(x)$ depend on x, G is a given set.

Such problems often appear in the engineering design theory. In recent years much attention was paid to the problems described. We mention only some books dealing with minimax theory [1], [5], [7], [9], [13]. It seems possible to claim that at present the minimax theory is formed. The minimax theory deals with the following problems:

1. Investigation of the properties of the functions (1)—(4) including their directional differentiability. For various types of functions conditions for the function to be directionally differentiable are obtained and formulae for the first and higher order directional derivatives are found (see for example [6], [8], [9], [10], [18], [20], [21], [22]).

2. Necessary and sufficient conditions and their geometric interpretation [2], [3], [9].

3. Steepest-descent directions and their applications to constructing numerical

* Not presented in person.

methods. Numerical methods of the first order (of the gradient type). These problems have been widely disucssed and studied for the function (1). For this case the first order methods [13] as well as various second order methods (see for example [25]) have been worked out. Some useful estimations have been obtained [12], [16]. Active research is under way to obtain numerical methods for minimizing the function (2) (see [24]) and the function (3) [17]. But there is too much to be done in this field. The main problem for the immediate future is to develop software for minimax problems and its practical applications.

Sometimes it is possible making use of special properties of the problem to develop an effective method for its solution.

4. Saddle points. The problem of finding saddle points is a special case of minimax problems. For this case it is possible to construct methods where it is not necessary to calculate the value of the function (1) at each step (see surveys [11], [14]).

5. Optimal control problems with a minimax criterion function.

6. Nonlinear approximation problems [15], [19].

Now we discuss Problems 1 and 5 in detail.

2. Let $\varphi(x) = \max_{y \in G(x)} f(x, y)$ where $x \in E_n$, $y \in E_m$. Fix x_0 and $g \in E_n$. Let

$$\gamma(y) = \{V \in E_m \mid \exists\, \alpha_0 > 0 : y + \alpha v \in G(x_0 + \alpha g)\ \forall\, \alpha \in [0, \alpha_0]\}.$$

The closure of $\gamma(y)$ we denote by $\Gamma(y) = \Gamma(x_0, y, g)$. It is known [9], [22] that under some additional conditions the function $\varphi(x)$ is differentiable at the point x_0 w.r. to the direction of g and

$$\frac{\partial \varphi(x_0)}{\partial g} \equiv \lim_{\alpha \to +0} \frac{1}{\alpha}[\varphi(x_0 + \alpha g) - \varphi(x_0)] = \sup_{y \in R(x_0)} \sup_{v \in \Gamma(y)} \left[\left(\frac{\partial f}{\partial y}, v\right) + \left(\frac{\partial f}{\partial x}, g\right)\right].$$

Higher order derivatives we define as follows. Suppose that $l \geqq 2$ and that it is already known that

$$\varphi(x_0 + \alpha g) = \varphi(x_0) + \sum_{k=1}^{l-1} \frac{\partial^k \varphi(x_0)}{\partial g^k} \frac{\alpha^k}{k!} + o(\alpha^{l-1})$$

where $\partial^k \varphi / \partial g^k$ are derivatives of the function φ w.r. to the direction of g at the point x_0. Then the limit

$$\frac{\partial^l \varphi(x_0)}{\partial g^l} = \lim_{\alpha \to +0} \frac{l!}{\alpha^l} \left[\varphi(x_0 + \alpha g) - \varphi(x_0) - \sum_{k=1}^{l-1} \frac{\partial^k \varphi(x_0)}{\partial g^k} \frac{\alpha^k}{k!}\right],$$

if it exists is called the lth derivative of the function $\varphi(x)$ at the point x_0 w.r. to the direction of g.

Now let us introduce the set $\Gamma^l(y, v_1, \cdots, v_{l-1})$ of feasible directions of the lth order. Suppose that sets $\Gamma^1(y)$, $\Gamma^2(y, v_1)$, $\cdots$, $\Gamma^{l-1}(y, v_1, \cdots, v_{l-2})$ have already been defined. Fix $y \in G(x_0)$, $v_1 \in \Gamma^1(y)$, $\cdots$, $v_{l-1} \in \Gamma^{l-1}(y, v_1, \cdots, v_{l-2})$ and define the set

$$\gamma^l(y, v_1, \cdots, v_{l-1}) = \left\{v_l \in E_m \mid \exists\, \alpha_0 > 0 : y + \sum_{k=1}^{l} \alpha^k \cdot v_k \in G(x_0 + \alpha g)\ \forall\, \alpha \in [0, \alpha_0]\right\}.$$

The closure of $\gamma^l(y, v_1, \cdots, v_{l-1})$ let us denote by $\Gamma^l(y, v_1, \cdots, v_{l-1})$ and call it the set of feasible directions of the lth order (or course γ^i and Γ^i depend on x_0 and g).

Suppose that the function f is l times continuously differentiable. Then for any $\sigma \in [1 : l]$ the following expansion is valid:

$$f\Big(x_0 + \alpha g, y + \sum_{k=1}^{\sigma} \alpha^k v_k + o(\alpha^\sigma)\Big) = f(x_0, y) + \sum_{k=1}^{\sigma} A_k(x_0, y, g, v_1, \cdots, v_k)\frac{\alpha^k}{k!} + o(\alpha^\sigma) \tag{5}$$

where A_k does not depend on σ and is a function of the derivatives of the function f of order $\leqq k$. Suppose that for any sequence $\{y_s\}$, $y_s \in R(x_0 + \alpha_s g)$, $\alpha_s \to +0$, there exists a subsequence $\{y_{s_i}\}$ which can be written as

$$y_{s_i} = \bar{y} + \sum_{k=1}^{l-1} \alpha_i^k \bar{v}_k + \alpha_i^l v_{li} + o(\alpha_i^l)$$

where $\bar{y} \in G(x_0)$, $\bar{v}_1 \in \Gamma^1(\bar{y})$, $\cdots$, $\bar{v}_{n-1} \in \Gamma^{n-1}(\bar{y}, \bar{v}_1, \cdots, \bar{v}_{l-2})$, $v_{li} \in \Gamma^l(\bar{y}, \bar{v}_1, \cdots, \bar{v}_{l-1})$, $\alpha_i v_{li} \to 0$ as $i \to \infty$.

THEOREM 1. *Let $l \geqq 2$. If there exists the first derivative of the function $\varphi(x)$ at the point x_0 w.r. to the direction of g and* (5) *is true, then under the assumptions above there exists the derivative of any order $\sigma \in [2 : n]$ and*

$$\frac{\partial^\sigma \varphi(x_0)}{\partial g^\sigma} = \sup_{[y, v_1, \cdots, v_{\sigma-1}] \in T^{\sigma-1}} \sup_{v^\sigma \in \Gamma^\sigma(y, v_1, \cdots, v_{\sigma-1})} A_\sigma(x_0, y, g, v_1, \cdots, v_\sigma) \tag{6}$$

where $A_\sigma(x_0, y, g, v_1, \cdots, v_\sigma)$ is taken from (5); *$T^{\sigma-1}$ is the set of elements of $[y, v_1, \cdots, v_{\sigma-1}]$ such that supremum in the formula for the $(\sigma-1)$th derivative is achieved at points $[y, v_1, \cdots, v_{\sigma-1}]$. Note that $T^{\sigma-1}$ is not empty for $\sigma \in [2 : n]$.*

3. Minimax problems in optimal control. Let $\dot{x}(t) = f(x, u, t)$, $x(0) = x_0$, where $x = (x_1, \cdots, x_n)$, $f = (f_1, \cdots, f_n)$, $u = (u_1, \cdots, u_r)$ and the functions f_i and $\partial f_i/\partial x$ are continuous in all variables. By U let us denote the class of piecewise continuous controls $u(t)$ such that $u(t) \in W \subset E_r$ for any $t \in [0, T]$. Let $I(u, z) = \int_0^T g(x, u, t, z)\, dt$ where $z \in Z \subset E_p$, the functions g and $\partial g/\partial x$ are continuous in all variables. Now let us consider the problem

$$\max_{z \in Z} I(u, z) \longrightarrow \min_{u \in U}. \tag{7}$$

Under some additional conditions the following result is valid.

THEOREM 2 (SEE [4]). *For a control $u^* \in U$ to be an optimal one it is necessary that*

$$\min_{u \in U} \max_{z \in R(u^*)} \int_0^T [H_z(u, \tau) - H_z(u^*, \tau)]\, d\tau = 0$$

where $R(u) = \{z \in Z \,|\, I(u, z) = \max_{v \in Z} I(u, v)\}$,

$$\frac{d\psi_z(\tau)}{d\tau} = -\left(\frac{\partial f(x(\tau, u^*), u^*, \tau)}{\partial x}\right)^* \psi_z(\tau) - \frac{\partial g(x(\tau, u^*), u^*, \tau, z)}{\partial x}, \quad \psi_z(T) = 0.$$

References

1. A. Auslander, *Problèmes de minimax via l'analyse convexe et les inégalités variationnelles. Théorie et algorithmes*, Lecture Notes in Economics and Math. Systems, Springer-Verlag, Berlin and New York, 1972.

2. V. V. Beresnev, *Necessary conditions for extremum in a convex maximin problem on connected*

sets, Kibernetika (Kiev) No. 2 (1972), 87–91. (Russian) MR **46** #3079.

3. V. G. Boltjanskiĭ and I. S. Čebotaru, *Necessary conditions in a minimax problem*, Dokl. Akad. Nauk SSSR **213** (1973), 257–260 = Soviet Math. Dokl. **14** (1973), 1665–1668.

4. T. K. Vinogradova and V. F. Demyanov, *On the minimax principle in optimal control problems*, Dokl. Akad. Nauk SSSR **213** (1973), 512–514 = Soviet Math. Dokl. **14** (1973), 1720–1722.

5. Ju. B. Germeĭer, *Introduction to the operations research theory*, "Nauka", Moscow, 1971. (Russian)

6. E. G. Gol'šteĭn, *Duality theory in mathematical programming and its applications* "Nauka", Moscow, 1971. (Russian) MR **48** #893.

7. J. M. Danskin, Jr., *The theory of max-min and its application to weapons allocation problems*, Econometrics and Operations Research, vol. 5, Springer-Verlag, New York, 1967. MR **17** #3843.

8. B. S. Darhovskiĭ and E. S. Levitin, *On differential properties of minimax in one distribution resources problem*, Ž. Vyčisl. Mat. i Mat. Fiz. **13** (1973), no. 5, 1219–1227. (Russian)

9. V. F. Demyanov, *Minimax: Directional differentiability*, Leningrad Univ. Press, Leningrad, 1974. (Russian)

10. V. F. Demyanov and A. B. Pevnyĭ, *First and second marginal values of mathematical programming problems*, Dokl. Akad. Nauk SSSR **207** (1972), 277–280 = Soviet Math. Dokl. **13** (1972), 1502–1506.

11. ———, *Numerical methods of searching saddle points*, Ž. Vyčisl. Mat. i Mat. Fiz. **12** (1972), no. 5, 1099–1127. (Russian)

12. ———, *Certain estimates in minimax problems*, Kibernetika (Kiev) No. 1 (1972), 107–112. (Russian) MR **46** #1354.

13. V. F. Demyanov and V. N. Malozemov, *Introduction to minimax*, "Nauka", Moscow, 1972. (Russian)

14. S. I. Zuhovickiĭ, R. A. Poljak and M. E. Primak, *Concave n-person games (numerical methods)*, Ékonom. i Mat. Metody **7** (1971), 880–900. (Russian) MR **45** #6460.

15. V. N. Malozemov and A. B. Pevnyĭ, *Alternation properties of solutions of nonlinear minimax problems*, Dokl. Akad. Nauk SSSR **212** (1973), 37–39 = Soviet Math. Dokl. **14** (1973), 1303–1306.

16. A. B. Pevnyĭ, *Minimization of a maxmin function*, Ž. Vyčisl. Mat. i Mat. Fiz. **12** (1972), no. 1, 227–230. (Russian) MR **46** #1089.

17. ———, *On convergence rate of some methods for finding minimax and saddle points*, Kibernetika (Kiev) No. 4 (1972), 95–98. (Russian)

18. B. N. Pšeničnyĭ, *Convex multivalued mappings and their conjugates*, Kibernetika (Kiev) No. (1972), 94–102. (Russian) MR **46** #7926.

19. J. R. Rice, *The approximation of functions*. Vol. 1: *Linear theory*, Addison-Wesley, Reading, Mass., 1964. MR **29** #3795.

20. A. I. Sotskov, *Necessary conditions for a minimum for a type of nonsmooth problem*, Dokl. Akad. Nauk SSSR **189** (1969), 261–264 = Soviet Math. Dokl. **10** (1969), 1410–1413. MR **41** #4353.

21. Ju. N. Tjurin, *A mathematical statement of a simplified model of industrial planning*, Economics and Math. Methods **1** (1965), 391–410. (Russian)

22. W. W. Hogan, *Directional derivatives for extremal value functions*, Working Paper no. 177, Western Management Science Inst., Los Angeles, Calif., 1971.

23. A. B. Pevnyĭ, *On the finding of minimax under connected constraints*, Optimization (Novosibirsk) **10** (**27**) (1973), 22–29. (Russian)

24. E. F. Voĭton and L. N. Poljakova, *On one class of extremal problems with connected constraints*, Optimization (Novosibirsk) **10** (**27**) (1973), 41–46. (Russian)

25. L. V. Vasiljev, *On finding the steepest descent direction in minimax problems*, Sixth All-Union Conf. on Extremal Problems, Tallinn, April 16–19, 1973, Abstracts of Reports, vol. I, pp. 61–62. (Russian)

LENINGRAD STATE UNIVERSITY
LENINGRAD, U.S.S.R.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Stochastic Differential Games with Stopping Times and Variational Inequalities

Avner Friedman

Consider a system of n stochastic differential equations

$$dz(t) = f(x(t), t, y, z)\,dt + \sigma(x(t), t)\,dw(t) \tag{1}$$

where $\sigma(x, t)$ is an $n \times n$ matrix, $f(x, t, y, z)$ is an n-vector, both uniformly Lipschitz continuous in $(x, t) \in R^n \times [0, \infty)$, and $w(t)$ is n-dimensional Brownian motion. The variables y, z are viewed as control functions. They are taken to be measurable functions $y = y(x, t)$, $z = (x, t)$ with values in compact sets Y and Z respectively. We say that Y and Z are the *control sets* for the players y and z respectively. We are also given a *pay-off*

$$\begin{aligned} P_{\xi\tau}(y, S; z, T) = E_{\xi,\tau}\Bigg\{ & \int_{\tau}^{S\wedge T} \exp\left[\int_{s}^{t} k(x, s, y, z)\,ds\right] h(x, t, y, z)\,dt \\ & + \exp\left[\int_{\tau}^{S} k(x, t, y, z)\,dt\right] g_1(x(S), S)\,\chi_{S\leqq T} \\ & + \exp\left[\int_{\tau}^{T} k(x, t, y, z)\,dt\right] g_2(x(T), T)\Bigg\}\,\chi_{T<S} \end{aligned} \tag{2}$$

where k, h, g_1, g_2 are, say, smooth functions with bounded second derivatives. Here S and T are stopping times with range in $[\tau, T_0]$ for the process (1) with $y = y(x, t)$, $z = z(x, t)$ and T_0 is a fixed positive number.

The aim of the player y is to choose $(y(x, t), S)$ so as to maximize the pay-off, and the aim of the player z is to choose $(z(x, t), T)$ so as to minimize the pay-off. We shall refer to (1), (2) as a *stochastic differential game with stopping time.*

A pair $\{(y^*(x, t), S^*), (z^*(x, t), T^*)\}$ is called a *saddle point if*

$$P_{\xi,\tau}(y, S; z^*, T^*) \leqq P_{\xi,\tau}(y^*, S^*; z^*, T^*) \leqq P_{\xi,\tau}(y^*, S^*; z, T)$$

for any pairs (y, S) and (z, T). If such a saddle point exists, then the number

$$V(\xi, \tau) = P_{\xi,\tau}(y^*, S^*; z^*, T^*)$$

is called the *value* of the game.

Let σ^* be the transpose of σ, and let $a = \sigma\sigma^*$. Set

$$Lu = \tfrac{1}{2} \sum a_{ij} \partial^2 u / \partial x_i \partial x_j$$

where $a = (a_{ij})$. It is easy to see that if a saddle point exists then

$$g_2(x, t) \geqq g_1(x, t), \qquad g_2(x, T_0) = g_1(x, T_0). \tag{3}$$

We shall now assume that (3) holds, that

$$\sum a_{ij}(x, t)\xi_i\xi_j \geqq \alpha |\xi|^2 \quad \text{for all } (x, t),\ \xi \in R^n\ (\alpha > 0), \tag{4}$$

and that the *minimax condition* holds:

$$\begin{aligned} &\max_{y \in Y} \min_{z \in Z} [h(x, t, y, z) + p \cdot f(x, t, y, z) + uk(x, t, y, z)] \\ &\quad = \min_{z \in Z} \max_{y \in Y} [h(x, t, y, z) + p \cdot f(x, t, y, z) + uk(x, t, y, z)] \\ &\quad = H(x, t, u, p). \end{aligned} \tag{5}$$

Consider the nonlinear parabolic variational inequality

$$\begin{aligned} u &\in L^p(0, T_0; W^{2,p,\mu}) \cap L^\infty(0, T_0; W^{1,2}), \\ \partial u/\partial t &\in L^p(0, T_0; W^{0,p,\mu}), \qquad g_1 \leqq u \leqq g_2, \\ [\partial u/\partial t + Lu + H(x, t, u, u_x)]\,(v - u) &\geqq 0 \quad \text{a.e. for every } v, g_1 \leqq v \leqq g_2, \\ u(x, T_0) &= g_1(x, T_0) \end{aligned} \tag{6}$$

with any $p > n$. Here $W^{i,p,\mu}$ is $W^{i,p}(R^n)$ with any density function $e^{-\mu|x|}$, $\mu > 0$.

THEOREM 1. (i) *There exists a unique solution u of* (6).

(ii) *Let $y^*(x, t)$, $z^*(x, t)$ be any control functions which realize the* $\max_y$ *and* $\min_z$ *in* (5) *when $p = u_x(x, t)$. Let S^* be the exit time from the set $\{u(x, t) > g_1(x, t)\} \cap \{\tau < t \leqq T_0\}$ and let T^* be the exit time from the set $\{u(x, t) < g_2(x, t)\} \cap \{\tau < t \leqq T_0\}$. Then $\{(y^*, S^*), (z^*, T^*)\}$ is a saddle point.*

(iii) $V(\xi, \tau) = u(\xi, \tau)$.

Theorem 1 is due to Bensoussan and Friedman [1]. The special case where there are no stopping times was proved earlier by Friedman [5]. The special case where there are no controls $y(x, t)$, $z(x, t)$ was proved by Bensoussan and Lions [2] and Friedman [6], [7] (in [2] there is only one player); Krylov [11] has considered the corresponding stationary case.

Consider next the case where $\sigma(x, t) \equiv 0$, *i.e.*, (1) is replaced by a deterministic dynamical system

$$dx(t) = f(x, t, y, z)dt \tag{7}$$

and $P_{\xi,\tau}(y, S; z, T)$ is defined by (2), but with $E_{\xi,\tau}$ removed. The stopping times S, T are now any numbers in the interval $[\tau, T_0]$. Since the assertion (i) of Theorem 1

is false in the present case, we proceed in a different manner:

In the theory of differential games [4] one takes control functions $y(t), z(t)$ and defines the concepts of upper value, value, δ-strategy, strategy, saddle point, etc. Then one proves that the upper value V^+ exists and, under some assumptions, the value exists. The differential game setting is that corresponding to taking $S \equiv T_0$, $T \equiv T_0$. Bensoussan and Friedman [1] have generalized the basic concepts and existence theorems in the theory of differential games to differential games with stopping times. In particular, they proved:

THEOREM 2. (i) *The upper value $V^+(x, t)$ exists and satisfies a.e. the first order nonlinear variational inequality*

$$\left\{\frac{\partial V^+}{\partial t} + \min_{z\in Z}\max_{y\in Y}\left[h(x,t,y,z) + \frac{\partial V^+}{\partial x}\cdot f(x,t,y,z) + V^+ k(x,t,y,z)\right]\right\}(v - V^+) \leqq 0 \quad \text{a.e. for any } v,\ g_1 \leqq v \leqq g_2.$$

(ii) *If $f = f_1(x, t, y) + f_2(x, t, z)$, $h = h_1(x, t, y) + h_2(x, t, z)$ then the value exists.*

Consider now the case where $\sigma = \varepsilon I$, I the identity matrix. Denote the value occurring in Theorem 1 by $V_\varepsilon(x, t)$.

In the special case where there are no stopping times, i.e., $S \equiv T_0$, $T \equiv T_0$, it is known [3], [8] that

$$V_\varepsilon(\xi, \tau) \to V(\xi, \tau) \quad \text{if } \varepsilon \to 0 \tag{8}$$

where $V(\xi, \tau)$ is the value of the deterministic differential game (with dynamics (7)). Bensoussan and Friedman [1] have proved (8) in case S is any stopping time but $T \equiv T_0$, provided

$$h + \frac{\partial g_1}{\partial t} + \frac{\partial g_1}{\partial x}\cdot f + k g_1 \leqq 0.$$

We return to the situation of Theorem 1. It is of great interest to study the *domains of continuation* $C_1 = \{u > g_1\}$, $C_2 = \{u < g_2\}$.

In the case where there is only one player, say y, and, furthermore, there are no control functions $y = y(x, t)$ in (1), (2), there are some recent results on the shape and smoothness of the boundary of C_1 [12], [13], [10], [9]. It would be of interest to obtain such results in the case where there is a control function y in the system (1) and in the pay-off (2).

References

1. A. Bensoussan and A. Friedman, *Nonlinear variational inequalities and differential games with stopping times*, J. Functional Analysis **16** (1974), 305–352.

2. A. Bensoussan and J.-L. Lions, *Problèmes de temps d'arrêt optimal et inéquations variationelles parabolique*, Applicable Anal. **3** (1973), 267–295.

3. J. Elliott and N. J. Kalton, *Upper values of differential games*, J. Differential Equations **14** (1973), 89–100.

4. A. Friedman, *Differential games*, Wiley, New York, 1971.

5. ———, *Stochastic differential games*, J. Differential Equations **11** (1972), 79–108. MR **45** #1613.

6. A. Friedman, *Stochastic games and variational inequalities*, Arch. Rational Mech. Anal. **51** (1973), 321–346.

7. ———, *Regularity theorems for variational inequalities in unbounded domains and applications to stopping time problems*, Arch. Rational Mech. Anal. **52** (1973), 134–160.

8. ———, *Differential games*, CBMS Regional Conference Series in Math., no. 18, Amer. Math. Soc., Providence, R. I., 1974.

9. ———, *Parabolic variational inequalities in one space dimension and the smoothness of the free boundary*, J. Functional Anal. **18**, 151–176.

10. A. Friedman and D. Kinderleherer, *A one-phase Stefan problem*, Indiana Univ. J. Math. (to appear).

11. N. V. Krylov, *Control of Markov processes and W-spaces*, Izv. Akad. Nauk SSSR Ser. Mat. **35** (1971), 224–255 = Math. USSR Izv. **5** (1971), 233–266. MR **45** #4493.

12. P. van Moerbeke, *Optimal stopping and free boundary problems*, Arch. Rational Mech. Anal. (to appear).

13. ———, *An optimal stopping problem for linear reward*, Acta Math. **132** (1974), 1–41.

NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60201, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Necessary and Sufficient Conditions for Local Controllability and Time Optimality

Henry Hermes*

Introduction. Let M be an analytic n manifold and $\mathscr{C} = \{X^\alpha : \alpha \in A\}$ a collection of analytic vector fields on M. A continuous map $\varphi : [0, T] \to M$, $T > 0$, is called a *solution of* $\mathscr{C}$ if it is a "piecing together" of integral curves of the X^α, specifically if there exists a partition $0 = t_0 < t_1 < \cdots < t_k = T$ and $\alpha_1, \cdots, \alpha_k \in A$ such that $d\varphi/dt = \dot{\varphi} = X^{\alpha_i}(\varphi(t))$ on $[t_{i-1}, t_i)$, $i = 1, \cdots, k$. We denote by $\mathscr{A}(t, p, \mathscr{C})$ the set of all points in M attainable at time $t \geqq 0$ by solutions of $\mathscr{C}$ initiating from p at time 0.

The set of all analytic vector fields on M, denoted $V(M)$, may be considered as a real Lie algebra with product the Lie product $[X, Y]$. For any collection of vector fields $\mathscr{C}$, let $L(\mathscr{C})$ denote the smallest subalgebra containing $\mathscr{C}$; let TM_q denote the tangent space to M at q and $\mathscr{C}_q = \{X(q) \in TM_q : X \in \mathscr{C}\}$. Associated with $L(\mathscr{C})$ is an ideal, $L_0(\mathscr{C})$, consisting of elements of the form $\sum_{i=1}^k \lambda_i X^{\alpha_i} + W$, where λ_i are real, $\sum \lambda_i = 0$ and W belongs to the derived algebra of $L(\mathscr{C})$. Sussmann and Jurdjevic [1] show that a necessary and sufficient condition that $\bigcup_{t \geqq 0} \mathscr{A}(t, p, \mathscr{C})$ have nonempty interior is that $\dim L(\mathscr{C})_p = n$. (See Krener [2] for a short proof of this.) Furthermore, they show that a necessary and sufficient condition that $\operatorname{int} \mathscr{A}(t, p, \mathscr{C}) \neq \emptyset$ for all $t > 0$ is that $\dim L_0(\mathscr{C})_p = n$.

Let $X \in \mathscr{C}$ and $T^X(\cdot)p$ be its associated flow, i.e., the solution of $\dot{x} = X(x)$, $x(0) = p$. This manuscript deals with the problem of giving necessary and sufficient conditions that $T^X(t)p \in \operatorname{int} \mathscr{A}(t, p, \mathscr{C})$ for all $t > 0$. If this occurs we say *the system* $\mathscr{C}$ *is locally controllable along* T^X *at* p. The Pontryagin maximum principle for a time optimal problem associated with $\mathscr{C}$ gives "first order" necessary conditions that $T^X(t)p \in \partial \mathscr{A}(t, p, \mathscr{C})$ (here ∂ denotes boundary) for $0 \leqq t \leqq t_1$. Our results are

* This research was supported by the National Science Foundation under grant GP-27957.

related to a higher order maximum principle. Throughout most of the paper we consider the special system

$$\mathscr{E} = \left\{X + \sum_{2}^{n} \alpha_i V^i : -1 \leqq \alpha_i \leqq 1\right\} \tag{1}$$

where X, V^i are analytic vector fields on M. We choose $T^X(\cdot)p$, i.e., all $\alpha_i = 0$, as the reference trajectory. Let $(\mathrm{ad}\, X, Y) = [X, Y]$ and inductively, $(\mathrm{ad}^k X, Y) = [X, (\mathrm{ad}^{k-1} X, Y)]$. Define

$$\mathscr{S}^0 = \{(\mathrm{ad}^j X, V^i) : i = 2, \cdots, n; j = 0, 1, \cdots\}. \tag{2}$$

One may show $L_0(\mathscr{E}) = L(\mathscr{S}^0)$ so dim $L(\mathscr{S}^0)_p = n$ is necessary and sufficient that int $\mathscr{A}(t, p, \mathscr{E}) \neq \emptyset$ for all $t > 0$. It is well known [3] that rank $\mathscr{S}^0_p = n$ is sufficient that $T^X(t)p \in \text{int}\, \mathscr{A}(t, p, \mathscr{E})\ \forall\ t > 0$. Theorem 1 gives, under certain assumptions, necessary and sufficient conditions that $T^X(t)p \in \text{int}\, \mathscr{A}(t, p, \mathscr{E})\ \forall\ t > 0$. This theorem is an improvement of the result announced in [4] in the case dim $n = 2$.

One may show that the "Pontryagin first order approximating cone" along T^X at $p^1 = T^X(t_1)p$, $t_1 > 0$, is span $\mathscr{S}^0_{p^1}$. If dim span $\mathscr{S}^0_{T^X(\tau)p} < n$ for $0 \leqq \tau \leqq t_1$, the solution $T^X(\cdot)p$ is called *singular* on $[0, t_1]$ in that the maximum principle gives no positive information. One can develop higher order maximum principles (see [5]) by constructing cones which, for this problem, include elements of $L(\mathscr{S}^0)$ not in $\mathscr{S}^0$. Theorem 1 considers such elements, and in the analytic case, resolves the question of whether or not a singular solution is time optimal. This is related to, but quite different from, the work of Kelley, Kopp and Moyer [6], who minimize a real function on $\mathscr{A}(T, p, \mathscr{E})$, $T > 0$ fixed.

In §1 we state the main Theorem 1, the simplification which occurs in two dimensions, and a brief outline of its proof. In §2 we illustrate the applications of the theorem via examples. Details of proofs and similar results for more general systems will appear elsewhere.

The author would like to acknowledge helpful conversations and correspondence with Professor A. J. Krener.

1. The main theorem. We make the following assumptions:

(a.1) $X(p)$, $V^2(p)$, $\cdots$, $V^n(p)$ are linearly independent.

(a.2) V^2, $\cdots$, V^n are involutive.

Let Z be a smooth one-form on M uniquely determined by the conditions

$$\langle Z(x), X(x)\rangle \equiv 1, \quad \langle Z(x), V^i(x)\rangle \equiv 0, \quad i = 2, \cdots, n, \tag{3}$$

for x in some nbd. of p. We introduce the notation: $\nu = (\nu_2, \cdots, \nu_n)$ with the ν_i nonnegative integers; $|\nu| = \sum \nu_i$; $a(\nu, j) = \langle Z(p), (\mathrm{ad}^j X, (\mathrm{ad}^{\nu_2} V^2, \cdots, (\mathrm{ad}^{\nu_n} V^n, X)\cdots)(p)\rangle$; $s_2 = (s, \cdots, s_n)$ with s_i real and in a nbd. of 0;

$$\varphi_{mj}(s) = \sum_{|\nu|=m} a(\nu, j)(-s_2)^{\nu_2} \cdots, (-s_n)^{\nu_n} \quad \text{an "}m\text{-form".}$$

THEOREM 1. *Assume* (a.1), (a.2). *A necessary and sufficient condition that* $\mathscr{A}(t, p, \mathscr{E})$ *have nonempty interior for all* $t > 0$ (*i.e., that* dim $L(\mathscr{S}^0)_p = n$) *is that some* m-*form* $\varphi_{mj}(s) \not\equiv 0$, $m \geqq 1$, $j \geqq 0$. *If* $m^* \geqq 1$ *is the smallest integer such that* $\varphi_{m^*j} \not\equiv 0$

for some $j \geqq 0$ and j^* the smallest j for which this occurs, a necessary and sufficient condition that $T^X(t)p \in \operatorname{int} \mathscr{A}(t, p, \mathscr{E})\ \forall\ t > 0$ is that $\varphi_{m^*j^*}(s)$ assumes both positive and negative values in every nbd. of $0 \in \boldsymbol{R}^{n-1}$. (Note that if m^* is odd, $\varphi_{m^*j}(s) = -\varphi_{m^*j}(-s)$; hence $T^X(t)p \in \operatorname{int} \mathscr{A}(t, p, \mathscr{E})\ \forall\ t > 0$.)

For dim $n = 2$, Theorem 1 admits a simpler statement, as announced in [4]. Indeed, for $n = 2$, assumption (a.2) is vacuous while $\nu = \nu_2$, $s = s_2$, $V = V^2$ and

$$\varphi_{mj}(s) = s^m\langle Z(p), (\mathrm{ad}^j X, (\mathrm{ad}^m\ V, X))(p)\rangle.$$

Let

$$\mathscr{S}^1 = \{V, (\mathrm{ad}\ V, X), (\mathrm{ad}\ X, (\mathrm{ad}\ V, X)), (\mathrm{ad}^2 X, (\mathrm{ad}\ V, X)), \cdots\},$$

$$\mathscr{S}^2 = \{V\ (\mathrm{ad}^2\ V, X), (\mathrm{ad}\ X, (\mathrm{ad}^2\ V, X)), (\mathrm{ad}^2 X, (\mathrm{ad}^2\ V, X)), \cdots\},$$

$$\mathscr{S}^3 = \{V, (\mathrm{ad}^3\ V, X), (\mathrm{ad}\ X, (\mathrm{ad}^3\ V, X)), \cdots\},$$

etc. Note that, other than for V, the elements of $\mathscr{S}^0$ and $\mathscr{S}^1$ differ only in sign; thus rank $\mathscr{S}^1_p$ = rank $\mathscr{S}^0_p$.

THEOREM 1 (CASE $n = 2$). *Assume $X(p)$, $V(p)$ are linearly independent. A necessary and sufficient condition that* $\operatorname{int} \mathscr{A}(t, p, \mathscr{E}) \neq \emptyset\ \forall\ t > 0$ *is that there exists an integer $m \geqq 1$ such that* rank $\mathscr{S}^m_p = 2$. *A necessary and sufficient condition that $T^X(t)p \in$* $\operatorname{int} \mathscr{A}(t, p, \mathscr{E})$ *for all $t > 0$ is that the smallest such integer m be odd.*

REMARK 1. This statement differs slightly from that given in [4]. The above, stronger statement is more useful for computation since one need not first verify that dim $L(\mathscr{S}^0)_p = 2$.

We next give a brief sketch of the ideas involved in the proof of Theorem 1. If ψ is a solution of $\mathscr{E}$ with $\psi(0) = p$, $\psi(t_2) = T^X(t_1)p = p^1$ and $\Gamma = \{\psi(t): 0 \leqq t \leqq t_2\}$ we have $t_2 = \int_\Gamma Z$. If $\varphi(\tau) = \psi(t(\tau)), 0 \leqq \tau \leqq t_1$, is a reparametrization of ψ,

$$t_2 = \int_0^{t_1} \langle Z(\varphi(\tau)), \varphi'(\tau)\rangle\, d\tau.$$

The first goal is to reparametrize so that the integrand in the above expression depends only on the point $\varphi(\tau)$. From assumption (a.2), through each point $T^X(\tau)p\ \exists$ an $(n-1)$-dim. integral manifold, $M^{n-1}(T^X(\tau)p)$, for $V^2, \cdots, V^n$ which is transverse to $T^X(\cdot)p$. Define $t(\tau)$ as that value t for which $\psi(t) \in M^{n-1}(T^X(\tau)p)$. Choose an ordering, say $V^n, \cdots, V^2$; then $\exists$ smooth functions $s_2(\tau), \cdots, s_n(\tau)$ such that $\psi(t(\tau)) \equiv \varphi(\tau) \equiv T^{V^n}(s_n(\tau)) \circ \cdots \circ T^{V^2}(s_2(\tau)) \circ T^X(\tau)p$. This reparametrization is as required. Indeed,

$$\langle Z(\varphi(\tau)), \varphi'(\tau)\rangle = \langle Z(\varphi(\tau)), DT^{V^n}(s_n(\tau)) \cdots DT^{V^2}(s_2(\tau)) \cdot DT^X(\tau)X(p)\rangle.$$

Now let $g(s, \tau, p) = T^{V^n}(s_n) \circ \cdots \circ T^{V^2}(s_2) \circ T^X(\tau)p$. For fixed s, τ in a nbd. of $0 \in \boldsymbol{R}^{n-1}$, $g(s, \tau, \cdot)$: $M \to M$ diffeomorphically. We let $g_*(s, \tau, p)$ be the induced tangent space isomorphism. Then $\langle Z(\varphi(\tau)), \varphi'(\tau)\rangle = \langle Z(g(s(\tau), \tau, p)), g_*(s(\tau), \tau, p)X(p)\rangle$. Define

$$H(s, \tau) = \langle Z(g(s, \tau, p)), g_*(s, \tau, p)X(p)\rangle.$$

Then H is real valued, analytic for s, τ in a nbd. of $0 \in \boldsymbol{R}^n$; also $H(0, \tau) \equiv 1$. One may show that $H(s, \tau) \equiv 1$ in a nbd. of zero implies dim $L(\mathscr{S}^0)_p < n$; also $H(s, \tau) \leqq 1$ (or $H(s, \tau) \geqq 1$) for all $0 \leqq \tau \leqq t_1$ and s in a nbd. of $0 \in \boldsymbol{R}^{n-1}$ implies $T^X(t)p \in \partial\mathscr{A}(t, p, \mathscr{E})$ for $0 \leqq t \leqq t_1$.

Assumption (a.1) implies that, for each $\tau \in [0, t_1]$, $\mathscr{A}(\tau, p, \mathscr{E})$ *contains* an $(n-1)$-manifold through $T^X(\tau)p$, transverse to T^X, and having $T^X(\tau)p$ as an interior point. (Indeed span $\mathscr{S}^0_{T^X(\tau)p}$ is the tangent space to such an $(n-1)$-manifold at $T^X(\tau)p$.) Assumption (a.1) also implies the existence of "comparison solutions" in a full nbd. of $\{T^X(\tau)p\colon 0 < \tau < t_1\}$. If $H(s, \tau) \not\equiv 1$ one can find comparison solutions whose values at time t_1 cover a set of the form $\{T^X(t_1 + \sigma)\colon \sigma \in [\alpha, \beta]\}$ where $H(s, \tau) \geqq 1 \Rightarrow \alpha < 0, \beta = 0$; $H(s, \tau) \leqq 1 \Rightarrow \alpha = 0, \beta > 0$ and $H(s, \tau) - 1$ changing sign $\Rightarrow \alpha < 0, \beta > 0$. Thus $H(s, \tau) \not\equiv 1$ is sufficient that $\mathscr{A}(t, p, \mathscr{E})$ have nonempty interior while $H(s, \tau) - 1$ changing sign $\Rightarrow T^X(t)p \in$ int $\mathscr{A}(t, p, \mathscr{E})$. An expansion of H is difficult; however one may proceed as follows. Let $f(s)q = T^{V^n}(s_n)\circ\cdots\circ T^{V^1}(s_2)q$ and define $h(s, \tau) = \langle Z(T^X(\tau)p), f_*^{-1} X(g(s, \tau, p))\rangle$. Then $H(s, \tau) = \gamma \Leftrightarrow h(s, \tau) = 1/\gamma$ showing $h - 1$ changes sign if and only if $H - 1$ does. We next show $T^X(\cdot)p$ is singular iff rank $\mathscr{S}^0_p = n - 1$. Thus T^X nonsingular $\Leftrightarrow \varphi_{1j}(s) \not\equiv 0 \Leftrightarrow$ rank $\mathscr{S}^0_p = n$ and the conclusion of Theorem 1 easily follows. If T^X is singular, $(d/dt)Z(T^X(t)p) = -Z(T^X(t)p)X_x(T^X(t)p)$ and $h(s,\tau) = \langle Z(p), g_*^{-1} X(g(s, \tau, p))\rangle$. Using the Campbell-Hausdorff formula for this case, i.e., $\varphi_{1j}(s) \equiv 0$,

$$h(s, \tau) = 1 + \sum_{|\nu|=2}^{\infty} (-s_n^{\nu_n}/\nu_n!) \cdots (-s_2^{\nu_2}/\nu_2!) \sum_{j=0}^{\infty} (\tau^j/j!)a(\nu, j).$$

Now, specifically, $h(s, \tau) - 1$ "changing sign" in the above discussion means that for any $\varepsilon > 0$ it is necessary and sufficient that $h(s, \tau) - 1$ change sign as a function of s, at $s = 0$, for some $\tau \in [0, \varepsilon)$. Thus the role of τ differs from that of s. Theorem 1 follows from the expansion of h.

REMARK 2. Consider the system $\mathscr{D} = \{X + \alpha Y\colon -1 \leqq \alpha \leqq 1\}$. If $\exists$ vector fields $V^3, \cdots, V^n$ such that $X(p)$, $Y(p)$, $V^3(p)$, $\cdots$, $V^n(p)$ are linearly independent while Y, $V^3, \cdots, V^n$ are involutive, one may expand $\mathscr{D}$ to a system of the form $\mathscr{E}$. Then a sufficient condition that $T^X(t)p \in \partial\mathscr{A}(t, p, \mathscr{E})$ is also sufficient that $T^X(t)p \in \partial\mathscr{A}(t, p, \mathscr{D})$. Such a condition is $H(s, \tau) \geqq 1$ (or $H(s, \tau) \leqq 1$) for s in a nbd. of zero, $0 \leqq \tau \leqq t_1$. Equivalently, from Theorem 1, if some $\varphi_{mj}(s) \not\equiv 0$ and $\varphi_{m^*j^*}(s)$ is definite, $T^X(t)p \in \partial\mathscr{A}(t, p, \mathscr{D})$ for sufficiently small $t > 0$.

The existence of $V^3, \cdots, V^n$ is equivalent to the existence of a foliation of co-dim 1 for a nbd. of $\{T^X(t)p\colon 0 \leqq t \leqq t_1\}$ such that the leaf through $T^X(\tau)p$ is transverse to T^X and contains the trajectory $T^Y(\cdot)\circ T^X(\tau)p$. Let $\mathscr{S}^0 = \{Y, (\text{ad } X, Y), (\text{ad}^2 X, Y), \cdots\}$ If rank $(X \cup \mathscr{S}^0)_p = n$ the existence of $V^3, \cdots, V^n$ is assured by a construction shown to me by Professor A. J. Krener. In this case, we also have $H(s, \tau) \not\equiv 1$ implies int $\mathscr{A}(t, p, \mathscr{D}) \neq \emptyset, \forall\, t > 0$.

2. Examples.

EXAMPLE 1. Let $X(x) = (8, x_2, x_2)$, $V^2(x) = (x_2^2, 1, 0)$, $V^3(x) = (0, 0, 1)$ and $p = 0$. Then $X(p)$, $V^2(p)$, $V^3(p)$ are linearly independent; V^2, V^3 commute, hence are

involutive. We compute $Z(x) = (1/8 - x_2^2)(1, -x_2^2, 0)$; $(\mathrm{ad}^j X, [V^2, X])(x) = (2(-2)^j x_2^2, -1, -1)$, $(\mathrm{ad}^j X, [V^3, X])(x) = 0$. Then $a(1, 0, j) = a(0, 1, j) = 0 \ \forall\, j \Rightarrow \varphi_{1j}(s) \equiv 0$; hence $T^X(t)p$ is singular. Since $[V^3, X] = 0$, $a(\nu_2, \nu_3, j) = 0$ if $\nu_3 \neq 0$. Also, $(\mathrm{ad}^j X, (\mathrm{ad}^2 V^2, X))(x) = ((-1)^j 6x_2, 0, 0)$ showing $a(2, 0, j) = 0 \ \forall j$. Thus $\varphi_{2j}(s) = 0 \ \forall\, j$. However $(\mathrm{ad}^3 V^2, X)(x) = (6, 0, 0) \Rightarrow a(3, 0, 0) = 2/3 \Rightarrow \varphi_{3,0}(s) = (2/3)s_2^3$. Here $m^* = 3$ is odd $\Rightarrow T^X(t)p \in \operatorname{int} \mathscr{A}(t, p, \mathscr{E}) \ \forall\, t > 0$.

EXAMPLE 2. Let X, V^3, p be as in Example 1 but now $V^2(x) = (x_2, 1, 0)$. Again $T^X(\cdot)p$ is singular. Now $Z(x) = (1/8 - x_2)(1, -x_2, 0)$, $\varphi_{1j}(s) \equiv 0 \ \forall\, j$; $a(\nu_2, \nu_3, j) = 0$ if $\nu_3 \neq 0$ but $(\mathrm{ad}^2 V^2, X)(x) = (-2, 0, 0) \Rightarrow a(2, 0, 0) = -1/4$ and $\varphi_{2,0}(s) = (-\frac{1}{4})s_2^2 \Rightarrow T^X(t_1)p \in \partial\mathscr{A}(t_1, p, \mathscr{E})$ for sufficiently small $t_1 > 0$. For such t_1, $h(s, \tau) < 1$ if $0 < \tau < t_1$ and $|s|$ is sufficiently small $\Rightarrow H(s, \tau) > 1 \Rightarrow T^X(\cdot)p$ minimizes transfer time from p to $T^X(t_1)p$.

EXAMPLE 3. Let $X(x) = (4, x_2, 0)$, $V^2(x) = (x_2, 1, x_2x_3)$, $V^3(x) = (0, 0, 1)$, $p = 0$. Conditions (a.1), (a.2) are satisfied. Computing gives $Z(p) = (1/4, 0, 0)$, $(\mathrm{ad}^j X, V^2)(x) = ((-x_2)^j, 1, x_2x_3)$, $(\mathrm{ad}^j X, V^3)(x) = 0$. Then $a(1, 0, j) = a(0, 1, j) = 0 \ \forall\, j \Rightarrow \varphi_{1j}(x) \equiv 0 \Rightarrow T^X(\cdot)p$ is singular. Next, $(\mathrm{ad}^2 V^2, X)(x) = (-2, 0, 0)$, $(\mathrm{ad} V^2, (\mathrm{ad} V^3, X))(p) = 0$, $(\mathrm{ad}^2 V^3, X)(p) = 0$ so $a(2, 0, 0) = -\frac{1}{2}$, $a(1, 1, 0) = a(0, 2, 0) = 0$. Thus $\varphi_{20}(s) = -s_2^2/4$ and $T^X(\cdot)p$ minimizes transfer time from p to $p^1 = T^X(t_1)p$ for sufficiently small $t_1 > 0$.

References

1. H. J. Sussmann and V. Jurdjevic, *Controllability of non-linear systems*, J. Differential Equations **12** (1972), 95–116.

2. A. J. Krener, *A generalization of Chow's theorem and the Bang-Bang theorem to nonlinear control systems*, SIAM J. Control **12** (1974), 43–52.

3. H. Hermes, *On local and global controllability*, SIAM J. Control **12** (1974), 252–261.

4. ———, *On necessary and sufficient conditions for local controllability along a reference trajectory*, Geometric Methods in System Theory, D. Q. Mayne and R. Brockett (editors), Reidel, Dordrecht, 1974.

5. A. J. Krener, *The high order maximal principle*, Geometric Methods in System Theory, D. Q. Mayne and R. Brockett (editors), Reidel, Dordrecht, 1974.

6. H. J. Kelley, R. E. Kopp and H. G. Moyer, *Singular extremals*, Topics in Optimization, Academic Press, New York, 1967, pp. 63–101. MR **35** #5996.

UNIVERSITY OF COLORADO
BOULDER, COLORADO 80302, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

A Finite Difference Method for Computing Optimal Stochastic Controls and Costs

Harold J. Kushner*

1. Introduction. Let U denote a compact set in some Euclidean space R^q, w_t a standard R^m-valued Wiener process on $[0, \infty)$, and let $f(\cdot,\cdot)$: $R^r \times U \to R^r$, $\sigma(\cdot)$: $R^r \to r \times r$ matrix valued, $k(\cdot,\cdot)$: $R^r \times U \to R$, $\phi(\cdot)$: $R^r \to R$ where the latter four functions are continuous and bounded and $\{f(x, U), k(x, U)\}$ is compact and convex for each $x \in R^r$. Suppose that G is a bounded open set. We will consider an interesting computational method for the control system governed by the Ito equation[1]

$$dx_s = f(x_s, u(x_s))ds + \sigma(x_s)dw_s, \qquad u(x) \in U. \tag{1}$$

If a solution to (1) exists up until $\tau(u) \equiv \min\{t: x_t \in \partial G, u(\cdot) \text{ used}\}$, and if $E_x^u\tau(u) < \infty$, define the cost functional

$$W(u, x) = E_x^u \int_0^{\tau(u)} k(x_s, u(x_s))\, ds + E_x^u\phi(x_{\tau(u)}), \tag{2}$$

and $W(x) = \inf_u W(u, x)$, where the infimum is over all U-valued $u(\cdot)$ of the above type. We can also consider the discounted case and the case where control is stopped at a finite time T.

Formally, if $W(\cdot)$ is smooth enough, Ito's lemma (where we abuse the notation u) implies that

$$\inf_{u \in U}[\mathscr{L}^u V(x) + k(x, u)] = 0, \qquad x \in G,\ V(x) = \phi(x),\ x \in \partial G \tag{3}$$

*This research was supported in part by the Office of Naval Research NONR N00014-67-A-0191-001804, by Air Force Office of Scientific Research AF-AFOSR 71-2078C and the National Science Foundation GK-40493X.

[1]Under an additional condition, σ could also depend on the control.

(where $\mathscr{L}^u$ is the differential generator of (1)). Conversely (Fleming [1], Rishel [2]), a "smooth" solution to (3) yields $W(\cdot)$ and the optimal control $\bar{u}(\cdot)$. Generally, $\mathscr{L}^u$ is degenerate, and suitable smoothness of the $V(\cdot)$ in (3) cannot be guaranteed, and, indeed, the use of (3) for computational purposes is in question.

Our plan here is the following. Finite difference (FD) approximations to (3) will be investigated. With a careful, but natural, choice of the FD approximations (difference interval h), the FD equations are the backward equations for the optimal control of a certain family of Markov chains. It turns out that this family of chains, when suitably interpolated to get a family of continuous parameter processes, converges "weakly" to a controlled diffusion (when the sequence of measures corresponding to a family of processes on some space converges weakly, we say that the processes converge weakly). Under some additional conditions, we can get an approximation to the optimal control for (1) and to the optimal costs. Such results are not frequently available, for our class of problems, by use of other current techniques. Space limitations prevent full development of the ideas of the physical intuition which the probabilistic approach provides. The probabilistic interpretation is often helpful in comparing and developing algorithms and interpreting numerical results. Only a brief and heuristic outline will be given. More details are in Kushner and Yu [3] or Kushner [4], or will be published elsewhere, and much still remains to be done.

2. Finite difference approximations. Define $a(\cdot) = \sigma(\cdot)\sigma'(\cdot)/2$. Purely for simplicity, let $a(\cdot)$ be diagonal. For each real $h > 0$, let R_h^r denote the FD grid on R^r with interval h, and write $G_h = G \cap R_h^r$. Define ∂G as the set of points on $R_h^r - G_h$ which are only one node removed from G_h. Let e_i denote the unit coordinate in the ith (orthogonal) direction of R^r. For reasons to be made clear below, choose the special finite difference approximations

$$V_{x_i}(x) \to \begin{cases} [V(x + e_ih) - V(x)]/h, & \text{if } f_i(x, u(x)) \geqq 0, \\ [V(x) - V(x - e_ih)]/h, & \text{if } f_i(x, u(x)) < 0, \end{cases} \tag{4a}$$

$$V_{x_ix_i}(x) \to [V(x + e_ih) + V(x - e_ih) - 2V(x)]/h^2. \tag{4b}$$

Define $Q_h(x, u) = \sum_i [a_{ii}(x) + h|f_i(x, u)|]$, and let $\inf_{x,u\in U} Q_h(x, u) \geqq Kh$ for some positive real K. Substituting (4) into (3), collecting terms and denoting the finite difference solution by $V^h(\cdot)$ yields (where $f_i^+ = \max(f_i, 0)$, etc.)

$$0 = \inf_{u\in U}\left[- V^h(x)Q_h(x,u) + \sum_{i,\pm} V^h(x \pm e_ih)[a_{ii}(x) + hf_i^{\pm}(x, u)] + h^2k(x, u)\right] \tag{5}$$

for $x \in G_h$, and $V^h(x) = \phi(x)$ for $x \in \partial G_h$.

At this point there are several directions that we can go in, but only one of them will be pursued. Define $\bar{Q}_h(x) = \sup_{u\in U} Q_h(x, u)$. Then add $\bar{Q}_h(x)V(s)$ to both sides of (5), divide (5) by $\bar{Q}_h(x)$, and define (for $x \in G_h$)

$$\Delta t^h(x) = h^2/\bar{Q}_h(x), \qquad p^h(x, x \pm e_ih|u) = [a_{ii}(x) - hf^{\pm}(x, u)]/\bar{Q}_h(x),$$

$p^h(x, x|u) = [\bar{Q}_h(x) - Q_h(x, u)]/\bar{Q}_h(x)$ and $p^h(x, x|u) = 1$ for $x \notin G_h$. For all other

(x, y), let $p^h(x, y) = 0$. Then[2] for $x \in G_h$,

$$(6)\quad V^h(x) = \inf_{u\in U}\Big[V^h(x)p^h(x, x|u) + \sum_{i,\pm} V^h(x \pm e_i h)p^h(x, x \pm e_i h|u) + k(x, u)\Delta t^h(x)\Big]$$

and $V^h(x) = \phi(x)$, $x \in \partial G_h$.

Now we are at the key point for the future development. The $p^h(\cdot, \cdot|u)$ are transition probabilities for a controlled Markov chain, which is absorbed on first exit from G_h. This Markov chain interpretation can be very helpful. Let $\{\xi_n^h\}$ denote the random variables of the chain, and define $N_h(u) = \min\{n: \xi_n^h \in \partial G_h, u \text{ used}\}$, and $\Delta t_i^h = \Delta t^h(\xi_i^h)$ and $t_i^h = \sum_{s=0}^{i-1} \Delta t_s^h$, and $\tau_h(u) = t_{N_h(u)}^h$. To facilitate the discussion, let us make an assumption:

(A1) for some i, $\inf_{x\in G}|a_{ii}(x)| > 0$.

(A1) is not necessary, but it does imply that $\sup_{u,x\in G_h} E_x N_h(u) < \infty$, and $\sup_{u,x\in G_h} E_x^u[\tau_h(u)]^k < \infty$, all $k > 0$. Define the process $\xi^h(t)$, $t \in [0, \infty)$, by $\xi^h(0) = \xi_0^h = x$, $\xi^h(s) = \xi_i^h$, $t \in [t_i^h, t_{i+1}^h)$. Then $\tau_h(u)$ is the escape time for $\xi_h(\cdot)$.

Now (6) is the value equation for optimal control of the $\{\xi_n^h\}$ with cost rate $k(x, u)\Delta t^h(x)$. Under (A1) and the prior conditions, it has a unique solution, which is

$$(7)\quad \begin{aligned} V^h(x) &= \inf_u E_x^u\Big[\sum_{n=0}^{N_h(u)-1} k(\xi_n^h, u(\xi_n^h))\Delta t_n^h + \phi(\xi_{N_h(u)}^h)\Big] \\ &= \inf_u E_x^u\Big[\int_0^{\tau_h(u)} k(\xi^h(s), u(\xi^h(s)))\, ds + \phi(\xi^h(\tau_h(u)))\Big]. \end{aligned}$$

$V^h(x)$ is no larger than $W^h(\tilde{u}, x)$, the cost due to a policy $\tilde{u}(\cdot)$ for which $E_x^{\tilde{u}} N_h(\tilde{u}) < \infty$. Let $u^h(\cdot)$ denote the optimal policy (defined in G_h).

We have $E_x^u[\xi_{n+1}^h - \xi_n^h|\ \xi_n^h = x] = f(x, u(x))\Delta t^h(x)$,

$$\operatorname{cov}[\xi_{n+1}^h - \xi_n^h|\ \xi_n^h = x,\ u \text{ used}] = 2a(x)\Delta t^h(x) + o(\Delta t^h(x))$$

which, together with the $\Delta t^h(x)$ coefficient on $k(x)$, further motivates our continuous time interpolation from $\{\xi_n^h\}$ to $\xi^h(\cdot)$. Write

$$\xi_{n+1}^h = \xi_n^h + f(\xi_n^h, u(\xi_n^h))\Delta t_n^h + \beta_n^h,$$

where $\{\beta_n^h\}$ is orthogonal.

3. Limits of $\{\xi^h(\cdot)\}$. Let D denote the space of R^r-valued functions on $[0, \infty)$ which are right continuous and have left-hand limits, and which is endowed with the Skorokhod topology (Billingsley [5, pp. 111–118]). The topology is metrizable, and is complete and separable with respect to that metric. Define $\tau^h(\cdot)$ as the piecewise constant process (on each Δt_i^h) with jumps $\Delta t_i^h I_{\{\xi_i^h \in G_h\}}$. Thus $\tau^h(\infty) = \tau_h$. Then $\{\xi^h(\cdot), \tau^h(\cdot)\}$ (for any sequence $\{\tilde{u}^h(\cdot)\}$) is a tight sequence, and each subsequence has a weakly convergent subsequence, i.e., for each subsequence there is a further subsequence and a process $x(\cdot)$, depending on the sequence and with paths in D

[2] It is sometimes better merely to divide (5) by $Q_h(x, u)$ (ignoring its u-dependence when doing that). Then p^h and Δt^h in (6) have Q_h rather than $\bar{Q}_h$ in the denominator. Under reasonable conditions both equations have the same solution, as $h \to 0$, and the latter sometimes has some computational advantages.

w.p.l., which is the weak limit. The limits are continuous w.p.l. Indeed, using a lemma of Skorokhod [6] (and also Dudley [8]), we can and will suppose (by properly choosing the underlying probability space) that all convergences are w.p.l. in the Skorokhod topology, hence uniform on any finite interval w.p.l.

Until further notice, let us fix a weakly convergent subsequence (corresponding to $\{u^h(\cdot)\}$)—denote it by $\{\xi^h(\cdot), \tau^h(\cdot)\}$ and the limit by $\{x(\cdot), \bar{\tau}(\cdot)\}$.

Write $\bar{\tau}(\infty) = \bar{\tau}$. $x(\cdot)$ is absorbed on ∂G at $\bar{\tau}$ (not necessarily its first hitting time) and (w.p.l.) $\tau^h(u^h) \to \bar{\tau} \geqq \inf\{t\colon x_t \in \partial G\}$. Also by (A1), $E_x \bar{\tau}^k < \infty$, for any $k > 0$. It turns out that $x(\cdot)$ is a controlled diffusion until $\bar{\tau}$. In particular, using the techniques of [3], [4] and the ideas of [7] concerning the nonanticipative form of the control, there is a Wiener process $w(\cdot)$, a nonanticipative control $\tilde{u}(\cdot,\cdot)$ with values $\tilde{u}(s, x(\cdot)) \in U$, for which ($x(\cdot)$ is also nonanticipative with respect to $w(\cdot)$)

$$x_t = x + \int_0^{t\cap\bar{\tau}} f(x_s u(s, x(\cdot))\, ds + \int_0^{t\cap\bar{\tau}} \sigma(x_s)\, dw_s, \tag{8}$$

$$E_x^{u^h}\left[\int_0^{\tau_h(u^h)} k(\xi^h(s), u^h(\xi^h(s)))\, ds + \phi(\xi^h(\tau_h(u^h)))\right] \to E_x^{\tilde{u}}\int_0^{\bar{\tau}} k(x_s, \tilde{u}(s, x(\cdot)))\, ds + \phi(x_{\bar{\tau}}).$$

Now, $x_t \in \bar{G}$, $t \leqq \bar{\tau}$, if $x_0 = x \in G$. If $\bar{\tau} \neq \tau$, at some ω then the path $x(\cdot)$ is tangent to ∂G at $t = \tau(\tilde{u})$: It hits G and does not leave for some positive time. Many conditions can be invoked to eliminate this possibility. Here we only assume

(A2) $x(\cdot)$ crosses ∂G w.p.l. at $t = \tau(\tilde{u})$.

Then $\tau = \bar{\tau}$ w.p.l., and $V^h(x) = W^h(x, u^h) \to W(x, u)$. (A2) can often be verified from the general form of $f(\cdot,\cdot)$, $\sigma(\cdot)$ on the boundary, without knowledge of $\tilde{u}(\cdot,\cdot)$.

A control $\hat{u}(\cdot)$ with values in U, and for which (1) has a unique solution up to $\tau(u)$, is said to be *discretizable* if there exists a sequence of policies $\hat{u}^h$ on G_h for which $\xi^h(\cdot)$ converges weakly to $x(\cdot)$ (corresponding to $\hat{u}(\cdot)$) on $[0,\tau(\hat{u}))$ and $E_x^{\hat{u}}\tau^h(\hat{u}^h) \to E_x^{\hat{u}}\tau(\hat{u}^h)$ and $\{\tau^h(\hat{u}^h)\}$ is uniformly integrable. (The latter condition is implied by (A1).) $W(x, \tilde{u}) \leqq W(x, \hat{u})$ for all discretizable $\hat{u}$. Suppose

(A3) There is an optimal control $\bar{u}(\cdot)$ which depends only on x_s at time s and is discretizable (with discretizations $\bar{\mu}^h(\cdot)$).

Then $W^h(x, \bar{u}^h) \geqq W^h(x, u^h)$. Since $V^h(x) = W^h(x, u^h) \to W(x, \tilde{u})$ and $W^h(x,\bar{u}^h) \to W(x, \bar{u}) = V(x) \leqq W(x, \tilde{u})$, we have that $V^h(x) \to V(x)$ as $h \to 0$.

(A1) guarantees that $\sup_{x,h} E_x\tau_h < \infty$ and that $\{\tau_h\}$ is uniformly integrable. This is essential for the desired convergence. If $\tau_h < \infty$ w.p.l., and we can prove that $\bar{\tau} = \tau$, then the uniform integrability condition can be eliminated by adding a discount factor. Let $\alpha(\cdot)$ be a bounded continuous nonnegative real-valued function on R^r. Adding the discount factor $\exp\{-\int_0^{\cdot}\alpha(x_s)\, ds\}$ to the cost implies the addition of $-\alpha(x)V(x)$ to the l.h.s. of (3). If $\inf_x \alpha(x) > 0$, and we are able to prove that $\tau_h(u^h) \to \tau(\tilde{u})$ w.p.l., then we do not need uniform integrability of $\{\tau_h(u^h)\}$, since, due to the discounting, the effective "extinction" times are uniformly bounded. Define τ for $x \in G$ by $\tau = \inf\{t\colon t > 0,\ x(t) \notin \bar{G}\}$. Then uniform integrability is implied by (A2) and either $\sup_{x,u\in\bar{G}} E_x^u\tau < \infty$ or, if $k(x, t, u) \geqq k_0 > 0$, by $\sup_{x\in\bar{G}} E_x^u\tau < \infty$ for some u satisfying (A3).

5. The optimal control. We are not able to prove a priori "smoothness" of the $V^h(\cdot)$, uniform in h. Thus, we cannot, in general, assert that there is a subsequence of the $u^h(\cdot)$ which converges pointwise to the optimal control. This is a serious difficulty, since a main object of the numerical solutions is to obtain a useful approximation to $\bar{u}(\cdot)$. However, in typical numerical calculations, the $u^h(\cdot)$ seem to converge to a smooth function pointwise, or to a discontinuous function where (in typical cases) the paths of the limiting process have "negligible" probability of being "near" the discontinuities at any time t. If this apparent convergence does take place, then the limit is indeed at least (under (A1)—(A2)) optimal with respect to any discretizable control.

The ideas can be extended to control problems with reflections on a part of a boundary, and to problems with many types of constraints of both the path and the expectation types.

6. Remaining problems. It would be very useful to have prior estimates on the smoothness of the $V^h(x)$. Certainly much needs to be done to understand fully the properties that the $\{\xi_n^h, u(\cdot)$ used$\}$ derive from the properties of $\{x(\cdot), u(\cdot)$ used$\}$, particularly on the behavior near the boundary and on the values of the exit times. An investigation of various finite difference and variable grid techniques (yielding weakly convergent sequences with the proper limit) would be helpful—for improving the numerical efficiency of the method.

References

1. W. H. Fleming, *Some Markovian optimization problems*, J. Math. Mech. **12** (1963), 131–140. MR **24** #1636.

2. R. W. Rishel, *Weak solutions of a partial differential equation of dynamic programming*, SIAM J. Control **9** (1971), 519–528. MR **47** #8187.

3. H. J. Kushner and C. F. Yu, *Approximations, existence and numerical procedures for optimal stochastic controls*, J. Math. Anal. Appl. **45** (1974), 563–587.

4. H. J. Kushner, *Convergence of finite difference approximations of a discounted problem for a diffusion*, Brown University Technical Report, CDS TR 74–3.

5. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR **38** #1718.

6. A. V. Skorokhod, *Limit theorems for stochastic processes*, Teor. Verojatnost. i Primenen. **1** (1956), 289–319 = Theor. Probability Appl. **1** (1956), 261–290. MR **18**, 943.

7. H. J. Kushner, *Existence of optimal stochastic controls*, J. Opt. Theory and Appl. **15** (1975), 347–359; Brown University, Technical Report, CDS, 74–3.

8. R. M. Dudley, *Distances of probability measures and random variables*, Ann. Math. Statist. **39** (1968), 1563–1572. MR **37** #5900.

BROWN UNIVERSITY
PROVIDENCE, RHODE ISLAND 02912, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Controllability in Topological Dynamics

Lawrence Markus

1. Global controllability for dynamic polysystems. Global periodicity and transitivity theorems are proved for control dynamical systems, say $dx/dt = f(x, u)$, in terms of an abstraction of topological dynamics. Here the state x lies in a given topological space M (the state or phase space), and the controllers $u(t)$ are piecewise constant functions with values, say $u_1, u_2, \cdots, u_p$ in a suitable restraint set Ω for finite durations $t_1, t_2, \cdots, t_p$.

Instead of actions of the real time line R on M consider, more generally, any collection of topological groups G_u, as indexed by a control parameter u in Ω, each acting on M as a topological dynamical system

$$\Phi_u : M \times G_u \to M : (x, g) \to x_g = xg.$$

We seek to control the orbit of a point $x \in M$ by commands, each of which is described by a word (say, abracadabra!)

$$w = (g_{u_1}, u_1)(g_{u_2}, u_2)\cdots(g_{u_p}, u_p) \qquad \text{(finite } p \text{ depending on } w)$$

with g_{u_i} in the group G_{u_i} (with repetitions allowed). Here the command w acts on $x \in M$ by

$$x \to xw = ((xg_{u_1})g_{u_2}) \cdots g_{u_p}.$$

We define the command group $\mathcal{G}$ using concatenation of reduced words (combining adjacent products from the same group G_u and omitting any factor containing the identity e_u of G_u), and consider the action of $\mathcal{G}$, the free product of $\{G_u\}$, on the space M; see [3].

DEFINITION. A topological dynamical polysystem

$$\Phi : M \times \mathcal{G} \to M : (x, w) \to xw = xg_{u_1}g_{u_2}\cdots g_{u_p},$$

as constructed from a family of dynamical systems Φ_u: $M \times G_u \to M$ for $u \in \Omega$,

is defined by the command group $\mathcal{G}$ of reduced words acting on M through concatenation.

We give $\mathcal{G}$ the strongest topology for which all maps $G_{u_1} \times G_{u_2} \times \cdots \times G_{u_p} \to \mathcal{G}$: $(g_{u_1}, \cdots, g_{u_p}) \to w$ are continuous. Then the map of $\mathcal{G}$ onto the orbit $x\mathcal{G} = \{xw | w \in \mathcal{G}\}$ given by $w \to xw$ is continuous.

Standing hypotheses. In this paper we consider only dynamical polysystems Φ for which:

(i) Each group G_u is a copy R_u of the real number line R.

(ii) The phase-space M is a compact metric space.

(iii) The control restraint set Ω is a compact neighbourhood of the origin $u = 0$ in the real number space R^m.

(iv) There exists a continuous map $\Psi : M \times R \times \Omega \to M$ which restricts to the dynamical system Φ_u for each fixed $u \in \Omega$. That is, for each $u_0 \in \Omega$, Ψ on $M \times R \times u_0$ is the map Φ_{u_0} on $M \times R$.

The dynamical system Φ_0 (where $u = 0$) is called the free system and the other Φ_u are the controlled systems.

REMARK. The orbit of $x \in M$ under the command group $\mathcal{G}$ consists of all points $xw = x(t_1, u_1)\,(t_2, u_2)\cdots(t_p, u_p)$ where $x(t_1, u_1) = \Phi_{u_1}(x, t_1)$ is obtained by the action of $G_{u_1} = R_{u_1}$ on M for a time duration t_1, and the other actions hold thereafter for durations $t_2, \cdots, t_p$. Thus the orbit of x consists of all points of M to which x can be steered by piecewise constant controllers constructed by a concatenation of actions of $\Phi_{u_1}, \cdots, \Phi_{u_p}$ for durations $t_1, t_2, \cdots, t_p$ (finite p depending on w).

In case all these $t_j \in R_+$, that is, $t_j \geqq 0$ for $j = 1, \cdots, p$, we declare that xw lies on the *future* orbit $x\mathcal{G}_+$ in M, and the past orbit is defined in a similar way. We emphasize that our results on controllability involve only future orbits. In this case each command word $w = (t_1, u_1) \cdots (t_p, u_p)$ all $t_j \in R_+$ can be specified by a piecewise constant controller on future times, $u(t) = u_j$ on $t_0 + t_1 + \cdots + t_{j-1} \leqq t \leqq t_0 + t_1 + \cdots + t_j$ (usually $t_0 = 0$). The corresponding response or trajectory of the point x_0 on $t_0 \leqq t \leqq t_0 + \cdots + t_p$ is defined in the usual way and is denoted by $\phi([u], x_0, t_0; t)$ with $\phi([u], x_0, t_0; t_0) = x_0$.

DEFINITION. Consider the topological dynamical polysystem $\Phi : M \times \mathcal{G} \to M$, as constructed from a family of dynamical systems $\Phi_u : M \times G_u \to M$, with $G_u = R_u$ for each $u \in \Omega$, according to the standing hypotheses. For each $x_0 \in M$ and time $\tau > 0$ the set of attainability is defined by

$$A(x_0; \tau) = \{x_0 w | w = (t_1, u_1) \cdots (t_p, u_p) \text{ with all } t_j \geqq 0 \text{ and } t_1 + t_2 + \cdots + t_p = \tau\}.$$

Also Φ is locally controllable along free trajectories, or Φ satisfies the uniform controllability hypothesis, in case

(uc) for each $\tau > 0$ there exists a radius $\eta = \eta(\tau) > 0$ such that each set of attainability $A(x_0; \tau)$ contains an open η-ball centred at the endpoint of the free trajectory point $\Phi_0(x_0; \tau) = \phi([0], x_0, 0; \tau)$.

EXAMPLE. Consider a control differential system $\dot{x} = f(x, u)$ for the state x in a compact Riemannian n-manifold M (without boundary and covered by C^∞ local coordinate charts), and the control u in a compact neighbourhood Ω of the origin in R^m. More precisely, $f(x, u)$ is a C^∞ map into the tangent bundle TM, $f: M \times \Omega$

$\to TM$, which is a cross-section over M for each fixed $u \in \Omega$. Hence, for each $u_0 \in \Omega$, the tangent vector field $f(x, u_0)$ is globally defined on all M, with solutions $\phi(u_0, x_0, t_0; t)$ in $C^\infty(\Omega, M, R, R)$ initiating at $x_0 \in M$ when $t = t_0$, and generating the flow

$$\Phi_{u_0} : M \times R \to M : (x_0, t) \to \phi(u_0, x_0, 0; t).$$

Using the command group $\mathcal{G}$ of words $w = (t_1, u_1) \cdots (t_p, u_p)$, say, with values $u_1, \cdots, u_p \in \Omega$ for finite durations $t_1, \cdots, t_p$, we define the dynamical polysystem $\Phi : M \times \mathcal{G} \to M$. If all the durations $t_j \geqq 0$ in $w \in \mathcal{G}$, so w can be represented by a piecewise constant controller, we obtain points on the future orbit $x_0\mathcal{G}_+$ of the initial state $x_0 \in M$, and then the corresponding set of attainability $A(x_0; \tau)$.

The general study of the uniform controllability hypothesis for such control dynamical systems can be found in earlier papers [4], [6], [7], [8], and a general theory in [1]. Below we shall indicate that uniform controllability is a *generic* property for such control dynamical systems.

If the free system of the above example $\dot{x} = f(x, 0)$ on the compact Riemannian manifold M is conservative, that is $\operatorname{div}_x f(x, 0) \equiv 0$, then the flow Φ_0 is measure-preserving on M. In this case there is a dense set of points in M each of which has a future recurrent orbit under Φ_0 (that is, the orbit is contained in its future limit set). Other types of differential systems also have such a dense set of recurrent points, for example Anosov hyperbolic flows—but some types definitely do not, for example Morse gradient flows.

By assuming the uniform controllability hypothesis on Φ, and that the free system Φ_0 has a dense set of future recurrent points in M, we shall prove the global controllability conclusion $A(x_0; T) = M$ for some duration $T > 0$ independent of the initial state $x_0 \in M$. This result is an easy case, simplified for exposition, of a much more difficult analysis of controlled differential systems [6], [8], as we comment below.

LEMMA. *Let the dynamical polysystem $\Phi : M \times \mathcal{G} \to M$ be constructed from a family of dynamical systems Φ_u for $u \in \Omega$, according to the standing hypotheses. Assume*

(i) *the free dynamical system Φ_0 has a dense set of future recurrent points, and*

(ii) *the uniform controllability hypothesis holds for $\eta(\tau) > 0$.*

Then each initial state $x_0 \in M$ lies on a controlled periodic orbit. That is, there exists a piecewise constant controller $u(t)$ on some finite interval $0 \leqq t \leqq T_0$ (moreover with $T_0 \geqq 1$ an integral period of $u(t)$) such that $\phi([u], x_0, 0; T_0) = x_0$.

SKETCH OF PROOF. Given $x_0 \in M$ take a future recurrent point x_1 so near that the corresponding free trajectories $\phi([0], x_0, 0; t)$ and $\phi([0], x_1, 0; t)$ remain within a distance of $\frac{1}{3}\eta(1)$ for the duration $-3 \leqq t \leqq 3$. In unit time control x_0 to the point $\phi([0], x_1, 0; 1)$, and thereafter follow the free trajectory until it recurs back at some rational time near to $\phi([0], x_1, 0; -2)$, then remains within $\frac{1}{3}\eta(1)$. Then control the trajectory to join $\phi([0], x_0, 0; -1)$ and thence onwards to x_0 after a total duration that is rational. By repeating this circuit a finite number of times we arrive back at x_0 in an integral duration T_0.

THEOREM 1. *Let the dynamical polysystem $\Phi : M \times \mathcal{G} \to M$ be constructed from a family of dynamical systems Φ_u for $u \in \Omega$, according to the standing hypotheses. Assume*

(i) *the free dynamical system Φ_0 has a dense set of future recurrent points, and*

(ii) *the uniform controllability hypothesis holds for $\eta(\tau) > 0$.*

Then there exists a duration $T > 0$ such that, for each $x_0 \in M$, the set of attainability $A(x_0; T) = M$.

SKETCH OF PROOF. Consider the reachable set, over integral times $k \geqq 1$,

$$A(x_0) = \bigcup_{k=1}^{\infty} A(x_0; k).$$

In any neighbourhood of a boundary point z of $A(x_0)$ in M there is a free recurrent point z_1 in some $A(x_0; k_1)$. Using the methods of the lemma, we show that z is necessarily an interior point of $A(x_0)$, so $A(x_0) = M$.

Next use the compactness of M to select a finite set of initial and terminal "transfer stations" $a_1, \cdots, a_r$ and $b_1, \cdots, b_s$, respectively, so that each $A(x; 1)$ for $x \in M$ contains some a_i, and each $y \in M$ lies in some $A(b_j; 1)$. Then using controlled periodic orbits through b_j, to waste time if necessary, we find an integral duration T_1 such that $b_j \in A(a_i; T_1)$ for all a_i and b_j. Then take $T = T_1 + 2$ for the required duration.

These two qualitative results on controlled dynamical systems are easy theorems of topological dynamics, but they are the prototypes of more complicated theories, related to nonautonomous control differential equations. For example $\dot{x} = f(x, t, u)$ for x in a compact Riemannian manifold M, and u in a compact neighbourhood Ω of the origin in R^m for each time $t \in R$, with the tangent vector field $f(x, t, u)$ and its first partial derivatives assumed bounded in $M \times R \times \Omega$. In addition we assume

(c) conservative hypothesis, $\operatorname{div}_x f(x, t, 0) = 0$;

(uc) uniform controllability hypothesis for $\eta(\tau) > 0$

and also some type of time-periodicity of $f(x, t, u)$.

If $f(x, t + 1, 0) = f(x, t, 0)$ is periodic in t, then the free dynamical system can be interpreted as an autonomous flow on the product of M with a circle S^1. However if we assume Bohr almost periodicity,

(ap) $f(x, t, 0)$ is uniformly almost periodic in t for $x \in M$,

then the related autonomous flow occurs in $M \times H(f)$. Here $H(f)$ is the hull of $f(x, t, 0)$ under the shift translational flow in a certain Banach space, so $H(f)$ is a compact infinite-dimensional space. In spite of these difficulties the conclusions of the above theorems on control periodicity and transitivity have been proved. For details see [8].

2. Generic controllability. Consider linear control processes $\dot{x} = A(t)x + B(t)u$ for $x \in R^n$, $u \in R^m$ and with coefficient matrices $(A(t), B(t))$ in C^∞ on $0 \leqq t \leqq T$. A sufficient condition for controllability in R^n on $0 \leqq t \leqq T$, with piecewise constant controllers $u(t)$ on $0 \leqq t \leqq T$ (regardless of $T > 0$) is [1]:

$$\text{rank } [B, \Gamma B, \Gamma^2 B, \cdots, \Gamma^{k-1} B] = n \quad \text{at } t = 0, \text{ for some integer } k \geqq 1,$$

$\Gamma V(t) = -AV + \dot{V}$ defined as a differential operator.

Consider the complete metric space $\mathscr{A} \times \mathscr{B}$ of all such C^∞ coefficient matrices $(A(t), B(t))$, with the usual metric defining uniform convergence of arbitrarily high order derivatives on $0 \leqq t \leqq T$. It is easy to show that the collection of all controllable members contains an open dense subset $\mathscr{C} \subset \mathscr{A} \times \mathscr{B}$; see [3].

For the proof take $k = n$ and the sufficiency condition can then be expressed by the nonvanishing of a single real polynomial P in the entries of the matrices $A(0)$, $B(0)$, $\dot{B}(0)$, $\cdots$, $B^{(n-1)}(0)$ (say, P is the sum of squares of the $n \times n$ subdeterminants of the controllability matrix). Hence the condition $P \neq 0$ defines an open set $\mathscr{C} \subset \mathscr{A} \times \mathscr{B}$. But by perturbing the elements of $(A(t), B(t))$ by very small real polynomials, with suitably chosen derivatives at $t = 0$, we find that $\mathscr{C}$ is dense in $\mathscr{A} \times \mathscr{B}$.

This result was noted for autonomous linear systems by the author [2] several years ago, and a corresponding theorem was also proved for local controllability for nonlinear systems near the origin. We here extend our results for nonlinear systems to the study of local controllability along trajectories, or equally well, the property of uniform controllability.

Thus consider the complete metric space $\mathscr{S}$ of all control dynamical systems $\dot{x} = f(x, u)$ for x in a given compact Riemannian n-manifold M, and u in a fixed compact neighbourhood Ω of the origin in R^m. That is, $\mathscr{S}$ is the space of all C^∞ cross-sections $f\colon M \times \Omega \to TM$ with the usual C^∞ metric of uniform convergence of arbitrarily high order partial derivatives on $M \times \Omega$; see [5].

THEOREM 2. *Consider the complete metric space $\mathscr{S}$ of all C^∞ control dynamical systems $\dot{x} = f(x, u)$ for $x \in M$ and $u \in \Omega$, as above. Then there exists an open dense subset $\mathscr{C} \subset \mathscr{S}$ of systems that satisfy the uniform controllability hypothesis.*

SKETCH OF PROOF. First we obtain a sufficient condition that the vector field $f(x, u)$ satisfies the uniform controllability hypothesis (uc) on M. Take $x_0 \in M$ and consider the free trajectory $\phi_0(t)$ for a given finite interval $0 \leqq t \leqq \tau$.

Choose a family of controllers

$$u(t, \xi) = \xi^1 u_1(t) + \xi^2 u_2(t) + \cdots + \xi^n u_n(t) + \cdots + \xi^{2n+m} u_{2n+m}(t)$$

for $0 \leqq t \leqq \tau$ and $\xi \in R^{2n+m}$, and with piecewise constant controllers $u_j(t)$ in Ω to be specified later. The corresponding responses are $x = \phi(t, \xi)$ with $\phi(t, 0) = \phi_0(t)$ and $\phi(0, \xi) = x_0$. We require the C^∞ map $R^{2n+m} \to M\colon \xi \to \phi(\tau, \xi)$ to cover an open neighbourhood of the point $\phi(\tau, 0) = \phi_0(\tau)$ for ξ near the origin. This will be assured by the implicit function theorem provided

$$\operatorname{rank} (\partial\phi/\partial\xi)(\tau, \xi)\big|_{\xi=0} = n.$$

Let $Z(t) = (\partial\phi/\partial\xi)(t, 0)$ be a continuous $(n, 2n + m)$ matrix (say computing all derivatives in one tubular neighbourhood of $\phi_0(t)$ on M), and we must show that $Z(\tau)$ has rank n. But the jth column $z_j(t)$ of $Z(t)$ satisfies, except at the finite set of control discontinuities,

$$\dot{z}_j(t) = A(t)z_j + B(t)u_j(t), \qquad z_j(0) = 0$$

where $A(t) = (\partial f/\partial x)(\phi_0(t), 0)$ and $B(t) = (\partial f/\partial u)(\phi_0(t), 0)$ are C^∞ matrices, and $\partial u/\partial\xi^j = u_j(t)$. Hence, provided the piecewise constant controllers $u_j(t)$ are suitably selected, the controllability condition (say with $k = 2n + m$)

$$\text{rank } [B(0), \Gamma B(0), \Gamma^2 B(0), \cdots, \Gamma^{2n+m-1} B(0)] = n$$

guarantees that $Z(\tau)$ has rank n, and that $\dot{x} = f(x, u)$ is locally controllable along the free trajectory $\phi_0(t)$.

Now compute

$$B(0) = f_u(x_0, 0), \qquad \Gamma B(0) = -AB + \dot{B} = -f_x f_u + f_{ux} f$$

and so forth, with $\Gamma^s B(0)$ given as an n-vector P_s with each component a nontrivial real polynomial in f and its partial derivatives of order $\leqq s + 1$, all evaluated at $(x_0, 0)$.

In terms of the $(2n + m)$-jet $J^{(2n+m)}_{n+m,n} f$ we express the controllability condition

$$\text{rank } [f_u, -f_x f_u + f_{ux} f, P_2, \cdots, P_{2n+m-1}(f, \cdots, f^{(2n+m)})] = n.$$

But f fails to meet this condition only when $n + m + 1$ determinants (among others) vanish simultaneously

$$\det[P_0, P_1, \cdots, P_{n-1}] = 0, \ \det[P_1, P_2, \cdots, P_n] = 0, \ \cdots, \ \det[P_{n+m}, \cdots, P_{2n+m-1}] = 0.$$

This set of determinantal conditions defines an algebraic variety of codimension $\geqq n + m + 1$ in the jet space $J^{(s)}_{n+m,n}$, for each $s \geqq 2n + m$. By standard transversality theory the set $\mathscr{C}$ of control systems in $\mathscr{S}$, satisfying the sufficiency condition for controllability, is open and dense in $\mathscr{S}$, at least for restrictions of $\mathscr{S}$ to a fixed coordinate neighbourhood of x_0 in M.

To prove that $\mathscr{C}$ is open and dense in $\mathscr{S}$, for control systems defined on the compact space $M \times \Omega$, which has dimension $n + m$, we refer to the general theory of transversality in jet bundles. Then, from the compactness of $M \times \Omega$, it follows that each $f \in \mathscr{C}$ satisfies the uniform controllability condition (uc).

References

1. H. Hermes and J. La Salle, *Functional analysis and time optimal control*, Academic Press, New York, 1969.

2. E. B. Lee and L. Markus, *Foundations of optimal control theory*, Wiley, New York, 1967. MR **36** #3596.

3. C. Lobry, *Dynamical polysystems and control theory*, Geometric Methods in System Theory, Reidel, Amsterdam, 1973, pp. 1–42.

4. L. Markus, *Control dynamical systems*, Math. Systems Theory **3** (1969), 179–185. MR **41** #3144.

5. L. Markus and K. Meyer, *Generic Hamiltonian dynamical systems are neither integrable nor ergodic*, Mem. Amer. Math. Soc. No. 144 (1974).

6. L. Markus and G. Sell, *Capture and control in conservative dynamical systems*, Arch. Rational Mech. Anal. **31** (1968/69), 271–287. MR **18** #2406.

7. ———, *Problems on capture and control in conservative dynamical systems*, SIAM Studies in Appl. Math. **5** (1970), 122–124.

8. ———, *Control in conservative dynamical systems: Recurrence and capture in aperiodic fields*, J. Differential Equations **16** (1974), 472–505.

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455, U.S.A.

UNIVERSITY OF WARWICK
WARWICK, ENGLAND

Труды Международного Конгресса Математиков
Ванкувер, 1974

Управление в Условиях Конфликта и Неопределенности*

А. И. Субботин

В данном сообщении рассматриваются некоторые вопросы теории дифференциальных игр. Основной задачей теории дифференциальных игр является построение такого способа управления, который обеспечивает требуемое качество управляемого процесса при любых неизвестных заранее помехах. В конфликтных ситуациях такими помехами являются управления противника. К настоящему времени опубликовано много работ по теории дифференциальных игр. Среди них можно упомянуть исследования Р. Айзекса, П. Варайи, Е.Ф. Мищенко, Л.С. Понтрягина, Б.Н. Пшеничного, У. Флеминга, А. Фридмана. В этих работах предлагаются различные трактовки игровых задач динамики и подходы к их решению. Ниже излагается подход, который сложился в Институте математики и механики Уральского научного центра АН СССР и опирается в основном на результаты Н.Н. Красовского и автора данного сообщения.

Введем основные определения, в рамках которых будем рассматривать решение позиционных игровых задач динамики. Пусть движение конфликтно управляемой системы описывается уравнением

$$\dot{x} = f(t, x, u, v).$$

Здесь x— n-мерный фазовый вектор; f—непрерывная функция, липшицева по x; u и v—векторы, выбор которых подчинен первому и второму игрокам и стеснен условиями $u \in P$, $v \in Q$, где P и Q—некоторые компакты. Предполагается, что рассматриваемые ниже движения продолжимы до любого момента времени $t = \vartheta$. Стратегии игроков $U \div u(t, x)$ и $V \div v(t, x)$ отождествляются с функциями $u = u(t, x)$ и $v = v(t, x)$, области значений которых содержатся

*Not presented in person.

в компактах P и Q соответственно. Относительно этих функций не предполагается выполнение каких-либо условий непрерывности. Движения, порожденные стратегией $U \div u(t, x)$, вводятся следующим образом. Сначала рассматриваются аппроксимационные движения $x_\Delta(t) = x_\Delta(t, t_0, x_0, U, v(\cdot))$, которые задаются уравнением

$$\dot{x}_\Delta(t) = f(t, x_\Delta(t), u_\Delta(t), v(t)), \qquad x_\Delta(t_0) = x_0, \quad t \geqq t_0, \tag{2}$$

где

$$u_\Delta(t) = u(\tau_i, x_\Delta(\tau_i)) \qquad (\tau_i \leqq t < \tau_{i+1},\ i = 0, 1, 2, \cdots),$$

$\Delta = \{[\tau_i, \tau_{i+1}): i = 0, 1, 2, \cdots, \tau_0 = t_0, \tau_i \to \infty \text{ при } i \to \infty\}$, $v(t) \in Q$ $(t \geqq t_0)$—некоторая суммируемая реализация управления второго игрока. Затем рассматриваются сходящиеся последовательности аппроксимационных движений $x^{(k)}(t) = x_{\Delta_k}(t, t_0, x_0^{(k)}, U, v^{(k)}(\cdot))$ $(k = 1, 2, \cdots)$, $x_0^{(k)} \to x_0$, $\sup_i(\tau_{i+1}^{(k)} - \tau_i^{(k)}) \to 0$ при $k \to \infty$. Равномерный на любом конечном отрезке $[t_0, \vartheta]$ предел такой последовательности функций $x^{(k)}(t)$ $(t \geqq t_0)$ называется движением системы (1), порожденным стратегией $U \div u(t, x)$, и обозначается символом $x(t, t_0, x_0, U)$, где отмечена эта стратегия и начальная позиция (t_0, x_0); управление противника в этом определении и обозначении не фиксируется. Аналогичным образом вводятся движения $x(t, t_0, x_0, V)$, порожденные стратегией $V \div v(t, x)$. При таком определении стратегий и отвечающих им движений справедливо следующее положение.

Пусть M и N—некоторые замкнутые множества в пространстве векторов $p = (t, x)$. Предположим, что для любой позиции $p = (t, x)$ и вектора s справедливо равенство

$$\min_{u \in P} \max_{v \in Q} s'f(t, x, u, v) = \max_{v \in Q} \min_{u \in P} s'f(t, x, u, v) \tag{3}$$

где $s'f$—скалярное произведение векторов s и f. Тогда имеет место следующая альтернатива.

Альтернатива I. Каковы бы ни были начальная позиция (t_0, x_0) и момент времени $t = \vartheta \geqq t_0$, выполняется одно из следующих двух положений: либо существует стратегия $U \div u(t, x)$ такая, что для любого движения $x(t) = x(t, t_0, x_0, U)$ точка $(t, x(t))$ попадет на множество M к моменту $t = \vartheta$, оставаясь при этом во множестве N вплоть до встречи её с множеством M; либо существует стратегия $V \div v(t, x)$, которая для любого движения $x(t) = x(t, t_0, x_0, V)$ гарантирует или уклонение точки $(t, x(t))$ от попадания на множество M вплоть до момента $t = \vartheta$ или нарушение фазового ограничения $(t, x(t)) \in N$ раньше, чем произойдет её встреча с множеством M.

К решению указанных задач о сближении и уклонении сводится исследование многих типов дифференциальных игр, например, тех, где плата задана функционалом одного из перечисленных ниже типов

$$\begin{aligned} &\gamma_1(x(\cdot)) = \min\{t: t \geqq t_0, x(t) \in M^*\}, \qquad \gamma_2(x(\cdot)) = \varphi(x(\vartheta)), \\ &\gamma_3(x(\cdot)) = \min_{t_0 \leqq t \leqq \vartheta} \varphi(x(t)), \qquad \gamma_4(x(\cdot)) = \int_{t_0}^{\vartheta} \varphi(x(t))\, dt. \end{aligned} \tag{4}$$

Здесь M^*—замкнутое множество, ϑ—заданное число, $\varphi(x)$—непрерывная функция. При выполнении условия (3) для этих игр из Альтернативы I выводится справедливость следующего положения.

Альтернатива I*. Для любой начальной позиции (t_0, x_0) и всякого числа $c \in (-\infty, \infty)$ либо существует стратегия $U \div u(t, x)$ такая, что для любого движения $x(t) = x(t, t_0, x_0, U)$ будет справедливо неравенство $\gamma(x(\cdot)) \leqq c$; либо существует стратегия $V \div v(t, x)$ такая, что для любого движения $x(t) = x(t, t_0, x_0, V)$ будет выполнено неравенство $\gamma(x(\cdot)) \geqq c$.

Для рассматриваемых типов дифференциальных игр из Альтернативы I* вытекает, что в рамках избранной формализации существует цена игры. Отметим, что при постановке позиционных игровых задач динамики возможны и другие определения движений и стратегий. Например, при рассмотрении разрывных стратегий иногда удобно заменять разрывную правую часть дифференциального уравнения неоднозначной и использовать аппарат теории дифференциальных уравнений в контингенциях. Однако такой подход может оказаться неудачным. Известны примеры, в которых решение игровых задач, полученное в рамках предложенной здесь формализации, не удается получить и даже аппроксимировать в рамках других формализаций, опирающихся на уравнения в контингенциях или непрерывные стратегии.

Обратимся теперь к случаю, когда условие (3) не выполняется. Введем в рассмотрение контрстратегии $U_v \div u(t, x, v)$ и $V_u \div v(t, x, u)$, которые отождествляются с функциями $u = u(t, x, v)$ и $v = v(t, x, u)$, удовлетворяющими включениям $u(t, x, v) \in P$ и $v(t, x, u) \in Q$ борелевскими по переменным v и u соответственно. Движения $x(t, t_0, x_0, U_v)$, отвечающие контрстратегии $U_v \div u(t, x, v)$, определяются предельным переходом от аппроксимационных движений $x_\Delta(t)$, которые удовлетворяют уравнению (2), где

$$u_\Delta(t) = u(\tau_i, x_\Delta(\tau_i), v(t)), \qquad \tau_i \leqq t < \tau_{i+1}, \quad i = 0, 1, 2, \cdots.$$

Аналогичным образом вводятся движения $x(t, t_0, x_0, V_u)$. Для дифференциальных игр, рассматриваемых в классе стратегий одного игрока и контрстратегий другого, справедливы Альтернативы II и II*, формулировки которых получаются из формулировок Альтернатив I и I* заменой стратегии одного из игроков на соответствующую контрстратегию.

При нарушении условия (3) возможна и другая постановка дифференциальных игр, в которой решение игровых задач требуется определить в классе смешанных стратегий $\tilde{U} \div \mu(t, x)$ и $\tilde{V} \div \nu(t, x)$. Здесь функция $\mu = \mu(t, x)$ ($\nu = \nu(t, x)$) позиции игры (t, x) ставит в соответствие вероятностную меру μ (меру ν), заданную на компакте P (компакте Q). Движения $x(t, t_0, x_0, \tilde{U})$ определяются предельным переходом от последовательности аппроксимационных движений $x_\Delta(t)$, удовлетворяющих уравнению

$$\dot{x}_\Delta(t) = \int_P \int_Q f(t, x_\Delta(t), u, v)\, d\mu_{\Delta, t}\, d\nu_t$$

где

$$\mu_{\Delta, t} = \mu(\tau_i, x_\Delta(\tau_i)), \qquad \tau_i \leqq t < \tau_{i+1}, \quad i = 0, 1, 2, \cdots,$$

ν_t —некоторая слабо измеримая функция. Аналогичным образом вводятся движения $x(t, t_0, x_0, \tilde{V})$. Для дифференциальных игр, рассматриваемых в классе смешанных стратегий справедливы Альтернативы III и III*, формулировки которых получаются из формулировок Альтернатив I и I* заменой детерминированных стратегий $U \div u(t, x)$ и $V \div v(t, x)$ смешанными стратегиями $\tilde{U} \div \mu(t, x)$ и $\tilde{V} \div \nu(t, x)$.

Содержание понятий смешанных стратегий и порожденных ими движений раскрывается при рассмотрении стохастических процедур, в которых мгновенное смешивание управлений заменяется последовательностью вероятностных испытаний. При условии взаимно независимого или слабо коррелированного выбора случайных управлений игроков эти стохастические процедуры порождают пучки случайных движений, которые с вероятностью, сколь угодно близкой к единице, попадают в наперед заданную окрестность соответствующих идеальных движений $x(t, t_0, x_0, \tilde{U})$ или $x(t, t_0, x_0, \tilde{V})$. Эти стохастические процедуры являются физически реализуемыми и аппроксимируют решения, полученные в рамках идеальной конструкции. Отметим, что обоснованием предположения о взаимно независимом или слабо коррелированном выборе случайных управлений игроков может служить наличие информационных помех в системе управления.

Доказательство Альтернатив I—III осуществляется по единой схеме. Опишем кратко доказательство Альтернативы I. Рассмотрим сначала задачу о сближении. Предлагаемое решение этой задачи можно наглядно истолковать следующим образом. В пространстве позиций (t, x) выделяется некоторое множество $W^{(u)}$, образующее мост, который соединяет начальную позицию (t_0, x_0) с целевым множеством M и лежит целиком во множестве N. Мост $W^{(u)}$ обладает свойством u-стабильности, что позволяет построить стратегию U^0, которая удерживает движение на этом мосту и тем самым доставляет решение задачи о сближении. Формально мост $W^{(u)}$ определяется как замкнутое множество, которое удовлетворяет условиям $W^{(u)} \subset N$, $W^{(u)}_\vartheta = \{(\vartheta, x): (\vartheta, x) \in W^{(u)}\} \subset M$ и обладает следующим свойством u-стабильности: каковы бы ни были точка $(t_*, x_*) \in W^{(u)}$, момент времени $t^* \in [t_*, \vartheta]$ и вектор $v_* \in Q$, существует решение $x(t)$ уравнения в контингенциях

$$\dot{x}(t) \in \operatorname{co}\{f(t, x(t), u, v_*): u \in P\}, \qquad x(t_*) = x_*,$$

удовлетворяющее или условию $(t^*, x(t^*)) \in W^{(u)}$ или условию $(\tau, x(\tau)) \in M$ при некотором $\tau \in [t_*, t^*]$.

Справедливо следующее положение.

Лемма. Если $(t_0, x_0) \in W^{(u)}$, то для любого движения $x(t) = x(t, t_0, x_0, U^0)$ точка $(t, x(t))$ остается на мосту $W^{(u)}$ вплоть до встречи её с множеством M. Здесь $U^0 \div u^0(t, x)$ —стратегия, экстремальная к мосту $W^{(u)}$, определяется соотношением

$$\min_{u \in P} \max_{v \in Q} (x - w)' f(t, x, u, v) = \max_{v \in Q} (x - w)' f(t, x, u^0(t, x), v)$$

где w—вектор, для которого (t, w)—точка множества $W^{(u)}$, ближайшая к позиции (t, x) (если $W_t^{(u)} = \emptyset$, то $u^0(t, x)$—произвольная точка компакта P).

Затем в пространстве позиций выделяется множество W^*—совокупность всех позиций (t_*, x_*), для которых как для начальных в классе стратегий $V \div v(t, x)$ разрешима задача об уклонении. Оказывается, что дополнение к множеству W^* является мостом $W^{0(u)}$ в задаче о сближении. Поэтому справедливость Альтернативы I вытекает теперь из определения множества $W^{0(u)}$ и сформулированной выше леммы.

Основным моментом предлагаемого решения игровых задач динамики является определение подходящих мостов, после чего построение экстремальных стратегий не доставляет больших трудностей. Для задач сближения и уклонения известны рекуррентные процедуры построения таких мостов, однако практическое использование этих алгоритмов ограничивается большими трудностями вычислительного характера. Поэтому представляет интерес исследование более простых способов построения стабильных мостов. Наиболее полно в этом отношении изучены регулярные случаи программного поглощения, здесь решение игровых задач удается получить в наиболее завершенном виде и сформулировать для него принцип минимакса, который можно рассматривать как обобщение принципа максимума Л.С. Понтрягина на случай игровых задач динамики. В регулярных случаях решение игровых задач динамики можно довести до алгоритмов, реализуемых на вычислительных машинах.

Решение игровых задач в форме экстремальной стратегии может оказаться неустойчивым по отношению к информационным помехам. Для стабилизации этого решения используется процедура управления с поводырем, содержание которой состоит в следующем. Наряду с реальной системой вводится в рассмотрение подобная ей эталонная система—поводырь, движение которой моделируется в вычислительной машине и известно с любой требуемой точностью. Управление игрока в реальной системе и движение поводыря формируются так, чтобы они взаимно отслеживались, причем поводырь остается на стабильном мосту вплоть до попадания его на целевое множество. Такая процедура управления устойчива к ошибкам измерения текущей позиции игры и помехам, действующим на систему.

Выше были указаны некоторые случаи (см. (4)), когда существует цена позиционной дифференциальной игры. Рассмотрим теперь общий случай, когда плата $\gamma(x(\cdot))$—произвольный полунепрерывный снизу (или сверху) функционал, определенный на непрерывных функциях $x(t)$ $(t \geqq t_0)$, т.е. $\overline{\lim}_{k\to\infty} \gamma(x_k(\cdot)) \leqq \gamma(x_*(\cdot))$ (или $\underline{\lim}_{k\to\infty} \gamma(x_k(\cdot)) \geqq \gamma(x_*(\cdot))$), где функция $x_*(t)$ $(t \geqq t_0)$—равномерный на каждом конечном отрезке $[t_0, \vartheta]$ предел последовательности функций $x_k(t)$ $(t \geqq t_0)$. Введем в рассмотрение стратегии $U \div u(x(\cdot, t_0, t))$ и $V \div v(x(\cdot, t_0, t))$, которые отождествляются с операторами $u = u(x(\cdot, t_0, t))$ и $v = v(x(\cdot, t_0, t))$, определенными при каждом $t \in [t_0, \infty)$ на

пространстве непрерывных функций $x(\tau)$ $(t_0 \leqq \tau \leqq t)$ и принимающие значения на компактах P и Q соответственно. Движения $x(t, t_0, x_0, U)$, порожденные стратегией $U \div u(x(\cdot, t_0, t))$, определяются предельным переходом от аппроксимационных движений $x_\Delta(t)$ которые удовлетворяют уравнению (2), где $u_\Delta(t) = u(x_\Delta(\cdot, t_0, \tau_i))$ $\tau_i \leqq t < \tau_{i+1}$, $i = 0, 1, 2, \cdots$, $x_\Delta(\cdot, t_0, \tau_i)$—аппроксимационное движение, реализовавшееся на отрезке $[t_0, \tau_i]$. Аналогичным образом вводятся движения $x(t, t_0, x_0, V)$, порожденные стратегией $V \div v(x(\cdot, t_0, t))$. При выполнении условия (3) для произвольного полунепрерывного функционала $\gamma(x(\cdot))$ справедлива Альтернатива IV, формулировка которой получается из формулировки Альтернативы I* заменой позиционных стратегий $U \div u(t, x)$ и $V \div v(t,x)$ стратегиями $U \div u(x(\cdot, t_0, t))$ и $V \div v(x(\cdot, t_0, t))$, которые опираются на информацию об истории игры. Если же условие (3) нарушается, то вводятся контрстратегии $U_v \div u(x(\cdot, t_0, t), v)$ и $V_u \div v(x(\cdot, t_0, t), u)$ или смешанные стратегии $\tilde{U} \div \mu(x(\cdot, t_0, t))$ и $\tilde{V} \div \nu(x(\cdot, t_0, t))$ и Альтернативы V и VI формулируются для дифференциальных игр, рассматриваемых в классе пар

$$\{U \div u(x(\cdot, t_0, t)),\ V_u \div v(x(\cdot, t_0, t), u)\}$$

или

$$\{U_v \div u(x(\cdot, t_0, t), v),\ V \div v(x(\cdot, t_0, t))\} \quad \text{и} \quad \{\tilde{U} \div \mu(x(\cdot, t_0, t)),\ \tilde{V} \div \nu(x(\cdot, t_0, t))\}.$$

При доказательстве Альтернатив IV—VI также используется экстремальная конструкция, при этом стабильные мосты строятся теперь не в конечномерном пространстве позиций (t, x) а в функциональном пространстве непрерывных функций $x(\tau)$, $\tau \geqq t_0$.

В заключение отметим, что экстремальная конструкция применяется также при решении игровых задач динамики с интегральными ограничениями на ресурсы управления игроков, при исследовании дифференциально-разностных игр и задач управления в условиях неполной информации о ходе процесса. Во всех этих случаях в описанную выше схему вносятся некоторые изменения, однако общее содержание подхода остается неизменным.

Институт Математики и Механики АН СССР
Свердловск 620000, ГСП-384, СССР

Section 16

Mathematical Physics and Mechanics

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Spectral Deformation Techniques and Application to N-Body Schrödinger Operators

J.-M. Combes

We describe a new tool in spectral analysis which supplements the standard method of perturbation theory for continuous or discrete spectrum. The emphasis is mostly on the problem of analytically continuing matrix elements of selfadjoint operators, motivated by classical and quantum scattering theory or the general theory of linear passive systems [3], although this method also appears as being very useful for analysing properties of eigenvectors. The basic idea is that similarity via unbounded closed operators U does not preserve continuous spectrum in general although a discrete spectrum is usually invariant if associated eigenvectors are in the domain of U. This theory can then be used in two ways:

(1) An active way where one tries to distort continuously the spectrum of an operator T in order to obtain larger analyticity domains for matrix elements of the resolvent. In practice such spectral deformations are performed by embedding T into an analytic family $T(\lambda)$, $\lambda \in \boldsymbol{C}$, such that $\sigma(T(\lambda)) = \Gamma(\lambda)$ where $\Gamma(\lambda)$ is a family of homotopic paths in the connected region of $\boldsymbol{C}$ where one wants to perform the analytic continuation. If this deformation is implementable by a group $U(\lambda)$, $\lambda \in \boldsymbol{R}$, such that

$$U(\lambda - \lambda_0)T(\lambda)\, U^{-1}(\lambda - \lambda_1) = T(\lambda_0) \quad \text{if } \lambda - \lambda_0 \in \boldsymbol{R},$$

then analytic continuation can be performed for matrix elements between analytic vectors with respect to this group.

(2) A passive way where one tries to control the deformation of the spectrum under similarity transformations. Invariance of a given eigenvalue usually implies domain properties of the corresponding eigenvectors.

In the abstract treament described below we try to synthesize a series of papers by

various authors which appeal more or less implicitly to the concept of spectral deformation, and summarize some of the most fruitful recent applications.

I. Let t be a densely defined sesquilinear closed sectorial form on a Hilbert space $\mathscr{H}$ and T the associated sectorial operator [1]. We will denote by $\mathscr{H}_+$ the Hilbert space canonically associated to the domain of t [2]. The space $\mathscr{H}_-$ of continuous linear functionals on $\mathscr{H}_+$ contains $\mathscr{H}$ as a dense subspace according to $|\langle \Phi, \Psi \rangle_{\mathscr{H}}| \leqq \|\Phi\|_{\mathscr{H}} \|\Psi\|_{\mathscr{H}_+}$. T is a continuous mapping from $\mathscr{H}_+$ to $\mathscr{H}_-$ and one has $t(\Phi, \Psi) = \langle \Phi, T\Psi \rangle$, $(\Phi, \Psi) \in \mathscr{H}_+ \times \mathscr{H}_+$, where $\langle \cdot, \cdot \rangle$ is the sesquilinear form on $\mathscr{H}_- \times \mathscr{H}_+$ extending $\langle \cdot, \cdot \rangle_{\mathscr{H}}$ defined on $\mathscr{H} \times \mathscr{H}_+$.

We now consider two such forms t_i on $\mathscr{H}^i$, $i = 1, 2$. Let $(\bar{U}, U)$ be a pair of densely defined closed linear mappings from $\mathscr{H}^1_+$ to $\mathscr{H}^2_+$; we denote by $D_{\bar{U}}$, D_U their domains.

DEFINITION. We will say that $(\bar{U}, U)$ intertwine (t_1, t_2) if:

(1) $\bar{U}$ and U have zero null set and dense range.

(2) $\Phi, \Psi \in D_{\bar{U}} \cap D_U$ one has

$$\langle \bar{U}\Phi, U\Psi \rangle_{\mathscr{H}^2} = \langle \Phi, \Psi \rangle_{\mathscr{H}^1}, \tag{A}$$

and

$$t_2(\bar{U}\Phi, U\Psi) = t_1(\Phi, \Psi). \tag{B}$$

(3) For a sufficiently large and positive,

$$(T_1 + a)\, \mathscr{D}_U = R_{\bar{U}^*}$$

where $R_{\bar{U}^*}$ denotes the range of $\bar{U}^*$. Some direct consequences of (1), (2) and (3) are (denoting by $\rho(T)$, $\sigma(T)$, etc. the resolvent set, spectrum, etc. of T):

THEOREM 1. *Let U^* (resp. $\bar{U}^*$): $\mathscr{H}^2_- \to \mathscr{H}^1_-$ be the adjoint of U (resp. $\bar{U}$). Then, $\forall\, \tilde{\Phi}, \tilde{\Psi} \in D_{U^*} \times D_{\bar{U}^*}$ and $\forall\, z \in \rho(T_1) \cap \rho(T_2)$ and* Re *z sufficiently negative, one has:*

$$\langle U^*\tilde{\Phi}, (T_1 - z)^{-1}\, \bar{U}\tilde{\Psi} \rangle = \langle \hat{\Phi}, (T_2 - z)^{-1}\, \tilde{\Psi} \rangle. \tag{C}$$

Under our assumption the ranges of U^* and $\bar{U}^*$ are dense in $\mathscr{H}^2_-$ and in practical applications have a dense intersection. This theorem furnishes then a basis of some of the analytic continuation techniques via the

COROLLARY. *$\forall\, \Phi \in R_{U^*} \cap R_{\bar{U}^*}$ the function $\langle \Phi, (T_1 - z)^{-1}\Phi \rangle$ has a piecewise analytic continuation in $\rho(T_1) \cup \rho(T_2)$.*

When T_1 is selfadjoint this corollary allows in some cases analytic continuation through the continuous spectrum for expectation values of the resolvent as we will see later. On the other hand, using the fact that $R(U^*)$ and $R(\bar{U}^*)$ are dense it is then possible to derive, e.g., results on the absence of continuous singular spectrum via Greenstein's theorem [3] asserting that a necessary and sufficient condition for

$$\langle \Phi, (T - z)^{-1}\Phi \rangle = \int_{-\infty}^{+\infty} \frac{1}{\lambda - z}\, d\langle \Phi, E(\lambda)\Phi \rangle$$

($E(\cdot)$: spectral family of T) to have an analytic continuation through $[a, b]$ in that $\langle \Phi, E(\lambda)\Phi \rangle$ is real analytic for λ in this interval.

Information about the discrete spectrum can also be obtained:

THEOREM 2. *Assume T_1 is selfadjoint and $E \in \sigma_p(T_1) \cap C\sigma_e(T_2)$. Then E is an eigenvalue of T_2 and has the same multiplicity as an eigenvalue of T_1 or T_2. If E has finite multiplicity and $T_1\Psi = E\Psi$, then $\Psi \in D_U \cap D_U$.*

EXAMPLE 1. $\mathscr{H}$ is a Hilbert space, $\alpha(z)$ an analytic function in a neighborhood $\mathscr{D}$ of $\boldsymbol{R}^+$ with positive derivative on $\boldsymbol{R}^+$. Let $\Gamma \subset \mathscr{D}$ be a continuous path and define:

$$\mathscr{H}^\Gamma = L^2(\Gamma, |d\alpha|; \mathfrak{h})$$

i.e., the Hilbert space of $\mathfrak{h}$-valued measurable functions Φ on Γ such that

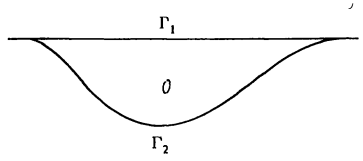
$$\int_\Gamma \|\Phi(z)\|_{\mathfrak{h}}^2 |d\alpha(z)| < \infty.$$

Let

$$t_\Gamma(\Phi, \Psi) = \int_\Gamma z^2 \langle \Phi(z), \Psi(z) \rangle_{\mathfrak{h}} |d\alpha(z)|. \tag{3}$$

If Γ is contained in a sector then t_Γ is sectorial and one has

$$\mathscr{H}_\pm^\Gamma = \left\{ \Phi \in \mathscr{H}^\Gamma; \int_\Gamma \|\Phi(z)\|_{\mathfrak{h}}^2 (1 + |z|^2)^{\pm 1} |d\alpha(z)| < \infty \right\}.$$



Let $\Gamma_1 = \boldsymbol{R}^+$ and Γ_2 be obtained by some local distortion of $\boldsymbol{R}^+$ and $\mathscr{V}$ the complex domain with boundary $\Gamma_1 \cup \Gamma_2$. Then define $\mathscr{D}_U = \{\Phi \in \mathscr{H}_+^{\Gamma_1}$ s.t. $\Phi(z)$ has an $\mathfrak{h}$-valued analytic continuation in $\mathscr{V}$ with boundary value $\Phi_{\Gamma_1} \in \mathscr{H}_+^{\Gamma_1}\}$ and

$$(U\Phi)(z) = \left(\frac{d\alpha(z)}{|d\alpha(z)|} \right)^{1/2} \Phi_{\Gamma_1}(z), \qquad z \in \Gamma_2. \tag{4}$$

Then U is closed on D_U. In the same way one defines, assuming $\bar{\Gamma}_2 \subset \mathscr{D}$,

$$(\bar{U}\Phi)(z) = \left(\frac{d\alpha(\bar{z})}{|d\alpha(z)|} \right)^{1/2} \Phi_{\bar{\Gamma}_1}(\bar{z}).$$

Then property (1) is an easy consequence of the Stone-Weierstrass theorem and properties (2) and (3) follow immediately from the Cauchy equality and other elementary properties of analytic functions.

Since $\sigma(T_\Gamma) = \Gamma^2$ it is clear that this example fulfills the analytic continuation program for resolvents of selfadjoint operators with uniform multiplicity like the Laplacian in $L^2(\boldsymbol{R}^n)$. It is strongly related to the local analytic continuations studied

by Balslev and Babbit [4], Dolph, McLeod and Thoe [5], Grossman [6], Thomas [8].

We give now examples of global deformations of the spectrum.

EXAMPLE 2. We now take $\Gamma_1 = \boldsymbol{R}^+$ and $\Gamma_2 = e^{-ia}\,\boldsymbol{R}^+$, $0 < a \leqq \pi/2$. If $\alpha(z) = z^n$ it is convenient to identify $\mathscr{H}^{\Gamma_1}$ and $\mathscr{H}^{\Gamma_2}$ via

$$(I\Phi)(z) = \Phi(e^{-ia}z), \qquad z \in \boldsymbol{R}^+.$$

Clearly this mapping from $\mathscr{H}_{\Gamma_1}$ to $\mathscr{H}_{\Gamma_2}$ is unitary. The connection between T_1 and T_2 associated to the quadratic forms $t_{\Gamma_1}(\Phi, \Psi)$ and $t_{\Gamma_1}(I^{-1}\Phi, I^{-1}\Psi)$ as defined by (3) is simply

$$T_2 = e^{-2ia}T_1. \tag{5}$$

So $\sigma(T_2) = e^{-2ia}\sigma(T_1)$. The construction of U and $\bar{U}$ can be made in the same way as before. A more elegant formulation can be obtained by noticing that in that case there exists an interpolating group $U(\theta)$, $|\operatorname{Im}\theta| < a$, such that $U = e^{aA}$ and $\bar{U} = e^{-aA}$ where, for $\theta \in \boldsymbol{R}$,

$$(U(\theta)\Phi)(z) = e^{-n\theta/2}\,\Phi(e^{-\theta}z), \qquad z \in \boldsymbol{R}^+. \tag{6}$$

Then D_U (resp. $D_{\bar{U}}$) is the set of Φ's in $\mathscr{H}_+^{\ddagger}$ such that $\Phi(\theta) = U(\theta)\Phi$ extends to an analytic $\mathscr{H}_+$-valued function in $-a \leqq \operatorname{Im}\theta \leqq 0$ (resp. $0 \leqq \operatorname{Im}\theta \leqq a$). (Notice that $U(\theta)$ is a bounded group for $\theta \in \boldsymbol{R}$.) The family $T(\theta) = e^{-2\theta}T_1$ is an analytic family of type (B) and one has, for $\Phi, \Psi \in D_{\bar{U}} \times D_U$,

$$\langle \Phi(\bar{\theta}),\, T_1(\theta)\Psi(\theta)\rangle = \langle \Phi,\, T_1\Psi\rangle,$$

since both sides are analytic for $-a \leqq \operatorname{Im}\theta \leqq 0$ and coincide for $\theta \in \boldsymbol{R}$ as can be seen by an elementary change of variables in the integral (3). Then property (B) follows by taking the limit $\theta = ia$. One shows (A) by the same kind of argument.

The group $U(\theta)$, $\theta \in \boldsymbol{R}^+$, so defined is called the *dilation group*. Complex dilations have been used in a number of papers on analytic continuation of scattering amplitudes. One of the first systematic investigations is due to the physicist Lovelace [9]. They have been used for the study of Schrödinger operators by Aguilar, Balslev and Combes [10], [11], Simon [12] and Van Winter [13]. Other applications include Weber [14], [15] for the Dirac and Lee Hamiltonians.

EXAMPLE 3. This example is given in view of the "passive use" of spectral deformation techniques.

Let $\mathscr{H}_1 = \mathscr{H}_2 = L^2(\boldsymbol{R}^n; \mathfrak{h})$ and

$$t_1(\Phi, \Psi) = \int k^2 \langle \Phi(\boldsymbol{k}), \Psi(\boldsymbol{k})\rangle_{\mathfrak{h}}\, d^n\boldsymbol{k}. \tag{7}$$

Let $\tau \in \boldsymbol{R}^n$ and

$$t_2(\Phi, \Psi) = \int (\boldsymbol{k} + i\tau)^2 \langle \Phi(\boldsymbol{k}), \Psi(\boldsymbol{k})\rangle_{\mathfrak{h}}\, d^n\boldsymbol{k}. \tag{7'}$$

We are again in a situation where there exists an interpolating analytic family

$$t_\alpha(\Phi, \Psi) = \int (\boldsymbol{k} + \alpha\tau)^2 \langle \Phi(\boldsymbol{k}), \Psi(\boldsymbol{k})\rangle_{\mathfrak{h}}\, d^n\boldsymbol{k}.$$

Analyticity in α is obvious for Φ, Ψ in the common form domain

$$\mathscr{H}_+ = \{\Phi \in \mathscr{H};\ \int k^2 \langle \Phi(\boldsymbol{k}), \Psi(\boldsymbol{k}) \rangle_{\mathfrak{h}}\, d^n\boldsymbol{k} < \infty\}.$$

There also exists an interpolating group $U(\alpha)$, $\alpha \in \boldsymbol{R}$, defined by

$$(U(\alpha)\Phi)(\boldsymbol{k}) = \Phi(\boldsymbol{k} + \alpha\tau). \tag{8}$$

Then analyticity arguments as those developed for the dilation group lead to the verification that $\bar{U} = U(-i)$ and $U = U(i)$ intertwine (t_1, t_2). The group (8) is called the boost group. It has been used in connection with the study of decay properties of eigenfunctions for Schrödinger operators by T. O. Connor [16], Combes and Thomas [17], Balslev [18].

Another interesting example is provided by Schrödinger operators $T_1 = -\Delta + V$ on $L^2(\boldsymbol{R}^3)$, where V is the multiplication operator by some function $V(x) \geqq 0$. Let E be an eigenvalue of T; then the class of operators $(\bar{U}, U)$ intertwining T_1 and any T_2 having E as an isolated eigenvalue contains the operator U of multiplication by $\exp(\theta \int_0^{|x|} W(r)\, dr)$, $0 \leqq \theta < 1$, where $W^2(r) = \inf_{|x|=r} V(x)$. If E has finite multiplicity this result of Simon [19] provides L^2-bounds of WKB type on eigenfunctions particularly useful in some problems of singular perturbation theory (Combes [20], Maslov [21]).

II. Perturbation theory. Most of the above examples bring really interesting results only after they are submitted to perturbations. Let v_1 be a t_1-bounded form with relative bound smaller than one; assume that $(\bar{U}, U)$ intertwine (t_1, t_2).

THEOREM 3. *If there exists a t_2-bounded form v_2 with relative bound smaller than one such that*

$$v_2(\bar{U}\Psi, U\Phi) = v_1(\Psi, \Phi), \quad \forall\, \Phi, \Psi \in \mathscr{D}_{\bar{U}} \times \mathscr{D}_U, \tag{9}$$

then $(\bar{U}, U)$ intertwine $(t_1 + v_1, t_2 + v_2)$.

Let V_i, $i = 1, 2$, be the bounded mappings from $\mathscr{H}^i_+$ to $\mathscr{H}^i_-$ associated to the forms v_i.

THEOREM 4. *If V_i, $i = 1, 2$, are compact then $\forall\ \Phi \in R_{\bar{U}^*} \cap R_{U^*}$ the function $(\Phi, (T_1 + V_1 - z)^{-1}\Phi)$ has a piecewise holomorphic continuation in $\rho(T_1) \cup \rho(T_2)$.*

We describe below some applications of Theorems 1, 2, 3, 4.

EXAMPLES 1—2. Assume that V_1 admits a kernel $\tilde{V}_1(z, z') \in B(\mathfrak{h})$ so that

$$v_1(\Phi, \Psi) = \int_{\boldsymbol{R}^+} \langle \Phi(z), \tilde{V}_1(z, z')\Psi(z') \rangle_{\mathfrak{h}}\, d\alpha(z)\, d\alpha(z').$$

Existence of a v_2 satisfying the conditions of the theorems is guaranteed if $\tilde{V}_1(z, z')$ extends as a $B(\mathfrak{h})$-valued analytic function in $\mathscr{V} \times \mathscr{V}$ having boundary value $\tilde{v}_{\Gamma_1}(z, z')$ and if

$$\tilde{V}_{\Gamma_1}(z, z') = V_{\Gamma_1}(z, z')\, \bar{\beta}(z)\beta(z')$$

where $\beta^2(z) = (z^2 + z_0)^{-1}\, d\alpha(z)/|d\alpha(z)|$, for some $z_0^2 \notin \Gamma_2^2$, defines a compact operator on $\mathscr{H}^{\Gamma_1}$.

As a more specific example let us consider perturbations of the three-dimensional

Laplacian $T_1 = -\Delta$ on $L^2(\boldsymbol{R}^3)$, i.e., the situation of Examples 1 and 2 with $\mathfrak{h} = L^2(\Omega)$, Ω unit sphere in $\boldsymbol{R}^3$, and $\alpha(z) = z^2$. A multiplicative perturbation $V_1(x)$ of $-\Delta$ has a kernel defined by

$$(\tilde{V}_1(z, z')\,\Phi(z))(\boldsymbol{w}) = \int_{\Omega} \tilde{V}_1(z\boldsymbol{w} - z'\boldsymbol{w}')\Phi(z', w')\,dw' \tag{10}$$

where $\tilde{V}_1$ denotes the Fourier transform of V_1. If one wants to perform an analytic continuation in (z, z') up to $\Gamma \times \Gamma$ for an arbitrary Γ contained in some strip $|\operatorname{Im} z| < a$ it is clear from (10) that exponential decay properties of V_1 are required via Paley-Wiener theorem. Actually the following criterion of Babbit and Balslev is sufficient for deformations in such a strip (and also covers nonlocal perturbations):

$$\begin{aligned} &V_1 = QBQ \text{ where } Q \text{ is the multiplication operator by} \\ &e^{-a|x|} \text{ and } B \text{ is compact from } \mathscr{H}_+^{\boldsymbol{R}^+} \text{ to } \mathscr{H}_-^{\boldsymbol{R}^+}. \end{aligned} \tag{11}$$

Actually QU^{-1} extends to a bounded operator from $\mathscr{H}_+^{\Gamma}$ to $\mathscr{H}_+^{\boldsymbol{R}^+}$ for any Γ contained in the strip $|\operatorname{Im} z| < a$, and $\bar{U}^*Q$ is bounded from $\mathscr{H}_-^{\boldsymbol{R}^+}$ to $\mathscr{H}_-^{\Gamma_1}$. This observation is more or less explicit but critical in [5], [6].

For contours $\Gamma = e^{-i\alpha}\boldsymbol{R}^+$ the restrictions on V_1 are totally different. Simon has shown that for radial perturbations, deformation up to $e^{-i\alpha}\boldsymbol{R}^+$ is allowed if

$$\begin{aligned} &V_1(|x|) \text{ has an analytic extension to } \{r; |\operatorname{Arg} r| \leqq a\} \\ &\text{such that } V_\theta(x) = V(e^{i\theta}|x|) \text{ is in } R + (L^\infty)_\varepsilon,\ |\theta| \leqq a. \end{aligned} \tag{12}$$

($V \in R$ iff $\int(|V(X)|\,|V(Y)|/|X - Y|^2)\,d^3X\,d^3Y < \infty$.) This class is denoted by $\mathscr{F}_a$.

So under condition (11) one obtains meromorphic continuation of resolvent matrix elements for $-\Delta + V_1$ in $\{z^2 \mid |\operatorname{Im} z| < a\}$, under condition (12) in $\{z^2 \mid |\operatorname{Arg} z| \leqq a\}$. The poles of these matrix elements have a direct physical interpretation as bound-states or resonance energies [22], [13].

EXAMPLE 3. Assume that V_1 has a kernel $\tilde{V}_1(\boldsymbol{k}, \boldsymbol{k}') \in L(\mathfrak{h}, \mathfrak{h})$ depending only on $\boldsymbol{k} - \boldsymbol{k}'$. Then

$$\begin{aligned} v_1(\Phi, \Psi) &= \int_{\boldsymbol{R}^n} \langle \Phi(\boldsymbol{k}), \tilde{V}_1(\boldsymbol{k} - \boldsymbol{k}')\,\Psi(\boldsymbol{k}')\rangle\,d^n\boldsymbol{k}\,d^n\boldsymbol{k}' \\ &= \int_{\boldsymbol{R}^n} \langle \Phi(\boldsymbol{k} + \alpha\boldsymbol{\tau}), \tilde{V}_1(\boldsymbol{k} - \boldsymbol{k}')\,\Psi(\boldsymbol{k} + \lambda\boldsymbol{\tau})\rangle\,d^n\boldsymbol{k}\,d^n\boldsymbol{k}'\ \forall\,\alpha \in \boldsymbol{R},\ \forall\,\boldsymbol{\tau} \in \boldsymbol{R}^n. \end{aligned}$$

Assuming that $\Phi, \Psi \in D_{\bar{U}} \times D_U$ one gets then by analytic continuation

$$v_1(\Phi, \Psi) = v_1(\bar{U}(\alpha)\Psi, U(\alpha)\Phi) = v_1(\bar{U}\Psi, U\Phi).$$

One is then in the conditions of Theorem 3. When applied to Schrödinger operators $-\Delta + V$, where V is a local small form perturbation of $-\Delta$, this example shows via Theorem 2 that any eigenstate of $-\Delta + V$ with eigenvalue $E < 0$ is in the domain of $K_\theta = e^{\theta|E|^{1/2}|X|}$, $0 \leqq \theta < 1$. This result is obtained by controlling the spectrum of T_2 as defined by (7′) when $\boldsymbol{\tau}$ varies. This property extends to positive eigenvalues by performing simultaneous dilation and boost transformations and

can be used in turn to show absence of such positive eigenvalues if $V_1 \in \mathscr{F}_\pi$ (Balslev [18], Simon [23]).

III. N-particle Schrödinger operators. To conclude we apply these techniques to N-particle Schrödinger operators on $L^2(\mathbf{R}^{3(N-1)})$ defined by $H_1 = T_1 + V_1$ where

$$T_1 = \sum_{\beta=1}^{N-1} (-\Delta_\beta/2\rho_\beta) \tag{13}$$

and

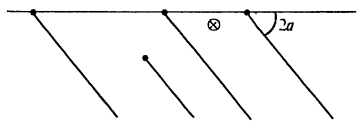
$$V_1 = \sum_\alpha V_\alpha. \tag{14}$$

Here β labels a set of three-dimensional relative coordinates $\boldsymbol{X}_\beta$ for the N-particles and Δ_β is the corresponding three-dimensional Laplacian; the decomposition (13) is nonunique if $N > 2$. As to the V_α's in (14) they are two-particle perturbations depending only on the relative distance (or more generally relative dynamical variables) of the pair α. This last fact prevents V_1 from defining a compact form perturbation of T_1 and one needs new methods to find the spectra of H_1 and $H_2 = T_2 + V_2$ where H_2 is obtained from H_1 by a complex dilation, or boost, or a combination of both. This can be done with the help of multiparticle resolvent equations like those of Weinberg and Van Winter (see [24]) which connect the resolvent of H and those of Hamiltonians of subsystems: $(H - z)^{-1} = D(z) + I(z)(H - z)^{-1}$ where $I(z)$ and $D(z)$ are analytic in $\bigcap_D \rho(H^D)$. Here H^D is obtained from H by subtracting the sum of interactions between particles belonging to different clusters in the k-cluster decomposition $D = \{C_1, C_2, \cdots, C_k\}$ of the system. H^D has the following tensor form

$$H^D = T^D \otimes \left(\bigotimes_{i=1}^{k} I_i\right) + I_D \otimes \left(\sum_{i=1}^{k} I_1 \otimes \cdots \otimes H^{C_i} \otimes \cdots \otimes I_k\right)$$

where H^C is the Hamiltonian for particles in cluster C noninteracting with the remaining particles and T^D is a Laplacian-like operator. So the spectrum of H^D can be found inductively using some recent results on the spectrum of operators having the form $I \otimes A + B \otimes I$ (Ichinose [25], Simon and Reed [26]). The results are as follows: Assume two-particle interactions $V_\alpha \in \mathscr{F}_a$; then

(1) Matrix elements of the resolvent $(H_1 - z)^{-1}$ for vectors in a dense set have a many-sheeted meromorphic continuation through $\sigma(H_1)$ with branch points at $\Sigma = \bigcup_D (\sum_{C \in D} \sigma_{\text{res}}^C)$ (where σ_{res}^C is the set of poles for the matrix elements of $(H_1^C - z)^{-1}$) up to the cuts $\Lambda + e^{-2ia}\mathbf{R}^+$, $\Lambda \in \Sigma$.



(2) Poles are isolated and can accumulate at most at Σ ([11], [12], [13]).

Direct consequences of (1) and (2) are absence of continuous spectrum for H_1, and finite degeneracy of eigenvalues of H_1 in Σ.

(3) Assume $H\Psi = E\Psi$, $E \notin \Sigma$. Then ([16], [17]):

$$\Psi \in \mathscr{D}(e^{\theta(|\Lambda - E| + \Gamma)^{1/2}|X|}), \qquad 0 \leqq \theta < 1,$$

where Λ is the closest point to E in Σ, $\Gamma = |\operatorname{Im} \Lambda|$ and $|x|^2 = \sum_\beta \rho_\beta |x_\beta|^2$.

(4) If $a \geqq \pi$ there are no positive eigenvalues ([18], [23]). Other results include analyticity properties of two-body scattering amplitudes [27] and proofs of unitarity of the s-matrix for $N = 2$ or 3 ([4], [29]).

References

1. T. Kato, *Perturbation theory for linear operators*, Die Grundlehren der math. Wissenschaften, Band 132, Springer-Verlag, New York, 1966. MR **34** #3324.
2. B. Simon and M. C. Reed, *Methods of modern mathematical physics*, Academic Press, New York, 1972.
3. C. Dolph, Ann. Acad. Sci. Fenn. Ser AI No. 336/9 (1963).
4. D. Babbit and E. Balslev, U.C.L.A. Spring, 1974 (preprint).
5. C. Dolph, J. McLeod and D. Thoe, J. Math. Anal. Appl. **16** (1966), 311. MR **35** #552.
6. A. Grossman, J. Math. Phys. **5** (1964), 1025. MR **29** #4450.
7. J. Nuttall, J. Math. Phys. **8** (1967), 873.
8. L. Thomas, Helv. Phys. Acta **45** (1973), 1057.
9. C. Lovelace, *Three-particle systems and unstable particles*, Strong Interactions and High Energy Physics (Scottish Universities' Summer School, 1963), Plenum Press, New York, 1964, pp. 437–475. MR **18** #5732.
10. J. Aguilar and J.-M. Combes, Comm. Math. Phys. **22** (1971), 269.
11. E. Balslev and J.-M. Combes, Comm. Math. Phys. **22** (1971), 280.
12. B. Simon, Comm. Math. Phys. **27** (1974), 1. MR **47** #9989.
13. C. Van Winter, University of Kentucky, 1974 (preprint).
14. R. A. Weber, University of Lourain, April, 1973 (preprint).
15. ———, University of Lourain, May, 1973 (preprint).
16. T. O. Connor, Comm. Math. Phys. **32** (1973), 319.
17. J.-M. Combes and L. Thomas, Comm. Math. Phys. **34** (1973), 251.
18. E. Balslev, U.C.L.A., 1973 (preprint).
19. B. Simon, Princeton, Summer, 1974 (preprint).
20. J.-M. Combes, Toulon, 1974 (preprint).
21. V. P. Maslov, *Theory of perturbations and asymptotic methods*, Izdat. Moskov. Gos. Univ., Moscow, 1965; French transl., Dunod, Paris, 1972.
22. B. Simon, Ann. of Math. (2) **97** (1973), 247.
23. ———, Math. Ann. **207** (1974), 133.
24. ———, *Quantum mechanics for hamiltonians defined as quadratic forms*, Princeton Series in Physics, 1971.
25. T. Ichinose, Math. J. **50** (1973), 185.
26. B. Simon and M. Reed, J. Functional Analysis **13** (1973), 107.
27. J.-M. Combes, *Proceedings of Denver conference on scattering theory*, Reidel, Dordrecht 1974.

INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47401, U.S.A.

CENTRE UNIVERSITAIRE DE TOULON
TOULON, FRANCE

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Time Evolution of Infinite Classical Systems

Oscar E. Lanford III

I will discuss in this article some recent progress in the problem of proving existence and uniqueness of solutions to Newton's equations of motion for infinite systems of classical particles interacting by two-body forces which go to zero reasonably quickly as the particle separation goes to infinity. For technical simplicity, I will assume that the interparticle potential Φ has a Lipschitz continuous derivative and finite range, but the results I will describe have extensions which require neither finite range nor the absence of singularities in the potential.

To establish notation: We consider systems of infinitely many particles with positions (q_i) and momenta (p_i), moving in $\boldsymbol{R}^\nu$. The equations of motion are

$$\frac{dq_i}{dt} = \frac{p_i}{m}, \qquad \frac{dp_i}{dt} = F_i = \sum_{j \neq i} F(q_i - q_j) \tag{1}$$

where m is the particle mass and $F = -\operatorname{grad} \Phi$ is the interparticle force. We assume that there are infinitely many particles, but that, initially at least, they are distributed so that there are only finitely many in each bounded region of space. Because of the infinite number of particles, these equations cannot be treated by the usual elementary techniques, and it is indeed not hard to imagine that some solutions may develop "singularities" in which, for example, infinitely many particles rush into a bounded region of space. What is needed is an existence result which assures us that such singularities are at least improbable.

The result to be described assumes, in addition to the regularity mentioned above, that the interparticle potential Φ has good thermodynamic properties. More specifically, we assume that Φ is *superstable* in the sense of Ruelle [6]. It is then possible to single out a class of probability measures on the phase space for the infinite system—the so-called Gibbs states—which represent thermodynamic

equilibrium for the interaction in question. (See [1], [4], or [6].) What we show is that there exists a set of solutions to the equations of motion forming a set of probability one for each Gibbs state. We are, however, not able to describe in any very explicit way the set of phase points which lie on such solution curves, nor are we able to prove existence of solutions for initial phase points representing situations which are globally not in thermodynamic equilibrium. In this respect, the results described here are much weaker than previous work on one-dimensional systems [2] which proved existence and uniqueness of solutions for all initial phase points satisfying some reasonable regularity conditions.

Mathematically, the main novelty in the argument we will give is that it exploits the formal fact that Gibbs states ought to be invariant under the flow we are trying to construct. This leads to an a priori estimate which is shown to hold almost everywhere with respect to each Gibbs state. The idea is that Gibbs states are concentrated on very well-behaved phase points, and the invariance ought to imply good behavior at all times. The following result illustrates the argument:

PROPOSITION 1. *Let μ be a Gibbs state, and assume that the equations of motion can be solved almost everywhere to give a flow T^t leaving μ invariant. Then, for almost every phase point $\boldsymbol{x} = (q_i, p_i)$, there exists a constant M such that*

$$|q_i(t) - q_i| \leqq M \log_+(q_i) \quad \textit{for all } i \textit{ and } |t| \leqq 1. \tag{2}$$

(Here $\log_+(q)$ denotes $\log(|q|)$ if $|q| \geqq e$ and 1 otherwise.)

To prove this result, we define a function B on the infinite system phase space by

$$B(\boldsymbol{x}) = \sup_i \left\{ \frac{|p_i/m|}{\log_+(q_i)} \right\}. \tag{3}$$

To say that $B(\boldsymbol{x})$ is finite says that velocity fluctuations grow at most like the logarithm of the distance from the origin. A simple argument, using the Maxwellian (i.e., Gaussian) character of the momentum distribution, shows that B is integrable with respect to μ. Now define

$$\bar{B}(\boldsymbol{x}) = \int_{-1}^{1} dt\, B(T^t \boldsymbol{x}).$$

By Fubini's theorem and the assumed invariance of μ under T^t,

$$\int \bar{B}\, d\mu = \int_{-1}^{1} dt \int B \circ T^t\, d\mu = 2 \int B\, d\mu < \infty.$$

Hence, $\bar{B}$ is finite almost everywhere. We now claim that, where $\bar{B}(\boldsymbol{x})$ is finite, there exists a constant M (depending only on $\bar{B}(\boldsymbol{x})$) such that (2) holds. In fact, we have, for each i,

$$\left| \int_0^t dt_1 \frac{|dq_i(t_1)/dt|}{\log_+(q_i(t_1))} \right| \leqq \int_{-1}^{1} dt_1 \sup_i \left\{ \frac{|p_i(t_1)/m|}{\log_+(q_i(t_1))} \right\} = \bar{B}(\boldsymbol{x}).$$

It is now a matter of elementary calculus to show that, for any number b, there exists an $M(b)$ such that

$$\int_0^t dt_1 \frac{|dq_i(t_1)/dt|}{\log_+(q_i(t_1))} \leqq b \quad \text{implies} \quad |q_i(t) - q_i| \leqq M(b) \log_+(q_i).$$

This proves the proposition. Of course, the proposition is of little direct use, since it assumes what was to be proved, the existence of solutions to the equations. Its usefulness derives from the fact that we can find approximate solutions to the equations of motion which leave μ invariant and to which we can apply the above argument. One way to do this is as follows: For each positive integer s, let Λ_s denote the ball of radius s centered about the origin, and let $T^t_{(s)}$ denote the solution flow for the following dynamics:

(a) particles initially outside Λ_s are frozen where they are (i.e., both positions *and momenta* remain fixed);

(b) particles initially inside Λ_s move under their mutual interaction, with constant external forces exerted by the particles outside and with elastic reflection at the boundary of Λ_s.

The definition of Gibbs state readily implies that every Gibbs state is invariant under $T^t_{(s)}$ for all s. We are going to construct solutions to the equations of motion as limits, as $s \to \infty$, of $T^t_{(s)}$.

To do this, we introduce functions

$$\bar{B}_{(s)}(\boldsymbol{x}) = \frac{1}{\pi} \int_{-\infty}^{\infty} dt \, \frac{B(T^t_{(s)}\boldsymbol{x})}{1 + t^2}$$

on the infinite system phase space. As before

$$\int \bar{B}_{(s)} \, d\mu = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dt}{1 + t^2} \int d\mu(\boldsymbol{x}) \, B(T^t_{(s)}\boldsymbol{x}) = \int B \, d\mu \tag{4}$$

for all s, so $\bar{B}_{(s)} < \infty$ almost everywhere. Also, if $\bar{B}_{(s)}(\boldsymbol{x}) \leqq b < \infty$, then, for any i and $|t| \leqq \tau$,

$$\left| \int_0^t dt_1 \frac{dq_i^{(s)}(t_1)/dt}{\log_+(q_i^{(s)}(t_1))} \right| \leqq \left| \int_0^t dt_1 \, B(T^{t_1}_{(s)}\boldsymbol{x}) \right|$$

$$\leqq (1 + \tau^2)\pi \frac{1}{\pi} \int_{-\tau}^{\tau} dt_1 \frac{B(T^{t_1}_{(s)}\boldsymbol{x})}{1 + t_1^2} \leqq (1 + \tau^2) \cdot \pi \cdot b,$$

and hence, in the notation of Proposition 1,

$$|q_i^{(s)}(t) - q_i| \leqq M((1 + \tau^2) \cdot \pi \cdot b) \log_+(q_i)$$

for all i and all t with $|t| \leqq \tau$. This inequality is a kind of localization condition which says that particles stay relatively near their initial positions. We would like to have a bound like this which is uniform in s, for almost all $\boldsymbol{x}$. To get this bound, let

$$\bar{B}_\infty(\boldsymbol{x}) = \liminf_{s \to \infty} \bar{B}_{(s)}(\boldsymbol{x}).$$

By Fatou's lemma, $\bar{B}_\infty$ is integrable and hence is finite almost everywhere. We will show that, if $\bar{B}_\infty(\boldsymbol{x})$ is finite, then there is a solution of the equations of motion with initial data $\boldsymbol{x}$. If $\bar{B}_\infty(\boldsymbol{x})$ is finite, there is a real number b and a sequence (s_n) increasing to infinity such that $\bar{B}_{(s_n)}(\boldsymbol{x}) \leqq b$ for all n. By the argument just given, this

means that, for any positive τ,

$$|q_i^{(s_n)}(t) - q_i| \leqq M((1 + \tau^2)\cdot\pi\cdot b) \log_+(q_i) \tag{5}$$

for all i and all t with $|t| \leqq \tau$.

It is now easy to finish the proof. The bound (5) together with the finite range of the interaction places a bound on the number of particles which can interact with the ith particle for times between $-\tau$ and τ. Since the potential has no singularities, the force which can be exerted by any single particle is bounded so this gives a bound on $|dp_i^{(s_n)}(t)/dt|$ for any fixed i which is uniform in t between $-\tau$ and τ and uniform in n provided s_n is large enough so the inequality (5) prevents the ith particle from colliding with the wall between time $-\tau$ and τ. The Arzelà-Ascoli theorem implies that a subsequence of the $p_i^{(s_n)}(t)$ converges uniformly for t between $-\tau$ and τ. But i and τ are arbitrary, so a diagonal procedure gives a subsequence along which each $p_i^{(s_n)}(t)$ converges uniformly on every bounded interval of times. Let us denote the limits by $(p_i(t))$. If we define

$$q_i(t) = q_i + \int_0^t dt_1\, p_i(t_1)/m \tag{6}$$

then a straightforward passage to the limit in the corresponding equation for the $p_i^{(s_n)}(t)$ gives

$$p_i(t) = p_i + \int_0^t dt_1 \sum_{j\neq i} F(q_i(t_1) - q_j(t_1)) \tag{7}$$

and equations (6) and (7) are simply the integral form of Newton's equations of motion. We have thus proved

THEOREM 2. *Let Φ be a finite-range superstable potential with Lipschitz continuous derivative, and let μ be a Gibbs state for Φ. Let $\bar{B}_\infty$ be defined as above. Then*

(i) $\int \bar{B}_\infty\, d\mu < \infty$, *and in particular $\bar{B}_\infty$ is finite almost everywhere.*

(ii) *If $\bar{B}_\infty(\boldsymbol{x})$ is finite, there exists a solution $\boldsymbol{x}(t) = (q_i(t), p_i(t))$ of the equations of motion with $\boldsymbol{x}(0) = \boldsymbol{x}$ which satisfies the localization condition*

$$\sup_{|t|\leqq\tau} \sup_i \frac{|q_i(t) - q_i|}{\log_+(q_i)} < \infty \tag{8}$$

for all finite τ.

An existence theorem like this one, without a corresponding uniqueness result, is of very little use. Furthermore, examples can be found in which solutions are nonunique, at least for systems of infinitely many hard spheres. Fortunately, the localization condition (8), together with some mild restrictions on the initial phase point $\boldsymbol{x}$ (which hold almost everywhere with respect to each Gibbs state), suffices to determine the solution uniquely. The proof of uniqueness is straightforward: The equations are rewritten as integral equations:

$$q_i(t) = q_i + t(p_i/m) + \int_0^t dt_1 \int_0^{t_1} dt_2\, F_i(t_2),$$

$$F_i(t) = \sum_{j \neq i} F(q_i(t) - q_j(t));$$

it is assumed that these equations have two solutions satisfying the same initial condition; the equations are subtracted and the localization condition (8) is inserted to obtain an integral inequality which is iterated to show that the two solutions must have been identical.

Once uniqueness has been established, the solution mappings give a flow T^t on the infinite system phase which is defined almost everywhere with respect to each Gibbs state. It may further be shown that T^t leaves each Gibbs state invariant, and that $T^t_{(s)}$ converges in measure to T^t as s approaches infinity.

The results described above are discussed in more detail in [3]; a slightly different proof is given in [5]. A manuscript giving the extension to long-range and singular interactions is in preparation. It should be mentioned that there is another approach to the problem of infinite system dynamics, due to Sinaĭ, in which it is shown that almost all initial phase points admit solutions in which, over any bounded interval of time, the particles break up into finite noninteracting clusters. This is proved for arbitrary densities in one dimension [7] and for small densities in more than one dimension [8]; it is surely not true at high densities in more than one dimension.

References

1. R. L. Dobrušin, *Gibbsian random fields. General case*, Funkcional. Anal. i Priložen. **3** (1969), no. 1, 27–35. (Russian) MR **39** #1151.

2. O. E. Lanford III, *The classical mechanics of one-dimensional systems of infinitely many particles.* I. *An existence theorem*; II. *Kinetic theory*, Comm. Math. Phys. **9** (1968), 176–191; ibid. **11** (1968/69), 257–292. MR **38** #2409; **48** #3856.

3. ———, *Lectures on the time-evolution of large classical systems*, Proc. 1974 Battelle Rencontres on Dynamical Systems, Lecture Notes in Physics, Springer-Verlag, Berlin and New York (to appear).

4. O. E. Lanford III and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. **13** (1969), 194–215. MR **41** #1343.

5. C. Marchioro, A. Pellegrinotti and E. Presutti, *Existence of time evolution for ν-dimensional statistical mechanics*, Comm. Math. Phys. (to appear).

6. D. Ruelle, *Superstable interactions in classical statistical mechanics*, Comm. Math. Phys. **18** (1970), 127–159. MR **42** #1468.

7. Ja. G. Sinaĭ, *Construction of the dynamics in one-dimensional systems of statistical mechanics*, Teoret. Mat. Fiz. **11** (1972), 248–258 = J. Theoret. Math. Phys. **11** (1972).

8. ———, *The construction of cluster dynamics for dynamical systems of statistical mechanics*, Vestnik Moskov. Univ. Ser. I Mat. Meh. **29** (1974), 152–158.

UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Thomas-Fermi and Hartree-Fock Theory*

Elliott H. Lieb

Quantum mechanics had its genesis in the attempt to understand the properties of atoms and molecules, particularly the nonclassical feature of a minimum, or ground state, energy. It should come as no surprise, despite the current interest in extending quantum mechanics to the nuclear and subnuclear domain, e.g., field theory, that there are still many important mathematical problems left over from the early days.

The chief triumph that established the Schrödinger equation was the analysis of the energy levels (i.e., discrete spectrum) of the one-electron hydrogen atom Hamiltonian ($Z>0$)

$$H_h = -\Delta - Z|x|^{-1}, \tag{1}$$

on $L^2(\boldsymbol{R}^3)$, where Z is the nuclear charge and $x \in \boldsymbol{R}^3$. The discrete spectrum of H_h is $\{-Z^2(2n)^{-2}: n = 1, 2, \cdots\}$ with degeneracy n^2, and this checks perfectly with experiment.

The good luck stops here. The N-electron Hamiltonian, when there are k nuclei of charges $Z_i > 0$ located at $R_i \in \boldsymbol{R}^3$, $i = 1, \cdots, k$, is

$$H_N = -\sum_{i=1}^{N} \Delta_i + \sum_{i=1}^{N} V(x_i) + \sum_{1\le i<j\le N} |x_i - x_j|^{-1} \tag{2}$$

where $V(x) = -\sum_{i=1}^{k} Z_i |x - R_i|^{-1}$. H_N acts on the N-fold antisymmetric tensor product of $L^2(\boldsymbol{R}^3)$. (Actually $L^2(\boldsymbol{R}^3; \boldsymbol{C}^2)$, but we shall ignore this complication here.) Not much can be done with (2), even when $N = 2$ and $k = 1$.

There exist two classical approximation methods for determining $E_N \equiv \inf \operatorname{spec} H_N$: Thomas-Fermi (TF) theory and Hartree-Fock (HF) theory. These replace (2) by much simpler, although nonlinear, minimization problems. Two

*This work partially supported by NSF grant GP-31674X at M.I.T.

questions arise: (i) Do these new problems actually have minima as distinct from infima? (ii) In what sense are these approximate solutions related to E_N? B. Simon and I [1], [2] have recently been able to answer (i) affirmatively using L^p techniques, and to show that the approximations converge to E_N as $Z \equiv \sum_{i=1}^k Z_i \to \infty$.

I. Thomas-Fermi theory. Let $\rho: \boldsymbol{R}^3 \to \boldsymbol{R}^+$, $\rho \in L^{5/3}(\boldsymbol{R}^3) \cap L^1(\boldsymbol{R}^3)$, and, ignoring some irrelevant constants, define

$$\mathscr{E}(\rho) = \tfrac{3}{5} \int \rho^{5/3} + \int \rho V + \tfrac{1}{2} \iint \rho(x)\, \rho(y)\, |x - y|^{-1}\, dx\, dy. \tag{3}$$

The TF approximation to E_N is

$$E_N^{\mathrm{TF}} \equiv \inf \left\{ \mathscr{E}(\rho) : \int \rho = N \right\}, \qquad N \in \boldsymbol{R}^+.$$

It is easy to prove (Hölder and Young inequalities) that the infimum exists; if it is a minimum, the minimizing ρ will be unique since $\rho \to \mathscr{E}(\rho)$ is strictly convex. In that case it is easy to see that ρ will satisfy the *TF equation*:

$$\rho_N^{2/3}(x) = \max\left[\varphi(x) + C_N, 0\right], \qquad \int \rho = N, \tag{4}$$

where $\varphi(x) = -V(x) - \int \rho_N(y)\, |x - y|^{-1}\, dy$ and $C_N \leqq 0$.

THEOREM 1. *$\mathscr{E}(\rho)$ has a unique, minimizing ρ_N iff $N \leqq Z$, i.e., in the subneutral case. This ρ_N satisfies* (4).

One first shows a minimum exists in $\mathscr{L}_N \equiv \{\rho : \int \rho \leqq N\}$. The idea is as follows:

(a) Observe that one can restrict one's attention to a bounded ball in $L^{5/3}(\boldsymbol{R}^3)$.

(b) $\rho \to \mathscr{E}(\rho)$ is lower semicontinuous in weak $L^{5/3}(\boldsymbol{R}^3)$ since the kernel $|x - y|^{-1}$ is positive definite and $|x|^{-1} \in L^{5/2}(\boldsymbol{R}^3) + L^{\infty}(\boldsymbol{R}^3)$.

(c) If $\lim_{n \to \infty} \mathscr{E}(\rho^{(n)}) = \hat{E}_N \equiv \inf \{\mathscr{E}(\rho) : \rho \in \mathscr{L}_N\}$ for $\rho^{(n)} \in \mathscr{L}_N$, then one can find a subsequence converging in weak $L^{5/3}(\boldsymbol{R}^3)$ to some ρ_N. As long as $\hat{E}_N$ is strictly decreasing in N, $\rho_N \in \{\rho : \int \rho = N\}$. $\hat{E}_N$ will be strictly decreasing until the absolute minimum is reached. This is when $C_N = 0$ and, from the form of $\varphi(x)$, which can be shown by potential theoretic arguments to be positive, this requires that $\int \rho_N = Z$.

By applying potential theory to (4) one can prove interesting properties of ρ_N, E_N^{TF} and C_N. For example, one can prove the Teller-Balasz theorem that if one adds the nuclear repulsion, $\sum_{1 \leq i < j \leq k} Z_i Z_j\, |R_i - R_j|^{-1}$, to E_N^{TF} then the total energy is *greater* than the sum of the energies obtained when the molecule is broken into two smaller molecules, i.e., there is no binding. Another typical result is that C_N, which is called the chemical potential, is a concave function of N on $[0, Z]$.

The connection between E_N^{TF} and the true E_N of (2) is as follows: Fix k and replace Z_i by $z_i N$, and R_i by $N^{-1/3} r_i$ with z_i, r_i fixed. Then it is easy to see that E_N^{TF} satisfies a scaling law: $E_N^{\mathrm{TF}} = N^{7/3} E_1^{\mathrm{TF}}$ and $N^{-2} \rho_N(N^{-1/3} x) = \rho_1(x)$.

THEOREM 2. *In the circumstances given above:*

(i) $\lim_{N \to \infty} E_N / E_N^{\mathrm{TF}} = 1$. *This limit depends, of course, on* $z = \sum_{i=1}^k z_i = Z/N$.

(ii) *If E_N is an eigenvalue of H_N, which will be the case when $\sum_{i=1}^k z_i \geqq 1$, define*

$$\hat{\rho}_N(x) = N \int |\Psi(x, x_2, \cdots, x_N)|^2 \, dx_2 \cdots dx_N,$$

with Ψ being a normalized, ground state eigenfunction of H_N. Then

$$\lim_{N\to\infty} N^{-2}\hat{\rho}_N(N^{-1/3}x) = \rho_1(x) \quad \text{in } L^1_{\text{loc}}(\boldsymbol{R}^3).$$

Theorem 2 is proved by obtaining upper (resp. lower) bounds for E_N by decomposing $\boldsymbol{R}^3$ into cubes of size $N^{-1/3}$ and using Dirichlet (resp. Neumann) boundary conditions on the cube boundaries. The fact that the nuclear charges, Nz_i, go to infinity presents some difficulty.

In summary, the solution to (4), when inserted into (3), gives a remarkably simple asymptotic formula for E_N in the large N limit. It seems to be very difficult to find corrections to this $N^{7/3}$ asymptotic formula; this is an important open problem.

II. Hartree-Fock theory. Let $\Psi = \{\varphi_1, \cdots, \varphi_N\}$ be a set of orthonormal functions in $H^1(\boldsymbol{R}^3)$, i.e., φ_i and $\nabla\varphi_i \in L^2(\boldsymbol{R}^3)$. Define the normalized function

$$D_\Psi(x_1, \cdots, x_N) = (N!)^{-1/2} \det[\varphi_i(x_j)]_{i,j=1}^N \tag{5}$$

in the antisymmetric tensor product of $L^2(\boldsymbol{R}^3)$, and let $\mathscr{E}_N(\Psi) = (D_\Psi, H_N D_\Psi)$. The HF approximation to E_N is

$$E_N^{\text{HF}} = \inf_\Psi \mathscr{E}_N(\Psi).$$

Is this infimum a minimum? If so, the φ_i will satisfy the coupled, nonlinear *HF equations*

$$(\hat{H}_\Psi \varphi_i)(x) = \sum_{j=1}^N e_{ij}\varphi_j(x), \qquad i = 1, \cdots, N, \tag{6}$$

where $\hat{H}_\Psi$ is the selfadjoint operator on $L^2(\boldsymbol{R}^3)$ given by

$$(\hat{H}_\Psi \varphi)(x) = \{-\Delta + V(x) + U_\Psi(x)\}\varphi(x) - (K_\Psi\varphi)(x),$$

$$U_\Psi(x) = \sum_{i=1}^N \int |\varphi_i(y)|^2 \, |x - y|^{-1} \, dy,$$

$$(K_\Psi\varphi)(x) = \sum_{i=1}^N \varphi_i(x) \int \varphi(y)\, \bar{\varphi}_i(y) \, |x - y|^{-1} \, dy.$$

Since $\hat{H}_\Psi$ is selfadjoint, e_{ij} is a selfadjoint matrix. Let T be the n-square unitary matrix that diagonalizes e_{ij}. The transformation $\varphi_i(x) \to \sum_{j=1}^N T_{ij}\varphi_j$ changes neither $\mathscr{E}_N(\Psi)$ nor $\hat{H}_\Psi$. Hence, we may assume $e_{ij} = e_i\delta_{ij}$. Also, the numbers $\{e_i\}$ are the N lowest eigenvalues of $\hat{H}_\Psi$; if f_j, the jth eigenvector, is missing for $j \leqq N$, replace f_{N+k} by f_j and note that $\mathscr{E}_N(\Psi)$ is lowered since $|x - y|^{-1}$ is a positive function on $\boldsymbol{R}^3 \times \boldsymbol{R}^3$.

Thus, if $\mathscr{E}_N(\Psi)$ has a minimum, the HF equations will have a solution. There is no claim of uniqueness this time.

THEOREM 3. *If $N < Z + 1$ then $\mathscr{E}(\Psi)$ has a minimum.*

The idea of the proof is this: Let $S = \{(\varphi_1, \cdots, \varphi_N) : M \leqq I, M_{ij} = (\varphi_i, \varphi_j)\}$. For $\Psi \in S$, let

$$\hat{\mathscr{E}}_N(\Psi) = \sum_{i=1}^N (\nabla\varphi_i, \nabla\varphi_i) + (\varphi_i, [V + \tfrac{1}{2}U_\Psi - \tfrac{1}{2}K_\Psi]\varphi_i)$$

and $\hat{E}_N = \inf\{\hat{\mathscr{E}}_N(\Psi) : \Psi \in S\}$. Let $\Psi^{(n)} \in S$ and $\hat{\mathscr{E}}_N(\Psi^{(n)}) \to \hat{E}_N$. Choose a subse-

quence such that each $\varphi_i^{(n)}$ converges weakly in H^1 to some φ_i and let $\Psi = (\varphi_1, \cdots, \varphi_N)$. Now $\Psi \to \hat{\mathscr{E}}_N(\Psi)$ is weakly lower semicontinuous, essentially because: (i) $U_\Psi - K_\Psi > 0$ and is bounded on $H^1(\boldsymbol{R}^3)$; (ii) V is a relatively form compact perturbation of $-\Delta$, so $(\varphi_i^{(n)}, V\varphi_i^{(n)}) \to (\varphi_i, V\varphi_i)$. Therefore Ψ minimizes $\hat{\mathscr{E}}_N(\Psi)$ and $\Psi \in S$. As before, the φ_i can be chosen so that $\hat{H}_\Psi \varphi_i = e_i \varphi_i$ and $(\varphi_i, \varphi_j) = \lambda_i \delta_{ij}$, $0 \leqq \lambda_i \leqq 1$. Clearly $\lambda_i = 0$ (resp. 1) if $e_i > 0$ (resp. $e_i < 0$), and to complete the proof one has to show that if n is the number of negative eigenvalues of $\hat{H}_\Psi$, then $n \geqq N$. The crucial remark in this regard is that, for large x, $V(x) + U(x) \sim (-Z + n)|x|^{-1}$, and this Coulomb tail can accommodate a bound state orthogonal to $\varphi_1, \cdots, \varphi_n$ when $Z > n$.

It can also be shown that the $\varphi_i(x)$ have exponential fall-off at infinity, i.e.,

$$|\varphi_i(x)| \leqq (\text{const}) \exp(-\lambda |x|) \quad \text{when } \lambda^2 < \min[-e_1, \cdots, -e_N].$$

Two important open questions about HF theory are these: (i) Can anything be said abou the uniqueness of the HF minimizing solution? (ii) Does $\mathscr{E}_N(\Psi)$ have a minimum when $N = Z + 1$? Perhaps not always, but it should in some cases.

The relationship of E_N^{HF} to E_N is the same as in the TF theory: $\lim_{N\to\infty} E_N^{\text{HF}}/E_N = 1$ if, as before, $Z_i = Nz_i$, $R_i = N^{-1/3} r_i$. This is so because a HF estimate, i.e., a particular choice of Ψ, was used to obtain the aforementioned upper bound to E_N and this bound was asymptotically (in N) equal to E_N.

In summary, the existence of solutions to the nonlinear TF and HF equations, (4) and (6), has been established by noting that these equations are gradients of certain functionals and by applying the Banach-Alaoglu theorem to those functionals. The equations have been studied by themselves [3]—[8], but the results obtained by such a direct attack have thus far not been as general as those given here.

References

1. E. H. Lieb and B. Simon, Phys. Rev. Lett. **31** (1973), 681; and papers (in preparation).
2. ———, J. Chem. Phys. **61** (1974), 735; and papers (in preparation).
3. G. Fonte, R. Mignani and G. Schiffrer, Comm. Math. Phys. **33** (1973), 293.
4. K. Gustafson and D. Sather, Rend. Mat. (6) **4** (1971), 723. MR **46** #9823.
5. E. Hille, J. Analyse Math. **23** (1970), 147. MR **43** #5098.
6. M. Reeken, J. Math. Phys. **11** (1970), 2505. MR **43** #5369; See also: Proc. R. Acad. Edinburgh (to be published).
7. C. Stuart, Battelle Geneva (preprint).
8. J. Wolkowsky, Indiana Univ. Math. J. **22** (1972/73), 551. MR **47** #2300.

PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08540, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Relations Between the Modulus and the Phase of Scattering Amplitudes

André Martin

1. Introduction. For physicists a very important object is the scattering amplitude which is at the interface between theory and experiment. Theories are supposed to predict scattering amplitudes. Experiment gives some information on scattering amplitudes. Now what kind of information? In the simplest case, the case of a collision between two particles of spin 0, the scattering amplitude is a complex function of two variables, which can be taken to be for instance the centre-of-mass energy and the scattering angle. What is now easily accessible to experimentalists is the *modulus* of the scattering amplitudes because the differential cross section is just the square of the modulus times some trivial factor. In elementary particle collisions the phase is seldom accessible (a notable exception is K_0 nucleon collisions). When I say that, I am taking a rather empirical point of view. I do not want to enter into a serious discussion of whether or not the phase is observable in some "gedanken" experiment (for a recent discussion of this, see Goldrich and Wigner [1]).

However, even if in practice the modulus is most of the time the only accessible quantity, the phase and the modulus are linked by very general relationships based on things that most physicists believe, which are:

(1) conservation of probability, which is technically called unitarity, and which implies in particular certain *positivity* properties;

(2) causality, more exactly microcausality as expressed by an underlying field theory; from this it follows that the physical scattering amplitude is the boundary value of an analytic function.

In the following, we shall give more details on (1) and (2). We shall give then two illustrations of the usefulness of these properties: (a) we shall see that perfect data

on the modulus of the scattering amplitude at all energies fix the phase; in some cases, it is enough to have measurement at *one* energy; (b) the general information we have on the phase allows us to make some statements on the high energy behaviour of the modulus, for instance to establish that the ratio of the moduli of the amplitudes for reactions $A + B \to A + B$ and $\bar{A} + B \to \bar{A} + B$ (where $\bar{A}$ designates the antiparticle of A) approaches unity or, more precisely, admits unity as one of its limiting values. This occurs for the forward scattering amplitude and also inside the region called the "diffraction peak".

2. Notations. The scattering amplitude $F(s, \cos\theta)$ for the reaction $A + B \to A + B$ is a function of two variables

$$s = \text{square of the centre-of-mass energy},$$
$$\theta = \text{centre-of-mass scattering angle}.$$

Other useful variables are

$$k = \text{centre-of-mass momentum}$$
$$((k^2 + M_A^2)^{1/2} + (k^2 + M_B^2)^{1/2})^2 = s,$$

$$t = \text{square of the momentum transfer}$$
$$t = -2k^2(1 - \cos\theta),$$

and, finally,

$$u = 2M_A^2 + 2M_B^2 - s - t.$$

The usefulness of u will appear later.

The scattering amplitude will be normalized if we say that $d\sigma/d\Omega$, differential cross section, is

$$\frac{d\sigma}{d\Omega} = \frac{4}{s}\,|F(s, \cos\theta)|^2.$$

3. Unitarity and positivity.

3.1. *The elastic region.* Here it may be useful to use a vector notation: $f(s, \boldsymbol{n}_1, \boldsymbol{n}_2)$ is the scattering amplitude from direction $\boldsymbol{n}_1$ to direction $\boldsymbol{n}_2$:

$$f(s, \boldsymbol{n}_1, \boldsymbol{n}_2) = \frac{2k}{s^{1/2}} F(s, \boldsymbol{n}_1 \cdot \boldsymbol{n}_2).$$

The unitarity condition reads

$$\text{Im} f(s, \boldsymbol{n}_1, \boldsymbol{n}_2) = \int \frac{d\Omega_3}{4\pi} f^*(s, \boldsymbol{n}_1, \boldsymbol{n}_3) f(s, \boldsymbol{n}_2, \boldsymbol{n}_3). \tag{1}$$

The imaginary part of the scattering amplitude is also called the absorptive part. In fact, the scattering amplitude can be expanded in partial waves

$$F(s, \cos\theta) = \sum (2l + 1) f_l(s) P_l(\cos\theta) \tag{2}$$

where the P_l's are Legendre polynomials. It has been proved from field theory, i.e., microcausality and existence of a minimum nonzero mass, that this expansion converges. Then the unitarity equation becomes

(3) $$\operatorname{Im} f_l(s) = \frac{2k}{s^{1/2}} |f_l(s)|^2$$

or

(4) $$f_l(s) = \frac{s^{1/2}}{2k} \exp(i\delta_l) \operatorname{Im} \delta_l.$$

Unitarity is expressed by the fact that the "phase shift" δ_l is a *real* quantity. The problem of obtaining the phase of the whole amplitude from the modulus is often called by physicists "phase shift analysis" because if one finds all the δ_l's one knows the amplitude entirely by summing the partial wave series. However, it is not necessary to go through the expansion. Equation (1) can be viewed as a nonlinear integral equation for the phase, once the modulus is known:

(5) $$f(\boldsymbol{n}_1, \boldsymbol{n}_2) = |f(\boldsymbol{n}_1, \boldsymbol{n}_2)| \exp i\phi(\boldsymbol{n}_1, \boldsymbol{n}_2),$$

$$|f(\boldsymbol{n}_1, \boldsymbol{n}_2)| \sin \phi(\boldsymbol{n}_1, \boldsymbol{n}_2)$$

(6) $$= \int \frac{d\Omega_3}{4\pi} |f(\boldsymbol{n}_1, \boldsymbol{n}_3)||f(\boldsymbol{n}_2, \boldsymbol{n}_3)| \cos[\phi(\boldsymbol{n}_1, \boldsymbol{n}_3) - \phi(\boldsymbol{n}_2, \boldsymbol{n}_3)].$$

We shall discuss later what can be extracted from this nonlinear equation.

3.2. *The inelastic region.* When the collision energy of A and B is sufficiently high other reactions compete with the elastic reaction $A + B \to A + B$. Then the expression of the conservation of probability is not so simple, and we do not have anymore an integral equation for the phase. However, one very simple and fundamental property survives: It is positivity which can be expressed as

(7) $$\int d\Omega_1 \, d\Omega_2 \, w(\boldsymbol{n}_1) \, w^*(\boldsymbol{n}_2) \operatorname{Im} f(s, \boldsymbol{n}_1, \boldsymbol{n}_2) \geqq 0,$$

or, if one prefers to use the partial wave amplitudes:

(8) $$\operatorname{Im} f_l(s) \geqq 0.$$

The most obvious consequence is

(9) $$\operatorname{Im} F(s, \cos\theta = 1) = \sum (2l + 1) \operatorname{Im} f_l(s) \geqq 0.$$

Therefore the phase of the forward amplitude stays between 0 and π. [The phase is defined by continuity. One can prove that the equality sign never happens in (9) unless the amplitude is identically zero!] However, a more subtle result has been obtained by Cornille and myself [2], which allows control of the phase at non-forward angles. Here one makes use of the positivity of the individual Im f_l's. One shows that

$$\left| \frac{\operatorname{Im} F(s, \cos\theta_1) - \operatorname{Im} F(s, \cos\theta_2) + \cdots + (-)^N \operatorname{Im} F(s, \cos\theta_N)}{\operatorname{Im} F(s, 1)} \right|$$

where $\theta_1, \theta_2, \cdots, \theta_N$ are ordered, $0 < \theta_1 < \theta_2, \cdots, < \theta_N$, is bounded by $C(N)^{1/2}$ where C can be calculated explicitly. With the help of this property, one deduces:

THEOREM I (CORNILLE-MARTIN). *The phase of the scattering amplitude at the angle θ, defined by continuity in the angle from $\theta' = 0$ to $\theta' = \theta$, is bounded by*

$$|\mathrm{Arg}\,(F(s,\cos\theta))| \leqq \pi + \frac{(\Gamma(\frac{1}{4}))^4}{\pi^3} \times \frac{|\mathrm{Im}\,F(s,\cos\theta = 1)|^2}{\min_{0\leqq\theta'\leqq\theta} |F(s,\cos\theta')|^2}.$$

The conclusion of this is that the knowledge of the differential cross section at a given energy allows us to get bounds on the phase.

A particular consequence is that "the phase is bounded inside the diffraction peak" [3]. The diffraction peak is defined as that angular interval in which, at high energies, the modulus of the scattering amplitude is a finite fraction of the forward amplitude. For instance, we could decide that the diffraction peak is that region where $d\sigma/d\Omega$ is larger than $1/e$ times the forward differential cross section; then the phase is bounded by $\pi + e(\Gamma(\frac{1}{4}))^4/\pi^3$ according to Theorem I.

4. Analyticity and the phase. We know that if we believe that particles are described by local fields of some kind, one can obtain analyticity properties of scattering amplitudes. In the simplest cases, like pion-nucleon scattering, these properties contain dispersion relations. In more complicated cases, the proof fails but Bros, Epstein and Glaser have found an excellent substitute [4]. For simplicity I shall assume that we are in a "good case" where dispersion relations hold. We use now as variables s and t. $F(s, t)$, for fixed t, $-T < t \leqq 4m_\pi^2$, is analytic in a cut plane in s. Its boundary value on the right-hand cut $F(s + i\varepsilon, t)$ is the physical scattering amplitude. The discontinuity across the cut is the imaginary part also called the absorptive part. The boundary value *below* the left-hand cut $F(s - i\varepsilon, t)$ is the physical scattering amplitude for the crossed process $\bar{A} + B \to \bar{A} + B$ where the square of the $(\bar{A}B)$ centre-of-mass energy is given by u defined in §2. This is the famous "crossing property". For $t \leqq 4m_\pi^2$ where m_π is the pion mass, the amplitude grows less fast than $|s|^2$ in the complex plane and it is possible therefore to write dispersion relations with two subtractions [5], i.e., to express the amplitude as an integral over the absorptive parts of reactions $A + B \to A + B$ and $\bar{A} + B \to \bar{A} + B$ plus two constants at fixed t.

In such a situation, we obtain a new link between the modulus and the phase, since the phase is conjugate to the logarithm of the modulus. However, analyticity alone is not very powerful because we can, for instance, introduce arbitrary zeros without changing the modulus on the cuts. This is very easily seen by mapping the twice cut plane on the unit circle. Then we know that multiplying by $(z - z_0)/(1 - zz_0^*)$ will not change the modulus on the circle.

One can, however, make statements in the other direction and get some restrictions on the behaviour of the modulus if $s \to \infty$. In particular we have:

THEOREM II (CORNILLE-MARTIN [6]). *Assume that for some fixed t the modulus of the scattering has no zero for sufficiently large, physical $|s|$ (i.e., $s \to +\infty$, $s \to -\infty$). Then it is meaningful to speak of the phase. Assume that the phases of the amplitude $A + B \to A + B$ and of the amplitude $\bar{A} + B \to \bar{A} + B$ are both $o(\log s)$. Then the set of limiting values of $|F(s,t)|/|F(-s,t)|$ for $s \to \infty$ contains unity.*

For physicists who believe that this limit exists, this means that, if the phases are $o(\log s)$ the ratio of the differential cross sections for reactions $A + B \to A + B$

and $\bar{A} + B \to \bar{A} + B$ approaches unity at infinite energy.

Now we shall see the power of the combination of unitarity with analyticity. From analyticity alone we have no way of knowing if the phase condition is satisfied. However, we find in §3.2 that the forward amplitude has a phase between 0 and π. So it is bounded and, for $t = 0$, the ratio of the differential cross sections for particles and antiparticles approaches unity (if it has a limit) at infinite energy [7].

In fact from Theorem I, this is true not only at $t = 0$ but also everywhere inside the diffraction peak. This implies that the widths of the diffraction peaks for particles and antiparticles tend to the same limit. A more refined version using two-variable analyticity can be used when the width of the diffraction peak goes to zero at infinite energy [2]. The conclusion is the same: The ratio of the widths must approach unity.

5. The phase problem.

5.1. *The phase problem at fixed energy.* Here we want to come back to the question of determining the phase from the modulus at a given energy at all angles, the energy being low enough so that no inelastic channels are present. Let me say that the problem is not completely solved. There exist sufficient conditions under which (6) has a unique solution [8] [modulo a trivial ambiguity $F(s, \cos\theta) \to -F^*(s, \cos\theta)$]. Such a condition is

$$\sup \frac{\int \frac{d\Omega_3}{4\pi} |f(\boldsymbol{n}_1, \boldsymbol{n}_3)|\,|f(\boldsymbol{n}_2, \boldsymbol{n}_3)|}{|f(\boldsymbol{n}_1, \boldsymbol{n}_2)|} \leqq .79. \tag{10}$$

This constitutes Theorem III.

Condition (10) is, from the point of view of physics, very strong and only satisfied in practice at very low energies. It can be shown that if any of the partial waves (except the S wave) is resonating (i.e., if δ_l is close to $\pi/2$) the condition is violated.

On the other hand, long ago, Crichton gave explicit examples of nontrivial ambiguities. More recent examples have been given [9]. These examples are polynomials in $\cos\theta$: *Two* different polynomials in $\cos\theta$ give exactly the same differential cross section. Atkinson and Johnson [10] have recently shown that these ambiguities occur also for the nonpolynomial case.

However, nobody has succeeded in producing more than a two-fold nontrivial ambiguity. The belief that this would be impossible is supported by:

THEOREM IV (ITZYKSON AND MARTIN [11]). *If the class of scattering amplitudes is restricted to nontrivial entire functions of finite order, there cannot be more than two amplitudes, differing in a nontrivial way, reproducing the same differential cross section.*

However, we know that physical scattering amplitudes are *not* entire functions in $\cos\theta$. The amplitudes produced by a finite range potential are entire functions, but this is not a physical, relativistic case. So the problem is open.

5.2. *The phase problem using various energies and angles.* Here I want to describe

some recent work by H. Burkhardt and myself [12] which is the continuation of an old paper by Bessis and myself [13].

If you assume that you know the modulus of the scattering amplitude at all energies and all angles, with perfect accuracy, the problem is overdetermined. Indeed you know the modulus at energies close to threshold, and then condition (10) is automatically fulfilled because the S wave ($l = 0$) dominates the scattering amplitude. Hence we have only to worry about the trivial ambiguity. This is easily removed, by using analyticity with respect to energy, together with the existence of a new singularity at the first inelastic threshold.

With Bessis first, in the case of a $\pi^0 \pi^0 \to \pi^0 \pi^0$ scattering amplitude and later with Burkhardt in the general case, we have tried another game: Assume that you know the modulus of the amplitude in two or three channels, say, $A + B \to A + B$ (s channel) and $\bar{A} + B \to \bar{A} + B$ (u channel), and try to find what is the minimum amount of unitarity that one should use to fix the scattering amplitude.

Here we want to present

THEOREM V (BURKHARDT-MARTIN). *Assume that particles A and B have equal masses m; assume that no unphysical threshold is present. Suppose you have two scattering amplitudes F and G such that F and G are analytic in the axiomatic analyticity domain* [5] *and that $|F| = |G|$ in the three physical regions corresponding to $A + B \to A + B$, $\bar{A} + B \to \bar{A} + B$, $A + \bar{A} \to B + \bar{B}$. Then*

$$\frac{F}{G} = \frac{\phi(s, t, u) + ((4m^2 - s)(4m^2 - t)(4m^2 - u))^{1/2}}{\phi(s, t, u) - ((4m^2 - s)(4m^2 - t)(4m^2 - u))^{1/2}}$$

where $\phi(s, t, u)$ is a real meromorphic function of s, t, u (with $s + t + u = 4m^2$), i.e., such that $\phi(s, t, u) = \phi^*(s^*, t^*, u^*)$.

It is almost obvious that if F/G has such a form, the ratio of the moduli in the physical regions is unity. The converse is less evident. Remarkably enough, one does not have to make extra analyticity assumptions, such as Mandelstam representation, to get this result. One sees that one has still an enormous freedom. However, as soon as we put in positivity requirements of the type described in §3.2, this freedom may disappear. For instance, positivity does not only imply that the *forward* ($t = 0$) scattering amplitude has a phase between 0 and π but this also holds for the amplitudes $A + B \to A + B$ and $\bar{A} + B \to \bar{A} + B$ for $0 \leqq t < 4m^2$. Therefore the amplitude $F(s, t)$ for $0 < t < 4m_\pi^2$ has a positive discontinuity on the right-hand cut and a negative continuity on the left-hand cut. This, combined with the fact that $F(s, t)$ grows less fast than s^2, makes it possible to show that F has at most two zeros in the complex s plane. Therefore, from the requirement of positivity in the s and u channels we deduce that the ratio of two amplitudes with the same modulus has at most two poles and two zeros for $0 < t < 4m_\pi^2$. That means

$$\frac{F}{G} = \frac{(A(t) + sB(t) + s^2 C(t))/(a(t) + sb(t)) + ((4m^2 - s)(4m^2 - t)(4m^2 - u))^{1/2}}{(A(t) + sB(t) + s^2 C(t))/(a(t) + sb(t)) - ((4m^2 - s)(4m^2 - t)(4m^2 - u))^{1/2}}.$$

The form of the meromorphic function is therefore enormously restricted.

If we have complete crossing symmetry as in the $\pi^0\pi^0 \to \pi^0\pi^0$ amplitude, we can go further. The only possible form of F/G is then

$$\frac{F}{G} = \frac{A + Bstu + C(s^2 + t^2 + u^2) + ((4m^2 - s)(4m^2 - t)(4m^2 - u))^{1/2}}{A + Bstu + C(s^2 + t^2 + u^2) - ((4m^2 - s)(4m^2 - t)(4m^2 - u))^{1/2}}$$

where A, B and C are constants.

However, with a few additional requirements, normal threshold behaviour, equality of the imaginary parts of F and G for $t = 0$ (which comes from the equality of total cross sections), one shows that $F/G \equiv 1$.

So in this case, only positivity inequalities have been sufficient to remove an ambiguity which corresponds to the choice of an arbitrary function of two variables.

References

1. F. Goldrich and E. P. Wigner, University of Baton Rouge, 1974 (reprint).
2. H. Cornille and A. Martin, Nuclear Phys. **B49** (1974), 413.
3. ———, Phys. Letters **40B** (1972), 671.
4. J. Bros, H. Epstein and V. Glaser, Comm. Math. Phys. **1** (1965), 240. MR **32** #2056.
5. A. Martin, Nuovo Cimento **42** (1966), 930; Y. S. Jin and A. Martin, Phys. Rev. **135** (1964), B1375. MR **30** #2840.
6. H. Cornille and A. Martin, Nuclear Phys. **B48** (1972), 104. This paper contains references to previous work, in particular that of M. Meiman.
7. A. Martin, Nuovo Cimento **39** (1965), 704. MR **33** #1031.
8. ———, Nuovo Cimento **59A** (1969), 131; R. G. Newton, J. Mathematical Phys. **9** (1968), 2050. MR **40** #2302.
9. J. H. Crichton, Nuovo Cimento **45A** (1966), 256; F. A. Berends and S. N. M. Ruijsenaars, Nuclear Phys. **B56** (1973), 507; H. Cornille and M. Drouffe, Nuovo Cimento **20A** (1974), 401.
10. D. Atkinson and M. Johnson, University of Groningen, 1973 (preprint). Their method uses the technique of D. Atkinson, G. Mahoux and F. Yndurain, Nuclear Phys. **B54** (1973), 263.
11. C. Itzykson and A. Martin, Nuovo Cimento **17A** (1973), 245.
12. H. Burkhardt and A. Martin, CERN preprint Th. 2017 (1975) submitted to Nuovo Cimento.
13. J. D. Bessis and A. Martin, Nuovo Cimento **52A** (1967), 719.

CERN
GENEVA, SWITZERLAND

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Markov Fields

Edward Nelson*

1. Introduction. Markov fields are a generalization to higher dimensions of the one-dimensional notion of Markov processes. Markov processes with appropriate symmetry properties are a useful tool in quantum mechanics, since analytic continuation in the time parameter yields the solution of the Schrödinger equation. Similarly, in the scalar case, Markov fields with Euclidean symmetry have been a useful tool in constructive quantum field theory (see [4]). Here we shall sketch some of the ideas which occur in extending this connection to the nonscalar case.

2. Markov fields. Let $\mathscr{V}$ be a finite-dimensional complex vector space, with dual $\mathscr{V}'$, and consider the vector bundle $\boldsymbol{E}^d \times \mathscr{V}$, where $\boldsymbol{E}^d$ is d-dimensional Euclidean space. (For the purposes of this section, any other C^∞ vector bundle would do as well.) Let $D(\boldsymbol{E}^d, \mathscr{V}')$ be the space of C^∞ functions with compact support in $\boldsymbol{E}^d$ taking values in $\mathscr{V}'$, topologized in the usual way, and let $D'(\boldsymbol{E}^d, \mathscr{V})$, the space of $\mathscr{V}$-valued distributions on $\boldsymbol{E}^d$, be its dual.

Let Σ be the σ-algebra of subsets of $D'(\boldsymbol{E}^d, \mathscr{V})$ generated by functions of the form $\varphi \mapsto \varphi(f)$ for f in $D(\boldsymbol{E}^d, \mathscr{V}')$. If Λ is a subset of $\boldsymbol{E}^d$, let Σ_Λ be the σ-algebra of all sets in Σ which are invariant under the action of all C^∞ diffeomorphisms of $\boldsymbol{E}^d$ which leave Λ pointwise fixed.

A $\mathscr{V}$-valued *Markov field* on $\boldsymbol{E}^d$ is a probability measure μ on $(D'(\boldsymbol{E}^d, \mathscr{V}), \Sigma)$ such that if Λ is an open set in $\boldsymbol{E}^d$ and u is a positive or integrable Σ_Λ-measurable random variable, then $E\{u \mid \Sigma_{\Lambda^c}\} = E\{u \mid \Sigma_{\partial\Lambda}\}$ where $\Lambda^c = \boldsymbol{E}^d \backslash \Lambda$ and $\partial\Lambda$ is the boundary of Λ, and $E\{\ \mid\ \}$ denotes the conditional expectation.

3. Euclidean covariance. Let $\mathscr{C}^d$ be the complex Clifford algebra of $\boldsymbol{E}^d$; that

*Support is gratefully acknowledged from the National Science Foundation, grant GP-33314, and from the Institute for Advanced Study.

is, the complex associative algebra with unit 1 which is generated by elements $\gamma \cdot x$, for x in $\boldsymbol{E}^d$, subject to linearity and the relation $\gamma \cdot x \, \gamma \cdot y + \gamma \cdot y \, \gamma \cdot x = 2x \cdot y$. Let $\tilde{O}(\boldsymbol{E}^d)$ be the group generated by all $\gamma \cdot x$ with $x \cdot x = 1$, and let $\Lambda(\gamma \cdot x)$ be the reflection in the hyperplane with unit normal x. Then Λ extends to be a homomorphism of $\tilde{O}(\boldsymbol{E}^d)$ onto the orthogonal group $\tilde{O}(\boldsymbol{E}^d)$ with kernel $\{1, -1\}$. We let $I\tilde{O}(\boldsymbol{E}^d)$ be the semidirect product $\tilde{O}(\boldsymbol{E}^d) \times \boldsymbol{E}^d$ with the multiplication law $(A, a)(B, b) = (AB, \Lambda(A)b + a)$.

If $\mathscr{V}$ is a finite-dimensional complex vector space and S is a continuous representation of $\tilde{O}(\boldsymbol{E}^d)$ on $\mathscr{V}$, then there is a natural action T of $I\tilde{O}(\boldsymbol{E}^d)$ on $D'(\boldsymbol{E}^d, \mathscr{V})$, and also a natural action S of $\tilde{O}(\boldsymbol{E}^d)$ on $D'(\boldsymbol{E}^d, \mathscr{V})$ which acts on the fiber $\mathscr{V}$ alone. A *Euclidean Markov field* of type S on $\boldsymbol{E}^d$ is a $\mathscr{V}$-valued Markov field such that the $T(A, a)$ for (A, a) in $I\tilde{O}(\boldsymbol{E}^d)$ are measure-preserving transformations (*Euclidean covariance*), and the $S(A)$ for A in $\tilde{O}(\boldsymbol{E}^d)$ are measure-preserving transformations (*fiber covariance*).

4. The transition semigroup. Let μ be a Euclidean Markov field of type S on $\boldsymbol{E}^d$, and let $\mathscr{K} = \mathscr{L}^2(\mu)$. Let $\mathscr{K}_0$ be the closed linear subspace of $\mathscr{K}$ consisting of all $\Sigma_{\boldsymbol{E}^{d-1}}$-measurable elements of $\mathscr{K}$, and let E_0 be the orthogonal projection operator from $\mathscr{K}$ to $\mathscr{K}_0$. Let

$$P^t u = E_0 T(1, -te_d)\, u, \quad u \in \mathscr{K}_0, 0 \leqq t < \infty,$$

where e_d is the unit vector perpendicular to $\boldsymbol{E}^{d-1}$. Then there is a unique positive selfadjoint operator K on $\mathscr{K}_0$ such that $P^t = e^{-tK}$. The proof that P^t is a contraction semigroup uses the Markov property (for half-spaces) and covariance under time translations in the familiar way, while the proof that P^t (and hence K) is self-adjoint makes use of both Euclidean covariance and fiber covariance under reflection.

5. Euclidean expectation values and their analytic continuation. Let $\bar{\mathscr{V}}$ be the complex conjugate vector space to $\mathscr{V}$, consisting of objects $\bar{u}$ for u in $\mathscr{V}$, with $\bar{u} + \bar{v} = \overline{u + v}$ and $\overline{\bar{a}u} = au$. If T is an endomorphism of $\mathscr{V}$ we define the endomorphism $\bar{T}$ of $\bar{\mathscr{V}}$ by $\bar{T}\bar{u} = \overline{Tu}$. We define $\bar{\varphi}(f)$, for f in $D(\boldsymbol{E}^d, \bar{\mathscr{V}}')$, by $\bar{\varphi}(f) = \overline{\varphi(\bar{f})}$. We adopt the general notational convention that if X is any object for which $\bar{X}$ is defined then $X^\varepsilon = X$ for $\varepsilon = i$ and $X^\varepsilon = \bar{X}$ for $\varepsilon = -i$, and if $X = (X_1, \cdots, X_n)$ and $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)$ then $X^\varepsilon = (X_1^{\varepsilon_1}, \cdots, X_n^{\varepsilon_n})$. Let $f = (f_1, \cdots, f_n)$ where f_j is in $D(\boldsymbol{E}^d, \mathscr{V}^{\varepsilon_j\prime})$ for $j = 1, \cdots, n$. If we assume that $\varphi^\varepsilon(f) = \varphi^{\varepsilon_1}(f_1) \cdots \varphi^{\varepsilon_n}(f_n)$ is integrable, then by the Schwartz nuclear theorem there is a unique $\mathscr{V}^\varepsilon$-valued distribution $\mathscr{S}_0^\varepsilon$ on $(\boldsymbol{E}^d)^n$ such that $\mathscr{S}_0^\varepsilon(f) = E\varphi^\varepsilon(f)$, where E denotes the expectation. Next we need a technical assumption, similar to Assumption (A) of [2] or the Osterwalder-Schrader growth condition [3], to ensure the possibility of analytically continuing $\mathscr{S}_0^\varepsilon$ with distribution boundary values $\mathscr{W}_0^\varepsilon$ on the n-fold product $(\boldsymbol{M}^d)^n$ of d-dimensional Minkowski space. Shorn of technicalities, this continuation is possible because the transition semigroup P^t admits a continuous extension to $\operatorname{Re} t \geqq 0$ which is holomorphic in the interior.

To state the basic properties of the $\mathscr{W}_0^\varepsilon$, we need some additional notation. The

double covering $\tilde{P}\uparrow_+$ of the restricted Poincaré group is imbedded in the complexification of $I\tilde{O}(\boldsymbol{E}^d)$, and the double covering $\tilde{L}\uparrow_+$ of the restricted Lorentz group is embedded in the complexification of $\tilde{O}(\boldsymbol{E}^d)$. The representation S extends to be holomorphic on $\mathscr{T}$, and $\bar{S}$ given by $\bar{S}(A) = S(\bar{A})$ is holomorphic on $\bar{\mathscr{T}}$ while $A \mapsto \overline{S(A)}$ is antiholomorphic on $\bar{\mathscr{T}}$.

We denote Fourier transforms by $\tilde{}$ and the closed forward light cone by V_+. By π_j we mean the permutation of $x = (x_1, \cdots, x_n)$ which interchanges x_{j-1} and x_j, by ε_{π_j} we mean the action of this permutation on $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)$, and by $\rho(\pi_j)$ we mean the corresponding linear map of $\mathscr{T}^{\varepsilon_{\pi_j}}$ to $\mathscr{T}^{\varepsilon}$.

The $\mathscr{W}_0^\varepsilon$ are tempered $\mathscr{T}^\varepsilon$-valued distributions on $(\boldsymbol{M}^d)^n$ with the following properties:

(a_0) Covariance:

$$S^\varepsilon(A)\, \mathscr{W}_0^\varepsilon(\Lambda(A, a)^{-1} x) = \mathscr{W}_0^\varepsilon(x), \qquad (A, a) \in \tilde{P}\uparrow_+(\boldsymbol{M}^d).$$

(b_0) Spectrality:

$$\tilde{\mathscr{W}}_0^\varepsilon(p_1, \cdots, p_n) = (2\pi)^d\, \delta\Big(\sum_{j=1}^n p_j\Big)\, \tilde{W}_0^\varepsilon(p_1, p_1 + p_2, \cdots, p_1 + p_2 + \cdots + p_{n-1})$$

where $\tilde{W}_0^\varepsilon(q_1, \cdots, q_{n-1}) = 0$ unless $q_j \in V_+$ for $j = 1, \cdots, n-1$.

(c_0) Symmetry:

$$\rho(\pi_j)\, \mathscr{W}_0^{\varepsilon_{\pi_j}}(x_{\pi_j}) = \mathscr{W}_0^\varepsilon(x), \qquad x_{j-1} - x_j \text{ space-like.}$$

(d_0) Positive definiteness:

$$\sum \int \cdots \int \bar{f}_m^*(y_1, \cdots, y_m)\, \mathscr{W}_0^{\bar{\delta}_m \cdots \bar{\delta}_1 \varepsilon_1 \cdots \varepsilon_n}(y_m, \cdots, y_1, x_1, \cdots, x_n) f_n(x_1, \cdots, x_n) \geqq 0.$$

6. Wightman distributions. At first sight, these properties (together with cluster decomposition, which we shall not discuss) look like the Wightman axioms for the vacuum expectation values of a quantum field. However, the covariance (a_0) is the wrong kind of covariance. In order for the adjoint field to transform correctly, Wightman distributions $\mathscr{W}^\varepsilon$ must satisfy

(a) Covariance:

$$S(A)^\varepsilon\, \mathscr{W}^\varepsilon(\Lambda(A, a)^{-1} x) = \mathscr{W}^\varepsilon(x), \qquad (A, a) \in \tilde{P}\uparrow_+(\boldsymbol{M}^d),$$

and in general $S^\varepsilon(A) \neq S(A)^\varepsilon$. (For scalar fields, S is the trivial representation and the above properties do suffice to construct a quantum field from a Euclidean Markov field—see [4].) To a field theorist, there is an even more glaring problem: The symmetry property (c_0) cannot hold for fields of nonintegral spin, $S(-1) = -1$, by the theorem on spin and statistics, and symmetry should be replaced by antisymmetry.

We shall very briefly sketch how to overcome these difficulties, for the two-point function. Suppose that $\varepsilon = (i, -i)$. We make the simplifying assumption that (b_0) is strengthened by assuming a mass gap. It can be shown that

$$\bar{A} = \gamma \cdot e_d\, A \gamma \cdot e_d, \qquad A \in \tilde{L}\uparrow_+.$$

Therefore $\mathscr{W}_1^\varepsilon = 1 \otimes (S(\gamma \cdot e_d))^* \mathscr{W}_0^\varepsilon$ has the correct covariance (a). However,

positive definiteness is now destroyed. We restore it, while maintaining covariance and spectrality, by using the projection given in momentum space by $\frac{1}{2}(1 + S(\gamma \cdot p/p \cdot p)^{1/2}))$. A study of symmetry leads to the projection $\frac{1}{2}(1 + S(-\gamma \cdot p/(p \cdot p)^{1/2}))$. The crucial point is that for integral spin, $S(-1) = 1$, this is the same projection operator whereas for nonintegral spin, $S(-1) = -1$, it is $\frac{1}{2}(1 - S(\gamma \cdot p/(p \cdot p)^{1/2}))$, which is a projection onto a negative subspace for $S(\gamma \cdot e_d)$. In this way we obtain the right connection between spin and statistics even though we started from classical commuting random variables in the Euclidean world.

Work extending this to the general n-point functions is nearing completion, and a complete account will be published elsewhere. The value of this approach will depend on its success in handling specific models, but we hope that in this way the techniques of classical statistical mechanics which Guerra, Rosen, and Simon, and others (see [4]) have applied to scalar fields will become available for quantum fields of higher spin, including Fermion fields.

In conclusion, I would like to urge probabilists to study Markov fields in their own right, aside from possible applications to quantum field theory. The central problem is the existence of non-Gaussian fields in higher dimensions. As Glimm and Jaffe have emphasized in their program of constructive quantum field theory (see e.g. [1] and the references therein), the central problem is the ultraviolet divergence. This problem remains, uncluttered by other complications, in the theory of Markov fields.

References

1. James Glimm, Arthur Jaffe and Thomas Spencer, *The particle structure of the weakly coupled $P(\varphi)_2$ model and other applications of high temperature expansions*, Lecture Notes in Physics, vol. 25, Constructive Quantum Field Theory, edited by G. Velo and A. Wightman, Springer-Verlag, Berlin, 1973.

2. Edward Nelson, *Construction of quantum fields from Markoff fields*, J. Functional Analysis **12** (1973), 97–112.

3. Konrad Osterwalder, *Euclidean Green's functions and Wightman distributions*, Lecture Notes in Physics, vol. 25, Constructive Quantum Field Theory, edited by G. Velo and A. Wightman, Springer-Verlag, Berlin, 1973.

4. Barry Simon, *The $P(\varphi)_2$ Euclidean (quantum) field theory*, Princeton Univ. Press, Princeton, N. J., 1974.

PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08540, U.S.A.

INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY 08540, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Approximation of Feynman Integrals and Markov Fields by Spin Systems

Barry Simon*

In this article I will discuss a similarity in the mathematical structures of two physically quite different classes of systems: the Markov processes associated with quantum mechanical anharmonic oscillators and field theories and the family of lattice models for ferromagnets. In fact, we will see that systems from the first class are limits of systems from the second class. This approximation is on two levels, the first due to Guerra, Rosen, and Simon [**1973**] and the second to Simon and Griffiths [**1973**]. These approximations, their extension and the development of the Ising model methods in constructive quantum field theory made available by them have been a major theme in constructive field theory during the past two and a half years. For a summary of applications up until January, 1974, I would refer you to my Zurich lectures(Simon [**1974**]). More recent work includes that of Glimm and Jaffe [**1974a,1974b**], Guerra, Rosen and Simon [**1974**], Newman [**1974**], and Spencer [**1974**].

Here I would like to describe the basic ideas of the approximation and illustrate their application by discussing the proof of the following result which is essentially due to Glimm, Jaffe, Spencer [**1973**].

THEOREM 1. *Let E_1, E_2, E_3 be the three smallest eigenvalues of the differential operator $H = -\frac{1}{2}(d^2/dx^2) + ax^2 + bx^4$ $(b > 0)$. Then*

$$E_3 - E_2 \geqq E_2 - E_1. \tag{1}$$

Perhaps the most interesting feature of Theorem 1 is that its proof is intimately connected with the fact that the magnetization in a ferromagnet induced by an external magnetic field is a concave function of the inducing field!

*A. Sloan Fellow, partly supported by NSF under grant GP-39048.

The basic systems. Let us begin by describing the basic systems which we will relate. The free Euclidean quantum field is the Gaussian random process $\phi(\cdot)$ indexed by $\mathscr{S}(\boldsymbol{R}^n)$ with mean zero and covariance

$$\int \phi(f)\,\phi(g)\,d\mu_0 = \langle f, (-\Delta + m^2)^{-1} g\rangle \tag{2}$$

where $\langle\ ,\ \rangle$ is an L^2 inner product. Due to work of Symanzik, Nelson and Osterwalder-Schrader, constructive quantum field theory is now concerned with constructing measures $d\nu = \lim_{\Lambda\to\infty} d\nu_\Lambda$ ($\Lambda \subset R^n$; compact) with

$$d\nu_\Lambda = \exp\left(-\int_\Lambda P(\phi(x))_{\text{ren}}\, d^n x\right) d\mu_0 \Big/ \text{Norm.} \tag{3}$$

where "Norm' represents a normalization factor chosen so that $\int d\nu_\Lambda = 1$. P is a polynomial bounded from below and "ren" indicates that when $n \geqq 2$ certain infinite subtractions are needed. As n increases, the local singularities of $\phi(x)$ become worse and, as a result, the renormalizations more complex. In any event, $d\nu_\Lambda$ has been defined in case $n = 2$ essentially due to work of Nelson and in case $n = 3$, deg $P = 4$, by recent work of Glimm and Jaffe and Feldman. In the physical case $n = 4$ (n is the number of space-time dimensions), there are still rather severe technical problems to be overcome in the definition of (3). (We expect that more details of this subject will be found in Glimm's and Nelson's contributions to these PROCEEDINGS.) When $n = 1$, there is a random process indexed by R, $q(t)$, with $\phi(f) = \int f(t)\,q(t)\,dt$ and this process is connected to the differential operator of Theorem 1 by the Feynman-Kac formula (for a proof, see, e.g., Simon [**1974**]):

THEOREM 2. *Let $n = 1$, $m = 1$ in* (2). *Let $P(x) = bx^4 + (a - \frac{1}{2})x^2$ and let $d\nu$ be the limit of the measures in* (3). *Let Ω_1 be the eigenvector of H normalized by $H\Omega_1 = E_1\,\Omega_1$, $\int \Omega_1(x)^2\,dx = 1$, $H \geqq E_1$. Then*

$$\int f(q(t))\,g(q(0))\,d\nu = \langle f(q)\Omega_1, \exp(-t(H - E_1))\,g(q)\,\Omega_1\rangle. \tag{4}$$

A classical Ising model is a probability measure on $\{-1, 1\}^N$ of the form

$$d\alpha = \exp\left(\sum_{i\neq j} a_{ij}\sigma_i\sigma_j + \sum_i \mu_i\sigma_i\right)\Big/\text{Norm} \tag{5}$$

where $\sigma_i = \pm 1$ are coordinates for $\{-1, 1\}^N$ thought of as the values of N spins which can point up ($\sigma_i = +1$) or down ($\sigma_i = -1$). The μ_i represent external magnetic fields and the measure α is called ferromagnetic if $a_{ij} \geqq 0$, all $i \neq j$, in which case there is a tendency for the spins to align in parallel. The study of Ising models has been much influenced by certain inequalities involving expectations of σ's and their products. The first of these "correlation inequalities" was proved by Griffiths [**1967**] and the subject has been developed extensively by Ginibre, Griffiths, Hurst, Kelley, Lebowitz, Percus, and Sherman. In particular, Theorem 1 depends on the following inequality of Lebowitz [**1974**] (the proof may be found in the original paper or in Simon [**1974**]):

THEOREM 3. *If $\langle\cdot\rangle$ denotes expectation with respect to a measure $d\alpha$ of the form* (5), *then whenever $a_{ij} \geqq 0$, $\mu_i = 0$,*

$$\langle\sigma_i\sigma_j\sigma_k\sigma_l\rangle \leqq \langle\sigma_i\sigma_j\rangle\langle\sigma_k\sigma_l\rangle + \langle\sigma_i\sigma_k\rangle\langle\sigma_j\sigma_l\rangle + \langle\sigma_i\sigma_l\rangle\langle\sigma_j\sigma_k\rangle. \tag{6}$$

To understand the physical significance of (6), we note that

$$\text{LHS} - \text{RHS} = \partial^3\langle\sigma_i\rangle/\partial\mu_j\partial\mu_k\partial\mu_l. \tag{7}$$

Now $\langle\sigma_i\rangle$ represents the magnetization of spin i in the external field $\{\mu_j\}$. The expected concavity of $\langle\sigma_i\rangle$ as a function of external fields is expressed by the Griffiths-Hurst-Sherman inequalities: If $\mu_i > 0$,

$$\partial^2\langle\sigma_i\rangle/\partial\mu_j\partial\mu_k \leqq 0. \tag{8}$$

Since the derivative in (8) is zero if all $\mu_i = 0$ (by $\mu_i \leftrightarrow -\mu_i$, $\sigma_i \leftrightarrow -\sigma_i$ symmetry), (8) imples that the derivative (7) is negative if all $\mu_i = 0$; this is the inequality (6).

We will also need to discuss generalized Ising models where $d\alpha$ on $\{-1, 1\}^N$ is replaced by a measure $d\beta$ on R^N with

$$d\beta = \exp\left(-\sum b_{ij}\sigma_i\sigma_j + \sum \mu_i\sigma_i\right)\prod_{i=1}^{N} d\gamma_i(\sigma_i)\Big/\text{Norm} \tag{9}$$

where $d\gamma_i$ is a finite measure on R and b_{ij} is a strictly positive definite matrix. The ferromagnetic case is the one with $b_{ij} \leqq 0$ for $i \neq j$. Certain correlation inequalities hold for expectations with respect to β but inequalities as detailed as (6) do not hold without additional restrictions on the $d\gamma$.

The approximation theorems. The basic approximation theorem of Guerra, Rosen Simon [**1973**] is:

THEOREM 4. *The measures $d\mu_\Lambda$ of form* (3) *are the limits of generalized ferromagnetic Ising models of form* (9) *in case* $n = 1, 2$. *The measures $d\gamma$ in* (9) *are of the form $d\gamma(\sigma) = \exp(-Q(\sigma))d\sigma$ where $Q = (\text{const})P +$ quadratic term.*

Guerra, Rosen, and Simon present a general scheme (the "lattice approximation") which formally approximates $d\mu_\Lambda$ and prove its convergence in case $n \leqq 2$. Recently, Park [**1974**] has proven convergence of this scheme in case $n = 3$, deg $P = 4$. The meaning of Theorem 4 is the following: For each $f \in \mathscr{S}(\boldsymbol{R}^n)$, $\phi(f)$ is a limit of suitable linear combinations of the σ_i in the approximating theory and, if $f \geqq 0$, the coefficients in this linear approximation are positive. Thus multilinear inequalities on expectations (such as (6)) carry over to measures $d\mu_\Lambda$. The basic idea of the proof of Theorem 4 is to replace $\boldsymbol{R}^n$ by a lattice $\boldsymbol{Z}^n\delta$, $\phi(f) \cong \sum_{m\in\boldsymbol{Z}^n} f(m\delta)\sigma_{m,\delta}$ and the Laplacian in (2) by a finite difference approximation. Since the *inverse* covariance matrix appears in the Gaussian for a joint probability distribution, $d\mu_0$ is approximated by general Ising models with $d\gamma$ Gaussian and with $b_{ij} \leqq 0$ $(i \neq j)$ because $-\Delta$ has finite difference approximations which are negative off-diagonal. $d\nu_\Lambda$ is then approximated by

$$\prod_{m\delta\in\Lambda} \exp\left(-\delta^n P(\sigma)_{\text{ren}}\right) d\mu_0$$

so only the $d\gamma$'s are affected by the change from $d\mu_0$ to $d\nu_\Lambda$.

The basic approximation theorem of Griffiths and Simon [**1973**] is

THEOREM 5. *A generalized Ising model of the form* (9) *with $d\gamma_i(\sigma) = \exp(-b_i\sigma_i^4 - a\sigma_i^2)d\sigma$ $(b_i > 0)$ is a limit of classical Ising models of the form* (5).

To explain the idea of their proof, take $N = mk$ in (5), replace σ_i by $s_{\alpha;r}$, $\alpha = 1, \cdots, m$, $r = 1, \cdots, k$ and let a_{ij} be of the form:

$$a_{\alpha r;\beta t} = -b_{rt}/m \qquad (b_{rr} = 0).$$

Then

$$\sum_{i \neq j} a_{ij} s_i s_j = -\sum_{r,t} b_{rt} \left(\sum_{\alpha} m^{-1/2} s_{\alpha;r} \right) \left(\sum_{\beta} m^{-1/2} s_{\beta,t} \right) + \text{constant}$$

so by the central limit theorem $\sigma = m^{-1/2} \sum s_{\alpha;r}$ will approach a generalized Ising model with $d\gamma$ Gaussian. By adding a term const$(1/m)\delta_{rt}$ to a we can cancel the Gaussian limit and by rescaling (i.e., taking $\sigma_r = m^{-3/4} \sum s_{\alpha;r}$) get the quartic limit.

The Griffiths-Simon theorem has recently been extended to the approximation of multicomponent fields by multicomponent Ising models (plane rotor and classical Heisenberg models) by Dunlop and Newman [**1974**].

PROOF OF THEOREM 1. By application of Theorems 3, 4, 5, the path space expectation $d\nu_\Lambda$ with $n = 1$, $P(x) = bx^4 + (a - \frac{1}{2})x^2$ obeys:

$$\langle q(t_1)\, q(t_2)\, q(t_3)\, q(t_4) \rangle \leqq \langle q(t_1)\, q(t_2) \rangle \langle (q(t_3)\, q(t_4) \rangle + 2 \text{ others};$$

so by taking $\Lambda \to \infty$, letting $t_1 = t_2 = 0$, $t_3 = t_4 = t$ and using Theorem 2:

$$\langle q^2 \Omega_1, \exp(-t(H - E_1))\, q^2 \Omega_1 \rangle - \langle q^2 \Omega_1, \Omega_1 \rangle^2 \leqq 2 \langle q\Omega_1, e^{-t(H-E_1)}\, q\Omega_1 \rangle^2. \tag{10}$$

Letting $\Omega_1, \Omega_2, \cdots$ be the eigenfunction of H, (10) says that

$$\sum_{m \geqq 2} |\langle q^2 \Omega_1, \Omega_m \rangle|^2 \exp(-t(E_m - E_1)) \leqq 2 \left[\sum_{m \geqq 1} |\langle q\Omega_1, \Omega_m \rangle|^2 \exp(-t(E_m - E_1)) \right]^2. \tag{11}$$

Since Ω_m is an even (odd) function of q for m odd (even), $\langle q^2 \Omega_1, \Omega_2 \rangle = 0$, $\langle q\Omega_1, \Omega_1 \rangle = 0$. Moreover, since Ω_m has $m - 1$ nodes, $\langle q\Omega_1, \Omega_2 \rangle \neq 0$, $\langle q^2 \Omega_1, \Omega_3 \rangle \neq 0$. Thus as $t \to \infty$, the leading behavior of the LHS and RHS respectively of (11) is $|\langle q\Omega_1, \Omega_3 \rangle|^2 \exp(-t(E_3 - E_1))$ and $|\langle q\Omega_1, \Omega_2 \rangle|^4 \exp(-2t(E_2 - E_1))$. It follows that $E_3 - E_1 \geqq 2(E_2 - E_1)$ completing the proof.

References

F. Dunlop and C. Newman (**1974**), Comm. Math. Phys. (to appear).

J. Glimm and A. Jaffe (**1974a**), Phys. Rev. **D10** (1974), 536–539.

J. Glimm and A. Jaffe (**1974b**), Phys. Rev. Lit. **33** (1974), 440–442.

J. Glimm, A. Jaffe and T. Spencer (**1973**), in *Constructive Quantum Field Theory*, edited by G. Velo and A. S. Wightman, Springer-Verlag, Berlin, 1973, pp. 133–198.

R. Griffiths (**1967**), J. Mathematical Phys. **8** (1967), 478–489.

F. Guerra, L. Rosen, and B. Simon (**1973**), Ann. of Math. **101** (1975), 111–259.

F. Guerra, L. Resen and and B. Simon (**1974**), Comm. Math. Phys. **41** (1975), 19–32.

J. Lebowitz (**1974**), Comm. Math. Phys. **35** (1974), 87–92.

C. Newman (**1974**), paper on Hadamard's Factorization Theorem and the Lee-Yang theorem (in preparation).

B. Simon (**1974**), *The $P(\phi)_2$ Euclidean (quantum) field theory*, Princeton Univ. Press, Princeton, N. J., 1974.

B. Simon and R. Griffiths (**1973**), Comm. Math. Phys. **11** (1973), 145–164.

PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08540, U.S.A.

Section 17

Numerical Mathematics

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Convergence in the Maximum Norm of Spline Approximations to Elliptic Boundary Value Problems

James H. Bramble

The subject matter in this article is based upon joint work with A. H. Schatz. A somewhat amplified version of this summary may be found in [3] and detailed proofs will be published elsewhere.

Let Ω be a bounded domain in Euclidean N-space R^N with smooth boundary $\partial\Omega$. For u a real valued function defined on Ω we shall consider the uniformly elliptic second order differential operator

$$Lu = -\sum_{i,j=1}^{N} \frac{\partial}{\partial x_j}\left(a_{ij}\frac{\partial u}{\partial x_i}\right) + cu$$

where a_{ij} and c are assumed smooth. The associated bilinear form is given by

$$B(v, w) = \sum_{i,j=1}^{N} \int_{\Omega} a_{ij}\frac{\partial v}{\partial x_i}\frac{\partial w}{\partial x_j}\,dx + \int_{\Omega} cvw\,dx.$$

For $f \in \mathscr{L}_2(\Omega)$, a weak solution u of

$$Lu = f \quad \text{in } \Omega \tag{1}$$

satisfies

$$B(u, \varphi) = \int_{\Omega} f\varphi\,dx \doteq (f, \varphi) \tag{2}$$

for all functions φ which are continuous and piecewise continuously differentiable in Ω and which vanish near $\partial\Omega$. We can associate with (1) various kinds of boundary conditions. Examples of these are

(a) $u = 0$ on $\partial\Omega$, or

(b) $\partial u/\partial\nu = 0$ on $\partial\Omega$, or

(c) $\partial u/\partial \nu + u = 0$ on $\partial\Omega$; where

$$\partial u/\partial \nu = \sum_{i,j=1}^{N} a_{ij} \frac{\partial u}{\partial x_j} n_i.$$

Here n_i is the component in the direction x_i of the outward normal to $\partial\Omega$ and $\partial/\partial\nu$ is called the conormal derivative.

Let S_h be a linear space of "finite elements" and $u_h \in S_h$ an approximate solution to the boundary value problem (1) with (a), (b), or (c) satisfied. For many different finite element methods proposed for these problems the interior equations are as follows:

$$B(u_h, \varphi) = (f, \varphi) \tag{3}$$

for all $\varphi \in S_h$ which vanish near $\partial\Omega$. In many different specific methods which have (3) in common, estimates for norms of the error $u - u_h$ in Sobolev spaces on all of Ω are well known. (For a summary of such results cf. [2].) Interior estimates in Sobolev norms for $u - u_h$ only satisfying (3) were given in [6] and in maximum norm in [1].

Here we shall consider, instead of u_h as an approximation to u, certain "local averages of u_h". These, as will be seen subsequently, are formed by computing $K_h * u_h$, where K_h is a fixed function and $*$ denotes convolution. As we shall see, the function K_h has the following properties:

(i) K_h has small support.

(ii) K_h is "independent" of the specific choice of S_h or the operator L.

(iii) $K_h * u_h$ is easily computable from u_h.

(iv) $K_h * u_h$ better approximates u than does u_h.

We shall need some standard notation. To this end denote by $C^s(\Omega)$, $s = 0, 1, 2, \cdots$, the space of functions defined on Ω with uniformly continuous partial derivatives of order up to and including s on Ω. For $v \in C^s(\Omega)$ we set

$$|v|_{s,\Omega} = \sup_{|\alpha| \leqq s;\, x \in \Omega} |D^\alpha v(x)|,$$

where α is a multi-index, $|\alpha| = \sum_{i=1}^{N} \alpha_i$ and $D^\alpha = \partial^{\alpha_1}/\partial x_1^{\alpha_1} \cdots \partial^{\alpha_N}/\partial x_N^{\alpha_N}$. For s real we define $H^s(\Omega)$, the Sobolev space with index s and, for $v \in H^s(\Omega)$, $\|v\|_{s,\Omega}$ will denote its norm (cf. [5]). For example, for $s = 0, 1, 2, \cdots$, $\|\cdot\|_{s,\Omega}$ is given by

$$\|v\|_{s,\Omega} = \left(\sum_{|\alpha| \leqq s} \int_\Omega |D^\alpha v|^2 \, dx \right)^{1/2}.$$

For s a positive noninteger $H^s(\Omega)$ may be defined by interpolating between successive integers and for $s < 0$ by duality (cf. [5]).

The one-parameter family of spaces $\{S_h\}$, $0 < h \leqq 1$, which we shall consider, will be assumed to have the following properties.

(i) For each h, $S_h \subset H^1(\Omega)$ and S_h is finite-dimensional.

(ii) For $x \in \Omega_1 \subset\subset \Omega$ and $U \in S_h$ there are functions $\varphi_1, \cdots, \varphi_k$ which are piecewise polynomials with compact support such that

$$U(x) = \sum_{j=1}^{k} \sum_{\alpha \in Z^N} a_\alpha^j \varphi_j(h^{-1} x - \alpha).$$

Here $\Omega_1 \subset\subset \Omega$ means $\bar{\Omega}_1 \subset \Omega$, a_α^j are real coefficients and Z^N are the multi-integers. (This property may be described as an interior translation invariance property.)

(iii) For some positive integer r there is a constant C such that, for $v \in H^s(\Omega)$, $1 \leqq s \leqq r$,

$$\inf_{\varphi \in S_h} (\|v - \varphi\|_{0,\Omega} + h\|v - \varphi\|_{1,\Omega}) \leqq Ch^s \|v\|_{s,\Omega}.$$

(iv) Let $\Omega_1 \subset\subset \Omega$ and let w be an infinitely differentiable function with support in Ω_1. There is a constant C such that for, $v \in S_h$,

$$\inf_{\varphi \in S_h; \operatorname{supp} \varphi \subset \Omega} \|wv - \varphi\|_{1,\Omega_1} \leqq Ch\|v\|_{1,\Omega}.$$

It was shown in [1] that subspaces consisting of tensor products of one-dimensional splines on a uniform mesh have all the requisite properties. Also in [6] it was demonstrated that the triangular element subspaces in R^2 defined in [4] are examples satisfying the above four conditions provided the triangulation is uniform. We emphasize that the uniformity is a condition which is only required locally. Thus we see that many of the finite element subspaces which are discussed in the literature satisfy the above conditions.

In order to define the function K_h we shall need to introduce the so-called smooth splines. In fact we shall choose K_h to be a particular smooth spline depending on the index r associated with the subspace S_h.

For t real define

$$\chi(t) = \begin{cases} 1, & |t| \leqq \frac{1}{2}, \\ 0, & |t| > \frac{1}{2}, \end{cases}$$

and for $x \in R^N$ set $\phi(x) = \prod_{j=1}^N \chi(x_j)$. For l a positive integer set $\phi^{(l)}(x) = (\phi * \cdots * \phi)(x)$, $(l-1)$ times. The function $\phi^{(l)}$ is just the N-dimensional B-spline of Schoenberg [7]. The space of smooth splines of order l on a mesh of width h consists of all functions of the form

$$U(x) = \sum_{\alpha \in Z^N} a_\alpha \phi^{(l)}(h^{-1}x - \alpha),$$

for some coefficients a_α.

The proof of the following will be given in a forthcoming paper by the author and A. H. Schatz.

PROPOSITION. *Let l and t be two given positive integers. The smooth spline*

$$K_h(x) = \sum_{\alpha \in Z^N} k_\alpha \phi^{(l)}(h^{-1}x - \alpha)$$

may be chosen so that

(a) $k_\alpha = 0$ *when* $|\alpha_j| > t - 1$ *for some* j,

(b) *for* $\Omega_0 \subset\subset \Omega_1$ *and* $v \in C^{2t}(\Omega_1)$ *there is a constant* C *such that*

$$|v - K_h * v|_{0,\Omega_0} \leqq Ch^{2t} |v|_{2t,\Omega_1}, \text{ and}$$

(c) *for* $v \in H^{2t}(\Omega_1)$ *there is a constant* C *such that*

$$\|v - K_h * v\|_{0,\Omega_0} \leqq Ch^{2t}\|v\|_{2t,\Omega_1}.$$

This function K_h is the aforementioned function in terms of which our local averages will be defined.

Let us denote by $\mathring{S}_h(\Omega_1)$ the subspace of S_h whose elements consist of functions in S_h with support in Ω_1. Let us suppose now that $u_h \in S_h$ satisfies $B(u - u_h, \varphi) = 0$ for all $\varphi \in \mathring{S}_h(\Omega_1)$. These equations are the same as (3) provided $Lu = f$. We now state our main result. The proof of this will be given in a forthcoming paper by the author and A. H. Schatz.

THEOREM. *Let $\Omega_0 \subset\subset \Omega_1 \subset\subset \Omega$ and p be an arbitrary but fixed real number. Let $l = r - 2$ and $t = r - 1$. Then there is a constant C such that, for $u \in H^{2r-2}(\Omega_1)$,*

$$\|u - K_h * u_h\|_{0,\Omega_0} \leqq C\{h^{2r-2}\|u\|_{2r-2,\Omega_1} + \|u - u_h\|_{-p,\Omega_1}\}$$

and, for $u \in H^{2r-2+\lfloor N/2\rfloor+1}(\Omega_1)$,

$$|u - K_h * u_h|_{0,\Omega_0} \leqq C\{h^{2r-2}\|u\|_{2r-2+\lfloor N/2\rfloor+1,\Omega_1} + \|u - u_h\|_{-p,\Omega_1}\}.$$

Let us consider some examples in order to illustrate the meaning of this result. Let $S_h^{(r)}$ consist of the smooth splines of order r (restricted to Ω). Then for $\varphi \in S_h^{(r)}$ we see that $K_h * \varphi \in S_h^{(2r-2)}$. It is known that, for some $U_h \in S_h^{(2r-2)}$, $u - U_h = O(h^{2r-2})$ as $h \to 0$ for smooth u. The theorem says that in fact the special smooth spline $K_h * u_h$ is such that $u - K_h * u_h = O(h^{2r-2})$ as $h \to 0$ in the interior of Ω provided that, for some p,

$$\|u - u_h\|_{-p,\Omega_1} = O(h^{2r-2}). \tag{4}$$

Let us consider a case where (4) is known to be true for $p = r - 2$. Let $c > 0$ in the operator L. Then the solution of the Neumann problem $Lu = f$ in Ω, $\partial u/\partial\gamma = 0$ on $\partial\Omega$ satisfies $B(u, \varphi) = (f, \varphi)$ for all $\varphi \in H^1(\Omega)$. The solution $u_h \in S_h^{(r)}$ of $B(u_h, \varphi) = (f, \varphi)$ for all $\varphi \in S_h^{(r)}$ exists and is unique. As may be found in [2] the estimate

$$\|u - u_h\|_{2-r,\Omega_1} \leqq \|u - u_h\|_{2-r,\Omega} \leqq Ch^{2r-2}\|u\|_{r,\Omega}.$$

More particularly if we choose $r = 4$ (cubic splines) we obtain, for $N = 2$,

$$|u - K_h * u_h|_{0,\Omega_0} \leqq Ch^6\{\|u\|_{8,\Omega_1} + \|u\|_{4,\Omega}\}.$$

Hence if u is locally smooth and globally less smooth ($u \in H^8(\Omega_1) \cap H^4(\Omega)$) we see that $K_h * u_h$ is a local 6th order approximation to u while u_h itself is in general only a 4th order approximation to u.

We emphasize that S_h need not be chosen to be the smooth splines (locally) but may be chosen from a much larger class of approximating subspaces of $H^1(\Omega)$.

References

1. J. H. Bramble, J. A. Nitsche and A. H. Schatz, *Interior maximum-norm estimates for Ritz-Galerkin methods*, Math. Comp. **29** (1975).

2. J. H. Bramble and J. E. Osborn, *Rate of convergence estimates for nonselfadjoint eigenvalue approximations*, Math. Comp. **27** (1973), 525–549.

3. J. H. Bramble and A. H. Schatz, *Higher order local accuracy by averaging in the finite element method*, Sympos. on Mathematical Aspects of the Finite Element Method, Madison, Wisconsin (editor, C. de Boor), April 1974.

4. J. H. Bramble and M. Zlámal, *Triangular elements in the finite element method*, Math. Comp. **24** (1970), 809–820. MR **43** #8250.

5. J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications*. Vol. I, Travaux et Recherches Math. no. 17, Dunod, Paris, 1968. MR **40** #512.

6. J. A. Nitsche and A. H. Schatz, *Interior estimates for Ritz-Galerkin methods*, Math. Comp. **28** (1974), 937–958.

7. I. J. Schoenberg, *Contributions to the problem of approximation of equidistant data by analytic functions*. Part A. *On the problem of smoothing or graduation. A first class of analytic approximation formulae*, Quart. Appl. Math. **4** (1946), 45–99. MR **7**, 487.

CORNELL UNIVERSITY
ITHACA, NEW YORK 14853, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

A Survey of Recent Progress in Approximation Theory

E. W. Cheney

Because of the limitations of time, this article cannot survey the entire field of approximation theory, but instead limits itself to certain aspects of nonlinear best approximation in the real domain. A more comprehensive survey including an extensive bibliography is available from the Center for Numerical Analysis, University of Texas, Austin, Texas 78712.

The nonlinear theory of best approximation has been cultivated in two ways which are interrelated. First, there has been an intensive investigation of particular approximating families, such as the rational functions and exponential polynomials. Second, there have been investigations of broad classes of nonlinear families, such as the Υ-polynomials and varisolvent families. The study of particular examples has, of course, served to direct the general theory into suitable and productive directions. Since the particular examples of approximating functions often arise from problems of science and technology, they provide proper motivation for the entire subject.

Typically, the problems of best approximation conform to the following outline. One prescribes a compact Hausdorff space X, which is often an interval on the real line. One denotes by $C(X)$ the space of continuous real-valued functions on X, normed by putting $\|f\| = \max_{x\in X}|f(x)|$. Next, a subset G is fixed in $C(X)$. The elements of G are the "approximants". For any $f \in C(X)$, the distance between f and G is the number $\operatorname{dist}(f, G) = \inf_{g\in G}\|f - g\|$. A complete theory would then attempt to answer the following eight questions. (1) For which f in $C(X)$ does there *exist* a best approximation in G? (2) Are best approximations in G *unique*? (3) How can $\operatorname{dist}(f, G)$ be *estimated* from only gross knowledge of f? (4) What properties distinguish a best approximation to f from all the other elements of G? (5) What *algorithms* can be devised for obtaining best approximations out of G for arbitrary

elements of $C(X)$? (6) If a "proximity map" A is defined from $C(X)$ to G by the requirement $\|f - Af\| = \operatorname{dist}(f, G)$ for all f, what is the behavior of A? (In particular, is it continuous?) (7) Do there exist well-behaved and convenient maps which are "near" A? (8) What specific approximations from G to special functions are there, and are these competitive with other possible types of approximation?

One of the first tasks in nonlinear approximation theory is to classify the types of approximating families. An ad hoc classification, based upon current research activity, contains the following ten categories (not mutually exclusive): (1) Generalized rational functions, (2) unisolvent families, (3) asymptotically convex families, (4) varisolvent families, (5) Υ-polynomials, (6) exponential polynomials, (7) constrained linear families, (8) splines with variable knots, (9) approximation from manifolds, (10) everything else.

Historically, the first important example of a nonlinear approximating family is the rational function class R_m^n. This consists of quotients p/q, in which p and q are polynomials of degrees $\leqq n$ and $\leqq m$ respectively. The theory here is well developed for the uniform norm, but work continues for L_p-norms. In the uniform case, there has been a flurry of activity on the subject of algorithms. Specifically, the differential correction algorithm, proposed in 1962 by H. L. Loeb and myself, has been proved to be quadratically convergent, and hence competitive with the Remez algorithm. The references are Dua and Loeb (SIAM J. Numer. Anal. **10** (1973)), Barrodale, Powell and Roberts (SIAM J. Numer. Anal. **9** (1972)), and Lee and Roberts (Math. Comp. **27** (1973)). Here, as in many nonlinear situations, the numerical problems can be blamed directly on the discontinuity of the proximity map defined above. It is known that the proximity map $A: C[0, 1] \to R_m^n$ is discontinuous at f if and only if $Af \in R_{m-1}^{n-1}$.

Rational approximation with other norms, particularly the L_p-norms, is not so well understood. Best approximations are generally not unique, and the quantity $\|f - r\|$ may have local minima as r ranges over R_m^n, if the norm is an L_p-norm. As a general rule in all of approximation theory, it seems that the advantages of L_2 approximation exist only for linear approximation, while the nonlinear theory seems to be more satisfactory in the L_∞-case. References for recent work are Dunham (J. Approximation Theory **10** (1974)), Wuytack (J. Approximation Theory **9** (1973)), and Lamprecht (Computing **5** (1970)).

The study of exponential approximation begins with functions of the form $a_1 \exp(\lambda_1 x) + \cdots + a_n \exp(\lambda_n x)$ with the a's and λ's variable. An immediate problem arises over existence of best approximation, since the family is not closed. For example, the function xe^x is in the closure of the family but not in the family. The uniform closure of the family consists of all functions which can be expressed in the form $\sum_{i=1}^{l} p_i(x) \exp(\lambda_i x)$, with p_i being polynomials whose degrees d_i satisfy $\sum_{i=1}^{l}(1 + d_i) \leqq n$. This enlarged family is denoted by V_n. The existence of best approximations from V_n was proved by Werner (Oberwolfach (1967), MR **40** #7689). Existence is not a triviality since closed and bounded sets in V_n need not be compact. For example, $e^{a(x-1)}$ on $[0, 1]$ converges to a discontinuous function as $a \to \infty$. The characterization of best approximations is due to Braess (Com-

puting **2** (1967)). Kammler (J. Approximation Theory **9** (1973)) has studied exponential sums by means of a different parametrization. For any $2n$ complex numbers $b_1, \cdots, b_n, c_1, \cdots, c_n$ there exists a unique function g which solves the initial value problem $\sum_1^n c_i D^{n-i} g(t) = 0$, $D^i g(0) = b_i$, $1 \leqq i \leqq n$. The mapping Φ_n so defined from $\boldsymbol{C}^{2n}$ into $C[0, 1]$ (complex functions) turns out to be a continuously differentiable homeomorphism when restricted to a suitable set of Baire category II. From this theorem, many existence theorems about constrained exponential approximation can be proved.

The two special classes that I have described so far are contained in one of the general classes, namely the γ-polynomials. Here we fix $X \subset \boldsymbol{R}$, $T \subset \boldsymbol{R}$, and $\gamma \in C(T \times X)$. A γ-polynomial is then a function of the form $\sum_{i=1}^n a_i \gamma(t_i, x)$ with $t_i \in T$. Important examples are obtained by taking γ to be e^{tx}, $\cosh tx$, x^t, $\arctan tx$, $(1 + tx)^{-1}$, or $(x - t)_+^n$. Even for closed T and fixed n, the set of γ-polynomials need not be closed. For example, $(\partial/\partial t)\gamma(t, x)$ can occur as a limit function. Hence the existence of best approximations cannot be proved without additional hypotheses. As in exponential approximation, one must consider the enlarged class, containing functions

$$\sum_{i=1}^{l} \sum_{j=0}^{m_i} a_{ij} \frac{\partial^j}{\partial t^j} \gamma(t_i, x)$$

for which $\sum_{i=1}^l (1 + m_i) \leqq n$. Existence of best approximations from the enlarged class can be proved if T is compact. See Braess (J. Approximation Theory **9** (1973), **11** (1974)). The γ-polynomials were first studied by Hobby and Rice (Arch. Rational Mech. Anal. (1967)). Here the theory exhibits a bifurcation which is typical: The unenlarged family affords elegant characterization theorems for best approximation, but is not large enough to assure the existence of best approximations for *all* continuous functions. The enlarged family contains best approximations for every continuous function, but they are not unique, nor so readily characterized.

Another wide class of approximating functions is the varisolvent family. A set $\mathscr{F}$ in $C[a, b]$ is termed "varisolvent" if each $f \in \mathscr{F}$ has a "degree" $n = n(f)$ with two properties: (1) If $g \in \mathscr{F}$ and $g \neq f$ then $f - g$ has at most $n - 1$ roots. (2) If $a \leqq x_1 < \cdots < x_n \leqq b$ and $\varepsilon > 0$ then there is a $\delta > 0$ such that the inequality $\max_i |f(x_i) - \lambda_i| < \delta$ implies the existence of g in $\mathscr{F}$ satisfying $g(x_i) = \lambda_i$ and $\|g - f\| < \varepsilon$.

The two properties just described are abstractions of properties of the polynomial class π_{n-1}, but in a varisolvent class, $n(f)$ may vary. For a varisolvent family the following result holds: Let $f \in C[a, b]$, $g \in \mathscr{F}$, $f - g$ nonconstant. In order that g be a best approximation to f it is necessary and sufficient that $f - g$ possess an alternant of length $n(g) + 1$. That means points $x_0 < \cdots < x_n$ for which $f(x_i) - g(x_i) = (-1)^i \lambda$, $|\lambda| = \|f - g\|$.

Recent work has been directed to removing the hypothesis in this theorem that $f - g$ not be constant. Recent references are Baraar and Loeb (J. Approximation Theory **1** (1968)), and Ling and Tornga (J. Approximation Theory **11** (1974)).

New directions in nonlinear approximation theory have been initiated in the past

three years by Wulbert and Braess, who in general have sought to bring global analysis to bear upon approximation problems. I refer the reader to Wulbert (Amer. J. Math. **93** (1971), 2 papers) and Braess (Math. Z. **132** (1973)). One of Wulbert's theorems describes those nonlinear manifolds on which best approximations can be characterized by the linear theory on the tangent space. As a sample of the work in this promising new direction, I shall describe his result. First of all, manifolds arise naturally in this subject since an approximating class is often defined by a map F: $\boldsymbol{R}^n \to C[a, b]$. The points of $\boldsymbol{R}^n$ are the parameter vectors in the approximation. Given $g \in C[a, b]$ one may ask for a choice of parameters $v \in \boldsymbol{R}^n$ which render $\|g - F(v)\|$ a minimum. The classical polynomial theory, for example, is recovered by taking $F(v)$ to be the function $\sum_{i=1}^{n} v_i x^{i-1}$. If M is a C^1-manifold in $C[a, b]$ there are two maps of importance in approximation. Let Ag denote the set of best approximations to g on M. Thus Ag consists of those $m \in M$ such that $\|g - m\| = \operatorname{dist}(g, M)$. Similarly, let Bg be the set of $m \in M$ for which $\|g - m\| = \operatorname{dist}(g, T(m))$, where $T(m)$ denotes the tangent manifold at m. It is always true that $Ag \subset Bg$, and the case of equality, $A = B$, is of special significance since in this case the linear characterization of best approximations on $T(m)$ also characterizes best approximations on M. Wulbert proved that in order for M to be a Chebyshev set satisfying $A = B$, it is necessary and sufficient that M be Haar-embedded, boundedly connected and boundedly compact. The terminology is as follows. A Chebyshev set is one from which best approximations exist and are unique. Haar-embeddedness is that each tangent space $T(m)$ be a Haar subspace, i.e., the number of roots of a nonzero element must be less than the dimension of $T(m)$. Finally, a set is boundedly connected (or compact) if and only if its intersection with each closed sphere is connected (or compact).

My final remarks concern an old and bizarre example in nonlinear approximation, namely the optimization of the Lagrange interpolation process. Some progress has recently been made by myself and a student, T. A. Kilgore. The problem is to select nodes $t_1 \cdots t_n$ in the order $-1 \leqq t_1 < t_2 < \cdots < t_n \leqq 1$ so as to minimize the norm of the corresponding Lagrange interpolation operator. This operator has the form

$$Lf = \sum_{i=1}^{n} f(t_i) l_i \quad \text{with } l_i(x) = \prod_{j=1; j \neq i}^{n} \frac{(x - t_j)}{(t_i - t_j)}.$$

As is well known, the operator norm of L is equal to the sup-norm of its corresponding Lebesgue function, $\Lambda(x) = \sum_{i=1}^{n} |l_i(x)|$. It was conjectured by Bernstein and others that there exists a choice of nodes for which $\Lambda(x)$ exhibits $n + 1$ equal maxima, and that such a choice of nodes solves the problem. We succeeded in establishing the existence part of this conjecture. Let λ_i denote the maximum of Λ on the interval $[t_i, t_{i+1}]$ $(0 \leqq i \leqq n,\ t_0 = -1,\ t_{n+1} = 1)$. A basic lemma asserts that if the nodes $t_1 \cdots t_{i-1} t_{i+1} \cdots t_n$ are held fixed then t_i can be varied to make $\lambda_{i-1} = \lambda_i$. The correct position of t_i is unique and depends continuously on $t_1 \cdots t_{i-1}\ t_{i+1} \cdots t_n$. This enables one to define maps M_i such that $M_i(t_1 \cdots t_n) = (t_1 \cdots t_{i-1} s t_{i+1} \cdots t_n)$ with s chosen to make $\lambda_{i-1} = \lambda_i$. A fixed point of the composite map $M =$

$M_1 \circ M_2 \circ \cdots \circ M_n$ would be a point for which $\lambda_0 = \lambda_1 = \cdots = \lambda_n$. The remainder of the proof is devoted to showing that M has indeed a fixed point on the (not closed) simplex which is its domain. This problem can be interpreted as one of minimizing $\|\Lambda(t_1, \cdots, t_n; x)\|$, and is thus a problem of best approximation with highly non-linear dependence on the parameters t_i. This work will appear in Aequationes Mathematicae.

UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712, U.S.A.

Труды Международного Конгресса Математиков
Ванкувер, 1974

Теория Устойчивости Разностных Схем и Итерационные Методы*

А. А. Самарский

Доклад посвящен устойчивости разностных схем с несамосопряженными операторами и сходимости итерационных процессов для операторного уравнения первого рода $Au = f$. Отправным пунктом является предложенная автором концепция устойчивости, изложение и развитие которой дано в книгах [1] и [2]. Нам понадобятся основные исходные посылки теории.

1. Двухслойная разностная схема. Разностная схема трактуется как операторно-разностное уравнение в линейном номированном пространстве $H = H_h$, зависящем от параметра h—вектора с нормой $|h| > 0$. Двухслойная схема записывается в канонической форме

$$B(y_{n+1} - y_n)/\tau + Ay_n = \varphi_n, \qquad n = 0, 1, \cdots, \text{задан } y_0 \in H \tag{1}$$

где A и B—линейные операторы, заданные в H, $y_n = y(t_n) \in H$—искомая, $\varphi_n = \varphi(t_n) \in H$—заданная абстрактные функции дискретного аргумента $t_n = n\tau$, A, B, y_n, φ_n зависят от параметров τ и, h, A и B могут зависеть от t_n. Существует оператор B^{-1}. Пространство H может быть как действительным, так и комплексным. Исходное семейство схем (1) задано, если заданы операторы A и B. В этом семействе ищутся классы устойчивых схем. Если H—конечномерное гильбертово пространство со скалярным произведением (y, v) и нормой $\|y\| = \sqrt{(y, y)}$, то необходимые и достаточные условия устойчивости схемы (1) имеют вид операторных неравенств между A и B (см. [1], [2]).

2. Устойчивость по начальным данным. Напомним определение устойчивости двухслойной схемы

*Not presented in person.

(2) $B(y_{n+1} - y_n)/\tau + Ay_n = 0, \qquad n = 0, 1, 2, \cdots$, задан $y_0 \in H$.

Для простоты считаем, что A и B не зависят от n. Пусть $D: H \to H$ линейный оператор, $D^* = D > 0$, H_D—энергетическое пространство, состоящее из элементов y, $v \in H$ со скалярным произведением $(y, v)_D = (Dy, v)$ и нормой $\|y\|_D = \sqrt{(Dy, y)}$. Схема (1) устойчива по начальным данным в H_D, если для решения задачи (2) при любых $y_0 \in H$ имеем:

(3) $$(Dy_{n+1}, y_{n+1}) \leqq (Dy_n, y_n), \qquad n = 0, 1, 2. \cdots.$$

Схема (1) ρ-устойчива, если для (2) имеет место оценка

(4) $$(Dy_{n+1}, y_{n+1}) \leqq \rho(Dy_n, y_n), \qquad n = 0, 1, 2, \cdots, \rho = e^{c_0\tau},$$

где $c_0 = \text{Const}$ не зависит от τ и h.

3. Необходимые и достаточные условия устойчивости. Укажем некоторые необходимые и достаточные условия устойчивости схемы (2) по начальным данным в H_D, предполагая, что либо один из операторов A и B, либо оба оператора являются несамосопряженными:

(5) $$B_0 = \operatorname{Re} B \geqq \tfrac{1}{2}\tau A, \qquad \text{если } A^* = A > 0,\ D = A,$$

(6) $$\operatorname{Re} A^{-1} \geqq \tfrac{1}{2}\tau B^{-1}, \qquad \text{если } B^* = B > 0,\ D = B,$$

(7) $$D(\operatorname{Re} A^{-1})D + (\sigma_0 - \tfrac{1}{2})\tau D \geqq 0, \qquad \text{если } B = D + \sigma\tau A,\ D = D^* > 0,$$

где $\sigma = \sigma_0 + i\sigma_1$—число, $\sigma_0 = \operatorname{Re}\sigma$, $B_0 = \operatorname{Re} B = \frac{1}{2}(B + B^*)$. Эти условия достаточны для устойчивости по правой части схемы (1) (при соответствующем выборе нормы для φ_n). Укажем еще один результат.

Явная схема $(y_{n+1} - y_n)/\tau + Ay_n = 0$, где A—кососимметрический оператор, $A^* = -A$, и A^{-1} существует—неустойчива при любом $D = D^* > 0$. В [1], [2] имеются много примеров, показывающих, что условие (5) удобно для проверки на практике. Условия (6), (7) проверить труднее, если $A^* \neq A$. Теоремы (6), (7) с успехом применяются к несамосопряженным системам уравнений. В этом случае A и B матрицы-операторы; их порядок равен порядку системы (см. [3]).

4. Трехслойные схемы с несамосопряженными операторами. Изложим некоторые результаты для трехслойной схемы

(8) $By_{\dot{t}} + \tau^2 R y_{\bar{t}t} + Ay = 0, \qquad t = n\tau, n = 1, 2, \cdots$, заданы $y^0, y^1 \in H$,

где A, B, R—линейные операторы, заданные в H,

$$y = y^n = y(t_n), \quad y_{\dot{t}} = (y^{n+1} - y^{n-1})/2\tau,$$
$$y_{\bar{t}t} = (y^{n+1} - 2y^n + y^{n-1})/\tau^2, \quad y^0 = y(0).$$

Решением задачи (8) в момент $t_n = n\tau$ называется вектор $y_n = \{y^n, y^{n+1}\} \in H^2$, $H^2 = H \oplus H$. Устойчивость по начальным данным схемы (8) означает, что $(Dy_{n+1}, y_{n+1}) \leqq (Dy_n, y_n)$ при любых $y_0 = \{y^0, y^1\} \in H^2$, где $D = D^* > 0$ — линейный оператор, заданный в H^2.

Приведем в виде таблицы некоторые теоремы.

Исходное семейство	Необходимое и достаточное условие устойчивости
(9) $B_0 > 0$, $A^* = A > 0$, $R^* = R$	$R \geqq \frac{1}{4}A$
(10) $B = E$, $A^* = -A$, $R^* = -R$, $AR = RA$ $(Ex = x)$	$E + \tau^2 A(A - 4R) \geqq 0$

Представляют интерес следующие частные случаи

Схема	$A^* = A$	$A^* = -A$
$y_t + Ay = 0$	Абсолютно неустойчива	Устойчива при $\tau \|A\| \leqq 1$
$y_{\bar{t}t} + Ay = 0$	Устойчива при $A > 0$, $(\tau^2/4)\|A\| \leqq 1$	Абсолютно неустойчива

Пример 1. Схема для уравнения Шредингера $\sqrt{-1}\, y_t + \sigma A y^{n+1} + (1 - 2\sigma)Ay^n + \sigma A y^{n-1} = 0$, $Ay = -y_{\bar{x}x} = -(y_{i-1} - 2y_i + y_{i+1})/h^2$, $y_i = y(x_i)$, $x = ih$, $i = 0, 1, 2, \cdots, N$, $y_0 = y_N = 0$, σ—действительное число—устойчива при $\sigma \geqq \frac{1}{4}(1 - 1/\tau^2\|A\|^2)$ или $\sigma \geqq \frac{1}{4}(1 - h^4/16\tau^2)$. В частности, явная схема устойчива при $\tau \leqq \frac{1}{4}h^2$.

Пример 2. Схема $y_t + \tau^2 y_{t\bar{t}\bar{x}}/h = y_{\bar{x}x}$, $x = ih$, $i = 0, 1, 2, \cdots, N$, $y_0 = y_N = 0$ устойчива при $\tau \leqq 0{,}5h^2$.

5. Асимптотическая устойчивость. Общее определение устойчивости не всегда обеспечивает нужные качественные свойства схемы. Так, схемы для параболических и гиперболических уравнений должны обладать разными свойствами. Для параболических уравнений характерно затухание при $t \to \infty$ (выход на регулярный режим) влияния начальных данных. Естественно требовать, чтобы и схема обладала этим свойством. Будем говорить, что схема (2) асимптотически устойчива в H_D, если существует такая постоянная $\delta > 0$, не зависящая от τ и h, что $\|y_n\| \leqq e^{-\delta t_n}\|y_0\|$ при любых $t_n = n\tau$. Если схема (2) ρ-устойчива с постоянной $c_0 = -\delta < 0$, то она асимптотически устойчива. Как показывает пример схемы с весами $(y_{n+1} - y_n)/\tau + A(\sigma y_{n+1} + (1 - \sigma)y_n) = 0$, где $A = A^* > 0$, требование асимптотической устойчивости может приводить к дополнительным ограничениям на шаг τ. Достаточное условие асимптотической устойчивости имеет вид $\tau\delta/(1 + \sigma\tau\delta) + \tau\Delta/(1 + \sigma\tau\Delta) \leqq 2$ где δ и Δ наименьшее и наибольшее собственные значения оператора A. Отсюда следует, что симметричная схема ($\sigma = 0{,}5$), абсолютно устойчивая в обычном смысле (3), асимптотически устойчива при $\tau \leqq \tau_0$, $\tau_0 = 2/\sqrt{\delta\Delta}$. В случае уравнения теплопроводности $\partial u/\partial t = \partial^2 u/\partial x^2$ на отрезке $0 \leqq x \leqq 1$ и $u(0, t) = u(1, t) = 0$ имеем $Ay = -y_{\bar{x}x}$ и $\tau_0 \approx h/\pi$.

Заметим, что схема (2) с $B = (E + \sigma\tau A_0/2)^2$, $A = A_0 + \sigma^2\tau A_0^2/4$, где $A_0^* = A_0 > 0$, $\sigma = 2 - \sqrt{2}$, асимптотически устойчива при любых τ.

Параболические разностные схемы изучались в [5], [6]. В [6] показано, что условие параболичности схемы (2) имеет вид

$$B \geqq \frac{1 + \varepsilon}{2}\tau A, \qquad \varepsilon = \text{Const} > 0.$$

6. Итерационные схемы. Методы общей теории устойчивости разностных схем позволяют построить единую теорию итерационных методов решения линейного операторного уравнения $Au = f$, где $A: H \to H$, $f \in H$, H— гильбертово пространство, A—вообще говоря, несамосопряженный оператор. Итерационная схема, по аналогии с п.1, записывается в каноническом виде

$$(11) \qquad B\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \qquad k = 0, 1, 2, \cdots \qquad \forall y_0 \in H,$$

где y_k—итерация номера k, τ_k—параметр, $B: H \to H$—линейный оператор; он имеет обратный B^{-1} и может зависеть от k. Здесь, для простоты, считаем, что B не зависит от k. Пусть в H задан линейный оператор $D = D^* > 0$. Задача теории состоит в получении оценок $\|y_n - u\|_D \leqq q_n \|y_0 - u\|_D$. Схема (11) сходится, если $q_n \to 0$ при $n \to \infty$, так что $q_n < \varepsilon$ при $n \geqq n_0(\varepsilon)$, $\varepsilon > 0$. Надо найти $\min q_n$ ($\min n_0(\varepsilon)$) путем выбора $\{\tau_k\}$ и B. Зададим исходное семейство схем (11) условиями

$$(12) \qquad (DB^{-1}A)^* = DB^{-1}A, \qquad \gamma_1 D \leqq DB^{-1}A \leqq \gamma_2 D, \qquad \gamma_2 \geqq \gamma_1 > 0.$$

Минимум q_n достигается, если $\{\tau_k\}$ есть чебышевский набор параметров: $\tau_k = \tau_0/(1 + \rho_0\mu_k)$, $\tau_0 = 2/(\gamma_1+\gamma_2)$, $\rho_0 = (1 - \xi)/(1 + \xi)$, $\xi = \gamma_1/\gamma_2$, $\mu_k \in \mathfrak{M}_n\{\mathrm{Cos}((2i - 1)/2n)\pi, i = 1, 2, \cdots, n\}$, $k = 1, 2, \cdots, n$, где $\mathfrak{M}_n$—упорядоченная последовательность, при которой схема (11) вычислительно устойчива. Устойчивые наборы $\{\tau_k\}$ даны в [7], [1], [8], [9]. При этом $n \geqq n_0(\varepsilon) = \ln(2/\varepsilon)/2\sqrt{\xi}$, $\xi = \gamma_1/\gamma_2$. Если $n = 1$, то $\tau_k = \mathrm{Const} = \tau_0$ и мы получаем неявную схему простой итерации; в этом случае $n_0(\varepsilon) = \ln(1/\varepsilon)/2\xi$. Вычислительная устойчивость схемы (11), т.е. устойчивость относительно возмущения всех входных данных: f, A, B и постоянных γ_1 и γ_2 исследована в [10].

7. Попеременно-треугольный метод (ПТМ). Важным является вопрос о выборе B. Предложенный автором [1] ПТМ определяется заданием

$$B = (E + \omega R_1)(E + \omega R_2), \qquad R_2^* = R_1, \qquad R_1 + R_2 = R = R^* > 0,$$

что соответствует представлению симметричной матрицы R в виде суммы нижней и верхней треугольных матриц, причем $c_1 R \leqq A \leqq c_2 R$, $c_2 \geqq c_1 > 0$. Параметр ω выбирается из условия минимума числа итераций. В случае модельной задачи Дирихле для p-мерного уравнения Лапласа ($p \geqq 3$) в параллелепипеде ($0 \leqq x_\alpha \leqq 1$, $\alpha = 1, 2, \cdots, p$) на сетке ω_h с шагами $h_1 = h_2 = \cdots = h_p = h$ ПТМ с устойчивым чебышевским набором $\{\tau_k\}$ требует $n_0(\varepsilon) \approx 0,28 \ln(2/\varepsilon)/\sqrt{h}$ итераций. Хотя для метода переменных направлений (МПН) с циклическим набором параметров $\{\tau_k\}$ имеем $n_0(\varepsilon) = O(\ln(1/h) \ln(1/\varepsilon))$, однако, на реальных сетках при $p \geqq 3$ для ПТМ с чебышевским набором $\{\tau_k\}$ надо в 3–4 раза меньше итераций, чем для МПН [11]. Кроме того, МПН применим только в случае прямоугольных областей, тогда как для ПТМ таких ограничений нет. ПТМ применим для разностных схем, соответствующих эллиптическим уравнениям и системам общего вида. В этом случае— R—разностный оператор Лапласа (в случае уравнений второго порядка).

Литература

1. А. А. Самарский, *Введение в теорию разностных схем*, "Наука", Москва, 1971.

2. А. А. Самарский и А. В. Гулин, *Устойчивость разностных схем*, "Наука", Москва, 1973.

3. А. В. Гулин и А. А. Самарский, *Об устойчивости разностных схем с несамосопряженными операторами*, Доклады АН СССР **206** (1972), 1280–1283.

4. А. В. Гулин, *Об устойчивости трехслойных разностных схем с несамосопряженными операторами*, Доклады АН СССР **210** (1973), 513–516.

5. В. Б. Андреев, *Об устойчивости по начальным данным разностных схем для параболических уравнений*, Журн. вычисл. мат. и мат. физ. **11** (1971).

6. Ю. И. Мокин, *Двухслойные параболические и слабо параболические разностные схемы*, Журн. вычисл. мат. и мат. физ. **14** (1974).

7. В. И. Лебедев и С. А. Финогенов, *О порядке выбора итерационных параметров в чебышевском циклическом итерационном методе*, Журн. вычисл. мат. и мат. физ. **11** (1971), 425–438.

8. Е. С. Николаев и А. А. Самарский, *Выбор итерационных параметров в методе Ричардсона*, Журн. вычисл. мат. и мат. физ. **12** (1972), 960–973.

9. В. И. Лебедев и С. А. Финогенов, *Решение проблемы упорядочивания параметров в чебышевских итерационных методах*, Журн. вычисл. мат. и мат. физ. **13** (1973), 18–33.

10. Е. С. Николаев и А. А. Самарский, *О вычислительной устойчивости двухслойных и трехслойных итерационных схем*, Журн. вычисл. мат. и мат. физ. **12** (1972), 1197–1207.

11. _____, *Методы численного решения задачи Дирихле для уравнения Пуассона любого числа измерений*, Доклады АН СССР **206** (1972), 815–818.

Институт Прикладной Математики АН СССР
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Recent Progress in the Numerical Treatment of Ordinary Differential Equations

Hans J. Stetter

I. Two-point boundary value problems. We consider the nonlinear boundary value problem

$$y'(t) = f(t, y(t)), \tag{D}$$

$$r(y(0), y(T)) = 0, \tag{R}$$

where $y: [0, T] \to \boldsymbol{R}^s$, $f: \boldsymbol{R} \times \boldsymbol{R}^s \to \boldsymbol{R}^s$, $r: \boldsymbol{R}^s \times \boldsymbol{R}^s \to \boldsymbol{R}^s$.

As the difficulties encountered with a reduction to initial value problems ("shooting") increase with the length T of the basic interval, a fundamental approach consists in breaking this interval into *segments* $[T_{m-1}, T_m]$, $m = 1(1)M$. Let $\phi(\eta_{m-1}, \eta_m)$ denote the numerical approximation for (D) on $[T_{m-1}, T_m]$; then (D)/(R) becomes

$$(\Delta) \qquad \boldsymbol{\Psi}\boldsymbol{\eta} := \begin{bmatrix} \phi_1(\eta_0, \eta_1) \\ \vdots \\ \phi_M(\eta_{M-1}, \eta_M) \\ r(\eta_0, \eta_M) \end{bmatrix} = 0.$$

The system (Δ) has to satisfy a *stability condition*

$$\|\boldsymbol{\eta}^{(1)} - \boldsymbol{\eta}^{(2)}\| \leqq S\|\boldsymbol{\Psi}\boldsymbol{\eta}^{(1)} - \boldsymbol{\Psi}\boldsymbol{\eta}^{(2)}\|,$$

so that the smallness of the local segmentation errors $\phi(y(t_{m-1}), y(t_m))$ implies the smallness of $\eta_m - y(t_m)$.

There are two basic strategies to make the local segmentation error small:

(i) to use large segments but a very accurate approximation to (D),

(ii) to use a crude approximation to (D) but rather small segments.

Strategy (i) has been implemented in Bulirsch's version of *parallel shooting* [1] where (D) is replaced on each segment by a GBS extrapolation algorithm. The lengths of the segments are chosen sufficiently small to inhibit an undue growth of the trial solutions. Strategy (ii) is found, e.g., in Keller's "box-scheme" *difference method* [2] based on the trapezoidal rule. The local error is kept small by the use of sufficiently short segments.

In each case, there remains the difficult problem of obtaining an accurate numerical solution to the nonlinear system (Δ) with its typical bidiagonal block structure. Very refined techniques, like damped Newton iterations with Broyden's rank one updating of the Jacobian, have to be used (see, e.g., [3]).

The asymptotic expansion results known for the difference method have been used by Pereyra [4] to design an *iterated difference correction technique*. By constructing increasingly better difference approximations to the local discretization error from previously gained solution values, one obtains more and more accurate solution values on the same segmentation, from essentially the same system of equations.

The *invariant imbedding* approach (see, e.g., [5]), though mathematically quite interesting, suffers from the same practical difficulties as simple shooting of which it is a systematized version. Assume separated boundary conditions

$$\text{(R')} \qquad \begin{aligned} r_1(y(0)) &= 0, \qquad r \text{ conditions}, \\ r_2(y(T)) &= 0, \qquad s - r \text{ conditions}, \end{aligned}$$

with (D). The components of $y(t)$ are collected into vectors $u(t) \in \mathbb{R}^r$ and $x(t) \in \mathbb{R}^{s-r}$ such that (D) splits into

$$\text{(C)} \qquad \begin{aligned} x'(t) &= F_1(t, x(t), u(t)), \\ u'(t) &= F_2(t, x(t), u(t)), \end{aligned} \qquad u(0) = (R_1(x(0)),$$

which is equivalent to the hyperbolic initial value problem

$$\text{(H)} \qquad u_t + u_x F_1(t, x, u) = F_2(t, x, u), \qquad u(0, x) = R_1(x).$$

The solution trajectory $(x(t), u(t))$ of (C) must be imbedded in the solution manifold $u(t,x)$ of (H); it can be distinguished by the boundary condition at the far end $r_2(x, u(T, x)) = 0$.

II. Stiff systems. A bibliography of the problem and of approaches to its solution up to 1970 is found in [6]; recently more than fifty percent of all papers on the numerical treatment of o.d.e.'s have been devoted to stiff systems.

The consideration of *uniform asymptotic stability* (for $T \to \infty$) was introduced by Dahlquist [7] and further developed by Stetter [8]; it points out the nonvalidity of "linear equations with constant coefficients" results for general systems and for discretizations with variable steps.

The necessity of distinguishing between slowly and quickly decaying solutions led to the *model problem*

$$\text{(S)} \qquad \begin{aligned} y_1' &= g_1(t, y_1, y_2)/\varepsilon, \qquad y_1 \in R_1 = \boldsymbol{R}^{s_1}, \\ y_2' &= g_2(t, y_1, y_2), \qquad y_2 \in R_2 = \boldsymbol{R}^{s_2}. \end{aligned}$$

In this singular perturbation problem, one may study the limit process $\varepsilon \to 0$ for fixed $h > 0$ and request that certain quantities of the discretization remain bounded uniformly for $\varepsilon \in (0, \varepsilon_0)$.

These ideas had already been present in earlier investigations (e.g., [9], [10]); very recently they have been elaborated in a thesis by van Veldhuizen [11]. He assumes that (S) may be transformed into

$$\text{(S')} \qquad x' = A(t, \varepsilon)x + Q(t, x, \varepsilon) + b(t, \varepsilon)$$

with

$$A(t, \varepsilon) = \begin{bmatrix} \varepsilon^{-1}A_1(t) & 0 \\ 0 & A_2(t) \end{bmatrix}$$

while Q and b remain harmless for $\varepsilon \to 0$. Then one may define *"regular" or "smooth" solutions* $\tilde{x}$ which decay slower than some exponential $e^{\varrho t}$ uniformly in ε while *"singular" or "fast" solutions* $\hat{x}$ do not possess this property. An initial condition $x(0)$ generates a unique decomposition into a regular and a singular component.

For one-step recursions of the type $x_i = P_{i-1}x_{i-1} + T_{i-1}(x_{i-1}) + W_i(x_i) + b_i$ with P_i a rational function of

$$\begin{bmatrix} \varepsilon^{-1}P_{1i} & 0 \\ 0 & P_{2i} \end{bmatrix}$$

and well-behaved T_i, W_i, b_i, an analogous analysis is made by van Veldhuizen. He then studies one-step discretizations of (S) resp. (S'). Consistency, convergence, and stability are defined w.r.t. smooth solutions only but required to hold uniformly for $\varepsilon \in (0, \varepsilon_0]$ and $h \in (0, h_0(\varepsilon)]$. If not $h_0(\varepsilon) = O(\varepsilon)$ as $\varepsilon \to 0$ the uniformity region contains arbitrarily large values of h/ε.

In this fashion, one can even obtain some results about asymptotic expansions for smooth solutions; see also [9]. The main goal of the theory is a characterization of discretizations where the step size is only restricted by accuracy requirements. In this connection it is important that the *decomposition* into smooth and fast solutions is accurately modelled by the discretization. In an example, the implicit Euler and the implicit midpoint methods perform well in this respect while the implicit trapezoidal rule generates a decomposition error of $O(h^2/\varepsilon)$! The decomposition problem has also been taken into account in Lindberg's IMPEX 2-package for stiff equations [12].

One of the most powerful construction principles for discretization methods suitable for stiff systems is *exponential fitting*. There are two systematic approaches which differ when applied to systems:

(i) reproduce exactly solutions of type $p_r(t)e^{qt}$,

(ii) solve exactly right-hand sides of type $qy + p_r(t)$;

p_r is a polynomial of maximal degree $r \geqq 0$.

Approach (i) may be generalized to solutions of type $\sum_i p_{r_i}(t)^{q_i t}$, with $q_0 = 0$. Mäkelä [13] has constructed linear multistep methods of this kind by considering Hermite interpolation with base functions $t^j e^{q_i t}$.

In approach (ii), one assumes the knowledge of the solution $y(t) = G(t, t_0)y(t_0)$ of the system $y' = Q(t)y$ and imitates the classical construction principles starting from the representation

$$y(t + h) = G(t + h, t)\, y(t) + \int_t^{t+h} G(t + h, \tau)\phi(\tau, y(\tau))\, d\tau$$

for a solution of $y' = Q(t)y + \phi(t, y)$. In this fashion, Meister [14] has constructed analogs to R-K methods while Liniger and Sarkany [15] have constructed Adams-type multistep methods. A practical difficulty of approach (ii) is the occurrence of matrix coefficients.

III. Error estimation and step size control. Even for nonstiff systems it is extremely difficult and expensive to obtain meaningful *bounds on the global discretization error*; see, e.g., Jackson [16]. Thus one may resort to *asymptotic estimates* in spite of their dependence on h being "sufficiently small". Here, a suggestion due to Zadunaisky [17] appears to be quite efficient: A "comparison problem" with known solution is constructed and solved by the same method, its error is used as an estimate.

The suggestion is implemented in this fashion:

$$y' = f(t, y), \quad y(0) = y_0 \xrightarrow[\text{step } h]{\text{method M}} \eta(t_i)$$

$$\xrightarrow[\text{interpolation}]{\substack{\text{piecewise}\\ \text{polynomial}}} \bar{y}(t) \longrightarrow d(t) := \bar{y}'(t) - f(t, \bar{y}(t))$$

$$\longrightarrow y' = f(t, y) + d(t),\ y(0) = y_0 \text{ ("comparison problem")}$$

$$\xrightarrow[\text{step } h]{\text{method M}} \bar{\eta}(t_i).$$

Then $\eta(t_i) - y(t_i) \approx \bar{\eta}(t_i) - \bar{y}(t_i) = \bar{\eta}(t_i) - \eta(t_i)$. With sufficiently accurate interpolation, asymptotic expansion theory shows [18] that for a method of order p

$$\eta(t_i) - y(t_i) = \bar{\eta}(t_i) - \eta(t_i) + O(h^{2p}).$$

Further analysis and experimental studies are in progress at the author's institute.

If the asymptotic estimate is turned into a correction one obtains a close analog to Fox-Pereyra's difference correction procedure which could be called *"differential correction"*. It can also be applied iteratively: see [18].

In the *step size control problem* one wishes to determine a step sequence such that the global discretization error is kept within a prescribed tolerance over the integration interval with a minimal number of steps. A recent study by Lindberg [19] has added much insight into the structure of the problem.

Lindberg assumes that the global error e is the solution of

$$e'(t) - f_y(t, y(t))e(t) = c[h(t)]^p\, y^{(p+1)}(t), \qquad e(0) = 0.$$

The continuous function $h(t) > 0$ is to be determined such that $\int_0^T (dt/h(t))$ becomes

a minimum while $\|e(t)\| \leqq \varepsilon$ for $t \in [0, T]$. (For a reasonable analytic treatment one has to assume certain sign conditions on the components of e which exclude a cancellation of error terms.)

The case $y' = Ay$, $y(t) \in \boldsymbol{R}^s$, with Euclidean norm, gives the following explicit result: Let $\alpha(t) :=$ convex hull of $\{(0, 0), (t, \phi(t)), t \in [0, T]\}$ where $\phi(t) := \varepsilon/(c\|y^{(p+1)}(t)\|)$. Then $h_{\text{opt}}(t) := [\alpha'(t)]^{1/p}$. Since α is convex, the optimal step sequence cannot decrease, a surprising result which holds even in more general cases!

For stiff systems of type $y' = Ay$, the optimal step sequence is obtained (except in the initial phase) with the control $\|l(t)\| = |q(t)|\,\varepsilon$ where l is the local discretization error (per unit step) and q the "*locally dominant eigenvalue*"

$$q(t) := \frac{d}{dt}\|y^{(p+1)}(t)\| \Big/ \|y^{(p+1)}(t)\|.$$

Since the quantity may change violently in stiff systems the control $\|l(t)\| \approx \varepsilon \cdot \text{const}$ is inefficient.

In all cases covered by Lindberg's theory the optimal step sequence corresponds to a certain control on the local error *per unit step* which is thus once more established as the relevant quantity.

IV. Software for ordinary differential equations. The average user should:

Specify his problem in a nonsophisticated and standarized form.

Not have to set parameter values which have no meaning w.r.t. the problem (e.g., step size).

Be able to specify qualitative information about his problem.

Receive reliable results of a specified accuracy within a reasonable computing time.

The creation and implementation of program packages meeting these demands require large-scale comparative testing and international collaboration and standardization. Major steps towards these goals have been made by the Toronto testing project (see, e.g., [**20**]) and by the establishment of Working Group 2.5 on Numerical Mathematical Software within IFIP.

Bibliography

1. R. Bulirsch, *Die Mehrzielmethode zur numerischen Lösung von nichtlinearen Randwertproblemen und Aufgaben der optimalen Steuerung*, Report, Carl-Cranz-Society, 1971. (A paper by Bulirsch et al., *Numerical solution of nonlinear two-point boundary value problems*, will appear in Numer. Math.).

2. H. B. Keller, *Accurate difference methods for nonlinear two-point boundary value problems*, SIAM J. Numer. Anal. **11** (1974).

3. J. Stoer and R. Bulirsch, *Einführung in die numerische Mathematik*. II, Heidelberger Taschenbuch 114, Springer-Verlag, Berlin, 1973, Chap. 7.3.

4. V. Pereyra, *A variable order, variable step, finite difference method for nonlinear multipoint boundary value problems*, Report 73–06, Comp. Dept., Univ. Centr. of Venezuela, 1973.

5. G. H. Meyer, *Initial value methods for boundary value problems*, Academic Press, New York, 1973.

6. G. Bjurel, et al., *Survey of stiff differential equations*, Report NA 70.11, Roy. Inst. Tech. Stockholm, 1970.

7. G. Dahlquist, *A special stability problem for linear multistep methods*, BIT **3** (1963), 27–43. MR **30** #715.

8. H. J. Stetter, *Analysis of discretization methods for ordinary differential equations*, Springer-Verlag, Berlin, 1973.

9. G. Dahlquist, *A numerical method for some ordinary differential equations with large Lipschitz constants*, Information Processing (Proc. IFIP Congress, Edinburgh, 1968), vol. 1: Mathematics, Software, North-Holland, Amsterdam, 1969, pp. 183–186. MR **41** #2937.

10. ———, *On some implicit onestep methods for stiff systems of differential equations*, Report NA 73.02, Roy. Inst. Tech. Stockholm, 1973.

11. M. van Veldhuizen, *Convergence of one-step discretizations for stiff differential equations*, Thesis, Univ. of Utrecht, 1973.

12. B. Lindberg, IMPEX 2, *a procedure for solution of systems of stiff differential equations*, Report NA 73.03, Roy. Inst. Tech. Stockholm, 1973.

13. M. Mäkelä, et al., *Exponentially fitted multistep methods constructed by generalized Hermite-Birkhoff interpolation*, Report MAT-A 31, Helsinki, Techn. Univ., 1973.

14. G. Meister, *Über die Integration von Differentialgleichungs systemen* 1. *Ordnung mit exponentiell angepassten numerischen Methoden*, Computing **13** (1974), 327–352.

15. E. F. Sarkany and W. Liniger, *Exponential fitting of matricial multi-step methods for ordinary differential equations*, Math. Comp. **28** (1974), 1035–1052.

16. L. W. Jackson, *Automatic error analysis for the solution of ordinary differential equations*, Tech. Rep. 28 Computer Science, Univ. of Toronto, 1971.

17. P. E. Zadunaisky, *On the estimation of errors propagated in the numerical integration of ordinary differential equations*, Dundee Conf., 1973 (communication).

18. H. J. Stetter, *Economical global error estimation*, Sympos. on Stiff Differential Systems, Wildbad, Plenum Press, New York, 1974.

19. B. Lindberg, *Optimal stepsize sequences and requirements for the local error for methods for (stiff) differential, equations*, Techn. Rep. 67 Computer Science, Univ. of Toronto, 1974.

20. W. H. Enright, et al., *Test results on initial value methods for non-stiff ordinary differential equations.*, Techn. Rep. 68 Computer Science, Univ. of Toronto, 1974.

INSTITUT FÜR NUMERISCHE MATHEMATIK,
TECHNISCHE HOCHSCHULE, WEIN
A-1040 WEIN, AUSTRIA

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Finite Element Method—Linear and Nonlinear Applications

Gilbert Strang

I. Numerical analysis is a crazy mixture of pure and applied mathematics. It asks us to do two things at once, and on the surface they do appear complementary: (i) to propose a good algorithm, and (ii) to analyze it. In principle, the analysis should reveal what makes the algorithm good, and suggest how to make it better. For some problems—computing the eigenvalues of a large matrix, for example, which used to be a hopeless mess—this combination of invention and analysis has actually succeeded. But for partial differential equations, which come to us in such terrible variety, there seems to be a long way to go.

We want to speak about an algorithm which, at least in its rapidly developing extensions to nonlinear problems, is still new and flexible enough to be improved by analysis. It is known as the finite element method, and was created to solve the equations of elasticity and plasticity. In this instance, the "numerical analysts" were all engineers. They needed a better technique than finite differences, especially for complicated systems on irregular domains, and they found one. Their method falls into the framework of the Ritz-Galerkin technique, which operates with problems in "variational form"—starting either from an extremum principle, or from the weak form of the differential equation, which is the engineer's equation of virtual work. *The key idea which has made this classical approach a success is to use piecewise polynomials as trial functions in the variational problem.*[1]

We plan to begin by describing the method as it applies to linear problems. Because the basic idea is mathematically sound, convergence can be proved and the error can be estimated. This theory has been developed by a great many numerical

[1]The most important applications are still to structural problems, but no longer to the design of airplanes; that has been superseded by the safety of nuclear reactors.

analysts, and we can summarize only a few of the most essential points—the conditions which guarantee convergence, and which govern its speed. This linear analysis has left everyone happier, and some divergent elements have been thrown out, but the method itself has not been enormously changed. *For nonlinear problems the situation is entirely different.* It seems to me that numerical analysts, especially those in optimization and nonlinear systems, can still make a major contribution. The time is actually a little short, because the large-scale programs for plasticity, buckling, and nonlinear elasticity are already being written. But everyone is agreed that they are tremendously expensive, and that new ideas are needed.

Nonlinear problems present a new challenge also to the analyst who is concerned with error estimates. The main aim of this paper is to describe some very fragmentary results (§III) and several open questions (§IV). We are primarily interested in those nonlinearities which arise, in an otherwise linear problem, when the solution is required to satisfy an *inequality constraint.* This is typical of the problems in plasticity. The solution is still determined by a variational principle, but the class of admissible functions becomes a convex set instead of a subspace. In other words, the equation of virtual work becomes a *variational inequality.*

At the end we look in still a different direction, at *linear programming constrained by differential equations.* Here we need not only good algorithms and a proper numerical analysis, but also answers to the more fundamental questions of existence, uniqueness, and regularity.

II. Linear equations. The finite element method applies above all to elliptic boundary value problems, which we write in the following form: Find u in the space of admissible functions V such that

$$a(u, v) = l(v) \quad \text{for all } v \text{ in } V. \tag{1}$$

STANDARD EXAMPLE. $\iint (u_x v_x + u_y v_y)\, dx\, dy = \iint fv\, dx\, dy$ for all v in $\mathscr{H}_0^1(\Omega)$. This is the weak form of Poisson's equation $-\Delta u = f$. Because the expression $a(u, v)$ is in this case symmetric and positive definite, the problem is equivalent to: Minimize $J(v) = a(v, v) - 2l(v)$ over the admissible space V. The "strain energy" $a(v, v)$ is the natural norm in which to estimate the error.

The error comes from changing to a finite-dimensional problem: Find u_h in S_h such that

$$a_h(u_h, v_h) = l_h(v_h) \quad \text{for all } v_h \text{ in } S_h. \tag{2}$$

It is this problem which the computer actually solves, once it is given a basis $\phi_1, \cdots, \phi_N$ for the space S_h. Very briefly, it has to form the stiffness matrix $K_{ij} = a_h(\phi_i, \phi_j)$ and the load vector $F_j = l_h(\phi_j)$, solve the linear system $KQ = F$, and print out the approximate solution $u_h = \sum Q_j \phi_j$. That sounds straightforward, but it is nearly impossible unless the basis functions ϕ_j are extremely simple, and nearly useless unless they can provide a good approximation to the true solution u. The finite element method manages to combine both properties.[2]

[2]We shall have to refer to the book [1] and to its bibliography, both for the construction of piece-

Our plan in this section is to summarize four of the main points in the theory of convergence. Each of them is concerned with the change in solution when there is a change in the problem—when the admissible space V is replaced by S_h, or the given a and l are approximated by a_h and l_h. To give some kind of order to the discussion, we formulate all four as applications of the "fundamental theorem of numerical analysis":

$$\text{Consistency} + \text{Stability} \Leftrightarrow \text{Convergence}.$$

1. *The classical Ritz-Galerkin case.* The energy $a(v,v)$ is symmetric positive definite; S_h is a subspace of V; $a_h = a$ and $l_h = l$.

Since every v_h is an admissible v, we may compare (1) and (2): $a(u, v_h) = a(u_h, v_h)$. This means that in the "energy inner product," u_h is the projection of u onto the subspace S_h. In other words, the positive definiteness of $a(v, v)$ implies two properties at once [1, p. 40]: The projection u_h is no larger than u itself,

$$a(u_h, u_h) \leqq a(u, u), \tag{3}$$

and at the same time u_h is as close as possible to u:

$$a(u - u_h, u - u_h) \leqq a(u - v_h, u - v_h) \quad \text{for all } v_h \text{ in } S_h\,. \tag{4}$$

Property (3) represents stability; the approximations are uniformly bounded. Given that u can be approximated by the subspace S —in this Ritz-Galerkin context, *consistency is the same as approximability*—convergence follows immediately from (4).

2. *The indefinite case.* u is only a *stationary point* of the functional $J(v)$. This corresponds to the use of Lagrange mutipliers in optimization; the form $a(v,v)$ can take either sign, and v may include two different types of unknowns—both displacements and stresses, in the "mixed method" and "hybrid method."

Consistency reduces as before to approximation by polynomials. But stability is no longer automatic; even the simplest indefinite form $J(v) = v_1 v_2$—which has a unique stationary point at the origin, if V is the plane R^2—will collapse on the one-dimensional subspace given by $v_2 = 0$. Therefore, for each finite element space S_h and each functional $J(v)$, it has to be proved that a degeneracy of this kind does not occur.

The proper stability condition is due to Babuska and Brezzi:

$$\sup_{\|v\|=1} |a(v, w)| \geqq c \parallel w \parallel. \tag{5}$$

Brezzi has succeeded in verifying this condition for several important hybrid elements. For other applications the verification is still incomplete, and the convergence of stationary points—which is critical to the whole theory of optimization—remains much harder to prove than the convergence of minima.

3. *The modified Galerkin method.* a and l are changed to a_h and l_h (numerical integration of the stiffness matrix and load vector), and v_h may lie outside V (non-

wise polynomials and for the proof of their approximation properties. Perhaps the favorites, when derivatives of order m appear in the energy $a(v, v)$, are the polynomials of degree $m+1$.

conforming elements).

The effect on u can be estimated by combining (1) and (2):

$$a_h(u - u_h, u - u_h) = (a_h - a)(u, u - u_h) - (l_h - l)(u - u_h). \tag{6}$$

Stability, in this situation, means a lower bound for the left side:

$$a_h(u - u_h, u - u_h) \geqq c a(u - u_h, u - u_h). \tag{7}$$

Consistency is translated into an upper bound for the right side, and *it is checked by applying the patch test*: Whenever the solution is in a "state of constant strain"—the highest derivatives in $a(u, u)$ are all constant—then u_h must coincide with u.[3]

The patch test applies especially to nonconforming elements, for which $a(v_h, v_h) = \infty$; the derivatives of v_h introduce delta-functions, which are simply ignored in the approximate energy a_h. This is extremely illegal, but still the test is sometimes passed and the approximation is consistent. Convergence was established by the author for one such element, and Raviart, Ciarlet, Crouzeix, and Lesaint have recently made the list much more complete.

4. *Superconvergence.* Extra accuracy of the finite element approximation at certain points of the domain. It was recognized very early that in some special cases—$u'' = f$ with linear elements, or $u'''' = f$ with cubics—the computed u_h is exactly correct at the nodes. (The Green's function lies in S_h.) And even earlier there arose the difficulty of interpreting the finite element output in a more general problem; u_h and its derivatives can be evaluated at any point in the domain, but which points do we choose? This question is as important as ever to the engineers.

In many problems the error $u - u_h$ oscillates within each element, and there must be points of exceptional accuracy. Thomée discovered superconvergence at the nodes of a regular mesh, for $u_t = u_{xx}$, and his analysis has been extended by Douglas, Dupont, Bramble, and Wendroff. It is not usually carried out in our context of consistency and stability, but perhaps it could be: Consistency is checked by a *patch test at the superconvergence points*, to see which polynomial solutions and which derivatives are correctly reproduced, and stability needs to be established in the pointwise sense.[4]

III. Variational inequalities. What happens when a constraint such as $v \leqq \psi$ is enforced on the admissible functions v, so that the functional $J(v)$ is minimized *only over a convex subset* K of the original space V? This occurs naturally in plasticity theory, when v represents the stress; wherever the yield limit ψ is reached, the differential equation (Hooke's law) is replaced by plastic flow. For the minimizing u, the "free boundary" which marks out this plastic region $u = \psi$ is not known in advance.[5] Since such a solution u lies on the edge of the convex set K, $J(u) \leqq J(u + \varepsilon(v - u))$ is guaranteed only for $\varepsilon \geqq 0$. This translates into *the variational*

[3]The patch test is also an ideal way to check that a finite element program is actually working.

[4]Convergence in L_∞ has been proved by Scott since the Congress; it was one of the outstanding problems in the linear theory.

[5]This boundary cannot be found by solving the original linear problem and then replacing u by $\min(u, \psi)$!

inequality which determines u:

$$a(u, v - u) \geqq l(v - u) \quad \text{for all } v \text{ in } K. \tag{8}$$

In the finite element method, we minimize an approximate functional $J_h(v) = a_h(v, v) - 2l_h(v)$ over a finite-dimensional convex set K_h. For example, the piecewise polynomials may be constrained by $v_h \leqq \psi$ at the nodes of the triangulation. Again the minimizing u_h is determined by a variational inequality,

$$a_h(u_h, v_h - u_h) \geqq l_h(v_h - u_h) \quad \text{for all } v_h \text{ in } K_h; \tag{9}$$

now a polygonal free boundary is to be expected.

The *practical* problem is to carry out this minimization and compute u_h; we are in exactly the situation described in the introduction, with many proposed algorithms and a difficult task of comparison and analysis. The *theoretical* problem, which assumes that u_h has somehow been found, is to estimate its distance from the true solution u. We want to report on this latter problem, and it is natural to ask the same four questions about convergence which were answered in the linear case.

The easiest way is to take the questions in reverse order:

4. Superconvergence is almost certainly destroyed by the error in determining the free boundary. Even in one dimension with $u'' = 1$, u differs from u_h by $O(h^2)$.

3. The approximation of a and l by a_h and l_h leads to no new difficulties; the identity (6) simply becomes an inequality, if we combine (8) and (9), and *the patch test is still decisive.* The same is true for nonconforming elements, and the extra term Δ in the error estimates [**1**, p. 178] is exactly copied from the linear case.

2. It is an open problem, both for K and for the discrete K_h, to show how stability can compensate for the indefiniteness of $a(v, v)$.

1. This is the basic question in the nonlinear Ritz-Galerkin method: If the trial functions in K_h can approximate u to a certain accuracy, how close is the particular choice u_h? It is no longer exactly optimal, because it is no longer the projection of u. But we hope to prove, in the natural norm $\|v\|^2 = a(v, v)$, that $\|u - u_h\| \leqq c \min \|u - v_h\|$.

First, we ask how large this minimum is, choosing v_h to be the piecewise polynomial u_I in S_h which interpolates u at the finite element nodes. The answer depends on the degree of the polynomial and on the regularity of u. For our obstacle problem, with $-\Delta u = f$ in the elastic part and $u = \psi$ in the plastic part, it is now known that u lies in $W^{2,\infty}$. (Brezis and Kinderlehrer announced this long-sought result in Vancouver.) At the free boundary there is a jump in the second derivative of u, which absolutely limits the accuracy of the interpolation. Courant's linear approximation, on triangles of size h, is still of order $\|u - u_I\| = O(h)$. But for polynomials of higher degree, and a smooth free boundary, this is improved only to $O(h^{3/2})$—and *no elements can do better.* There are $O(1/h)$ triangles in which the gradient is in error by $O(h)$. Therefore there is no justification for using cubic polynomials, and the question is whether quadratics are worthwhile; we don't know.

To prove that the actual error $u - u_h$ is of the same order as $u - u_I$, we depend on an a priori estimate of Falk [**2**]. It resembles (4), but the change in (8) and (9)

from equations to inequalities produces a new term:

$$\|u - u_h\|^2 \leqq \|u - v_h\|^2 + 2\int(f + \Delta u)\,(u - v_h + u_h - v). \tag{10}$$

We may choose any v_h in K_h and any v in K—and for simplicity we have specialized to $l = l_h = \int fv$ and $a = a_h = \int|\nabla v|^2$. The new term is automatically zero in the elastic part, where $-\Delta u = f$, but elsewhere $f + \Delta u > 0$.

To estimate (10), we take $v = \phi$ and $v_h = u_I$—which lies in K_h because it cannot exceed ϕ at the nodes, where it agrees with u. (In the case of quadratic polynomials, some members of K_h will go above the yield limit ϕ within the triangles, but we have to be generous enough to permit that; it does not hurt the error estimate, and anyway it is only constraints on v_h at nodal "checkpoints" which can be enforced in practice.) With this choice of v and v_h, the terms in (10) are

(i) With Courant's linear finite elements:

$$\|u - u_I\|^2 \sim h^2, \quad \int(f + \Delta u)\,(u - u_I) \sim h^2, \quad \int(f + \Delta u)\,(u_h - \phi) \leqq 0.$$

(ii) With quadratic finite elements:

$$\|u - u_I\|^2 \sim h^3, \quad \int(f + \Delta u)\,(u - u_I) \sim h^3, \quad \int(f + \Delta u)\,(u_h - \phi) \sim h^3.$$

(The next-to-last integral is split into a part completely within the plastic region, where $u - u_I \sim h^3$, and a part formed from those triangles which cross the free boundary. This transition region has area $O(h)$, and the integrand $u - u_I$ is $O(h^2)$.) Substituting back into (10), the rates of convergence are h and $h^{3/2}$ in the two cases—and these rates are confirmed by experiment.

IV. Open problems. True plasticity theory is a deeper mathematical problem than the model we have used above. The reason is that the *history* of the loading f has to be taken into account; a part of the domain can go from elastic to plastic and back again, as the external loads are increased. Therefore incremental theory introduces a time parameter, and a *rate* of loading $\dot{f}$ in the functional J—and it computes the stress rate $\dot{\sigma}$. In other words, as Maier and Capurso have shown, we have *a time-dependent variational inequality*,

$$\min_{\dot{v} \in K(t)} J(\dot{v}) = J(\dot{u}), \quad \text{with } K = \{\dot{v} \in \mathscr{H}_0^1,\ \dot{v} \leqq 0 \text{ where } u(t) = \phi\}. \tag{11}$$

Notice that *at each instant the convex set depends on the current state u.* In a practical problem the state is actually a vector of stresses and plastic multipliers, but we hope that this *quasi-static obstacle problem* will serve as a reasonable model. We also hope that the new results on regularity can be extended to $u(t)$. But even on this assumption, there remain three new problems in numerical analysis:

(i) Keeping time continuous, to prove convergence of the finite element approximations. The difficulty is that the convex set K, and therefore the minimizing $\dot{u}$, depend discontinuously on the current state u; therefore it is not true that $\dot{u}_h$ is close to $\dot{u}$ whenever u_h is close to u.

(ii) To admit finite difference approximations in time, and to determine the stability limits on the interval Δt.

(iii) To find a quick way of solving, with adequate accuracy, the obstacle problem which arises at each time step.

We believe that these are among the most important questions in nonlinear finite element analysis and that answers can be found.

A second class of problems, of an entirely different type, arises if we are interested only in the multiple λ of the load f which will induce plastic collapse. This is known as *limit analysis*, and no longer requires us to follow the loading history. In place of minimizing a quadratic functional, the problem falls into the framework of *infinite-dimensional linear programming*. Here is a typical example, with unknown stresses $\sigma_{ij}(x, y)$ and multiplier λ: *Maximize* λ, *subject to*:

equilibrium: $\sum (\partial\sigma_{ij}/\partial x_i) = \lambda f_j$ in Ω, $\sum \sigma_{ij}\, n_i = \lambda g_j$ on $\partial\Omega$, and

piecewise linear yield conditions: $\sum b_{ij}^{\alpha}\, \sigma_{ij} \leqq c^{\alpha}$ in Ω, $1 \leqq \alpha \leqq M$.

Suppose we make this problem finite-dimensional by assuming that the stresses (and also the displacements, which are the unknowns in the dual program) belong to piecewise polynomial spaces S_h. The continuous linear programming problem is then approximated, in a completely natural way, by a discrete one [3]. But we know nothing about the rate of convergence.

References

1. G. Strang and G. Fix, *An analysis of the finite element method*, Prentice-Hall, Englewood Cliffs, N. J., 1973.

2. R. Falk, *Error estimates for the approximation of a class of variational inequalities*, Math. Comp. (to appear).

3. E. Anderheggen and H. Knöpfel, *Finite element limit analysis using linear programming*, Int. J. Solids and Structures **8** (1972), 1413–1431.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139, U.S.A.

Труды Международного Конгресса Математиков
Ванкувер, 1974

Численные Методы в Теории Дифракции*

А. Г. Свешников

1°. В настоящее время численные методы являются весьма эффективным аппаратом исследования во многих областях естествознания. Большую роль играют они и в электродинамике и, в частности, находят широкое применение при исследовании многих проблем теории дифракции и распространения волн. Особенно эффективными численные методы оказываются в тех случаях, когда их применение связано с разработкой общих универсальных алгоритмов решения достаточно широких классов задач.

На ряде таких общих алгоритмов, с успехом применяющихся в последнее время для решения некоторых традиционных классов задач дифракции и распространения волн мы и остановимся в настоящем докладе.

2°. В качестве первого класса задач рассмотрим задачи дифракции на теле T, находящемся в среде с локально неоднородными характеристиками. Если длина падающей волны соизмерима с размерами области, в которой характеристики среды являются переменными функциями координат, то для решения этого класса задач оказываются наиболее эффективными прямые численные методы, основанные на идеях метода Галеркина. Задача сводится к определению решения системы уравнений Максвелла

$$\text{rot } \vec{H} = -i\omega\hat{\varepsilon}\vec{E} + \vec{j}; \qquad \text{rot } \vec{E} = i\omega\hat{\mu}\vec{H} \tag{1}$$

удовлетворяющего граничному условию импедансного типа

$$[\vec{n}\vec{E}]\big|_S = w[\vec{n}[\vec{n}\vec{H}]]\big|_S \tag{2}$$

на поверхности S тела T и условиям излучения на бесконечности. В силу локального характера неоднородности тензоры $\hat{\varepsilon}$ и $\hat{\mu}$ являются переменными

*Not presented in person.

функциями координат в области D, ограниченной поверхностью S и некоторой поверхностью Σ, содержащей тело T. В качестве поверхности Σ может быть, в частности, выбрана сфера Σ_R достаточно большого радиуса R с центром 0 внутри тела T. Вне Σ тензоры $\hat{\varepsilon} \equiv \varepsilon_0 = \text{Const}$, $\hat{\mu} \equiv \mu_0 = \text{Const}$. Тем самым исходная задача является внешней краевой задачей для эллиптической системы с локально переменными коэффициентами, причем в силу условий излучения задача оказывается несамосопряженной. При достаточных условиях гладкости функций $\hat{\varepsilon}$ и $\hat{\mu}$ и поверхности S решение поставленной задачи существует и единственно [1].

3°. Прямое применение численных методов к решению внешней задачи затруднительно. Можно перейти к внутренней задаче для ограниченной области. Для этого надо вместо обычных условий излучения, определяющих асимптотическое поведение решения на бесконечности, сформулировать так называемые "парциальные" условия излучения [2]. Последние формулируются в виде некоторых интегральных соотношений, которым должно удовлетворять решение на поверхности Σ. Наиболее простой вид эти соотношения имеют в том случае, когда Σ—сфера Σ_R радиуса R. В этом случае "парциальные" условия излучения принимают вид [3]

$$\int_{\Sigma_R} [\vec{E}\vec{H}_k^*]_n \, d\sigma = C_k\beta_k \qquad (k = 1, 2, \cdots), \tag{3}$$

где $\vec{H}_k$—поперечная часть магнитного вектора парциальной сферической волны, C_k—коэффициент разложения поля $\{\vec{E}, \vec{H}\}$ вне Σ_R по расходящимся сферическим волнам $\{\vec{E}_k, \vec{H}_k\}$, β_k—нормировочный коэффициент.

4°. С помощью метода Галеркина задача (1)—(3) может быть сведена к краевой задаче для системы обыкновенных дифференциальных уравнений. При этом данный метод удобно применять в регулярной области. Поэтому исходную область D отобразим на шаровой слой K между двумя концентрическими сферами S_{r_1} и S_{r_2}. При этом в уравнениях (1) тензоры $\hat{\varepsilon}$ и $\hat{\mu}$ надо заменить на новые значения $\hat{\varepsilon}'$ и $\hat{\mu}'$, выражающиеся через $\hat{\varepsilon}$, $\hat{\mu}$ и метрический тензор $\hat{g}$ преобразования координат.

Введем системы вектор-функций $\{\vec{e}_k\}$ и $\{\vec{h}_k\}$ полные на любой концентрической сфере S_r $(r_1 \leqq r \leqq r_2)$ внутри K и будем искать поперечные части приближенного решения в виде

$$\vec{E}_t^N = \sum_{n=1}^{N} a_n(r)\vec{e}_n; \qquad \vec{H}_t^N = \sum_{n=1}^{N} b_n(r)\,\vec{h}_n. \tag{4}$$

Радиальные компоненты приближенного решения определим из уравнений (1).

Для неизвестных коэффициентов $a_n(r)$ и $b_n(r)$ получим систему обыкновенных дифференциальных уравнений, потребовав выполнение на любой сфере S_r $(r_1 < r < r_2)$ следующих условий ортогональности

$$\int_{S_r} (\operatorname{rot} \vec{E}^N - i\omega\hat{\mu}'\vec{H}^N)_t \, \vec{h}_n^* \, d\sigma' = 0, \tag{5}$$

$$\int_{S_r} (\operatorname{rot} \vec{H}^N + i\omega\ddot{\varepsilon}'\vec{E}^N - \vec{j})_t \, \vec{e}_n^* \, d\sigma' = 0 \qquad (n = 1, 2, \cdots, N).$$

Граничные условия для этой системы получим из (2) и (3), заменив их следующими интегральными условиями:

$$\int_S ([\vec{n}\vec{E}^N] - w[\vec{n}[\vec{n}\vec{H}^N]])\vec{h}_k^* \, d\sigma = 0 \qquad (k = 1, 2, \cdots, N) \tag{6}$$

и

$$\int_{\Sigma_R} [\vec{n}\vec{E}^N]\vec{h}_m^* \, d\sigma = \sum_{k=1}^{\infty} \alpha_{mk}^* \beta_k C_k^N \qquad (m = 1, 2, \cdots, N), \tag{7}$$

где коэффициенты α_{mk} и C_k^N определяются из соотношений

$$\vec{H}_t^N\big|_{\Sigma_R} = \sum_{m=1}^{N} b_m(r_2)\vec{h}_m = \sum_{k=1}^{\infty} C_k^N \vec{H}_k, \tag{8}$$

$$C_k^N = \sum_{m=1}^{N} \alpha_{mk} \, b_m(r_2). \tag{9}$$

Задача (5)—(7) является краевой задачей для системы обыкновенных дифференциальных уравнений на ограниченном промежутке $r_1 \leqq r \leqq r_2$.

Если тензоры $\ddot{\varepsilon}$ и $\ddot{\mu}$ имеют вид $\ddot{\varepsilon} = \varepsilon I + \ddot{\varepsilon}_1$, $\ddot{\mu} = \mu I + \ddot{\mu}_1$, где $\ddot{\varepsilon}_1$ и $\ddot{\mu}_1$ —эрмитовы тензоры, а ε и μ—скалярные функции с положительной мнимой частью, то оператор краевой задачи (5)—(7) является положительным в подходящим образом определенном энергетическом пространстве. Отсюда следует, что данная задача всегда разрешима. Более того решение задачи (5)—(7) оказывается равномерно ограниченным по N и при $N \to \infty$ сходится к решению исходной задачи (1)—(3) в соответствующей норме [3]. При этом коэффициенты в разложении (8) сходятся к соответствующим коэффициентам разложения точного решения, что представляет особый интерес при решении многих практических задач.

Для численного решения системы (5)—(7) применимы хорошо разработанные методы прогонки с ортогонализацией. Примеры практической реализации данного алгоритма для решения конкретных задач имеются, например, в работах [4], [5].

Эффективность предложенного алгоритма зависит как от способа отображения исходной области D на шаровой слой K, так и от выбора системы базисных функций $\{\vec{e}_n, \vec{h}_n\}$. Обычно эти вопросы решаются исходя из соображений простоты реализации алгоритма. Вопрос определения "оптимального" алгоритма решения данного класса задач остается открытым.

5°. В настоящее время достигнут значительный прогресс в разработке устойчивых методов решения некорректно-поставленных задач математической физики [6]. Полученные здесь результаты широко используются при решении задач синтеза электродинамических устройств, возбуждающих электромагнитное поле с заданными характеристиками. Эти задачи по постановке очень близки к обратным задачам электродинамики, которые формулируются как задачи восстановления источников поля или системы рассеивающих тел по наблюдаемым характеристикам поля. Однако, имеются и

существенные отличия в постановке задач синтеза и обратных задач. В частности, возможны и такие постановки задач синтеза, где не требуется единственность решения, удовлетворяющего поставленным условиям. Это позволяет проводить отбор решений с помощью дополнительных условий предпочтительности [7], [8].

Пусть электромагнитное поле u, возбуждаемое электромагнитным устройством z, определяется выражением $u = A[z]$, причем оператор A^{-1} —неограниченный. Широкий класс задач синтеза электродинамических устройств состоит в оптимизации некоторого функционала $\Phi(u)$ и удовлетворении ряда требований на устройство z. Корректная постановка такой задачи заключается в минимизации функционала

$$M = \Phi(A[z]) + \sum_i \alpha_i \Omega_i[z] \tag{10}$$

при дополнительных ограничениях. Функционалы Ω_i описывают поставленные требования на устройство z, а весовые множители α_i определяют цену каждого из поставленных требований.

6°. Важным классом задач синтеза являются задачи возбуждения электромагнитного поля с заданными характеристиками системой параллельно расположенных полубесконечных открытых волноводов с общим плоским фланцем.

Здесь возможны различные постановки задачи. Например, потребуем, чтобы отношение энергии $L(\omega)$, излученной в заданном угловом секторе ω, ко всей отраженной Q и рассеянной энергии $L(\Omega)$ было максимальным при дополнительном условии задания полной энергии возбуждения L_0. Задача сводится к определению

$$\sup_{a_m^i} \{L(\omega)/(L(\Omega) + Q)\}$$

при дополнительном условии $L_0[a_m^i] = 1$. Для вычисления значений функционалов $L(\omega)$, $L(\Omega)$ и Q необходимо сначала найти во всем пространстве электромагнитное поле при заданных параметрах возбуждения a_m^i.

7°. Дополнительные трудности при решении задач этого класса связаны с отсутствием явного аналитического выражения решения соответствующей прямой задачи даже в простейшем случае. Требуется разработка специальных численных алгоритмов их решения. Рассмотрим этот вопрос на примере плоской скалярной задачи.

Прямая задача в этом случае состоит в определении решения уравнения Гельмгольца

$$\Delta u + k^2 u = -f \tag{11}$$

удовлетворяющего однородному условию $\partial u/\partial n = 0$ на границе области и условиям излучения на бесконечности. Решение этой задачи существует и единственно, причем в окрестности угловых точек границы $\partial u/\partial \rho$ имеет

особенность порядка $\rho^{-\alpha_i}$, где $0 \leqq \alpha_i \leqq \frac{1}{2}$ и ρ —расстояние до угловой точки.

Эта задача является частным случаем общей задачи об электромагнитной связи областей D_i через отверстия S_j ($j = 1, 2, \cdots, N$) в общих участках границы. На S_j должны выполняться условия сопряжения

$$[u]_{S_j} = 0, \qquad \sigma_i \frac{\partial u_i}{\partial n}\bigg|_{S_j} = \sigma_k \frac{\partial u_k}{\partial n}\bigg|_{S_j} = \mu_j(x_j), \tag{12}$$

где σ_i определяются характеристиками среды i-ой области, а отверстие S_j связывает i-ую и k-ую области. Задача (11)—(12) сводится к системе сингулярных интегральных уравнений первого рода

$$\sum_j \int_{S_j} K_{lj}(x_l, \xi_j)\, \mu_j(\xi_j)\, d\xi_j = \varphi_l(x_l), \qquad \xi_j \in S_j,\ x_l \in S_l \quad (l = 1, 2, \cdots, N). \tag{13}$$

Функции $\varphi_l(x_l)$ определяются способом возбуждения, ядра K_{ll} являются суммой функций Грина областей, связанных через отверстие S_l, ядра K_{lj} ($j \neq l$) выражаются через функцию Грина области, в границу которой входят участки S_l и S_j. В более общем случае [9] получим систему интегро-дифференциальных уравнений типа (13).

Уравнения системы (13) являются уравнениями первого рода. Устойчивый алгоритм их численного решения основан на методе, использующем особенность ядра K_{ll} [10]. При реализации алгоритма существенно используются методы обращения дифференциального оператора [9], методы выделения особенности ядра, имеющего сложную аналитическую структуру [11], и методы выделения особенности решения интегрального уравнения [12].

8°. Рассмотренные алгоритмы решения задач дифракции и распространения волн, основанные на проекционных методах и методах интегральных уравнений, успешно переносятся на задачи дифракции на периодических структурах. В случае периодического возбуждения [13] задача сводится к определению решения в полосе, на границах которой должны выполняться условия Флоке

$$u(x + a, y, z) = e^{i\alpha} u(x, y, z). \tag{14}$$

Здесь a—период структуры.

Случай локального возбуждения, когда решение уже не удовлетворяет условиям Флоке, также может быть сведен к задаче в полосе [14], определяемой периодом структуры, но уже не для исходного решения, а для его специального преобразования

$$U(t, x, y, z) = \sum_{n=-\infty}^{\infty} u(x + na, y, z) e^{-int}. \tag{15}$$

Очевидно, функция $U(t, x, y, z)$ удовлетворяет условиям Флоке.

Литература

1. C. Müller, *Grundprobleme der mathematischen Theorie der electromagnetischer Schwingungen*, Springer-Verlag, Berlin, 1957.

2. А.Г. Свешников, Доклады АН СССР **184** (1969), 63–65.
3. А.Г. Свешников и А.С. Ильинский, Журн. Вычисл. Мат. и Мат. Физ. **11** (1971), 960–968.
4. А.Г. Свешников, А.С. Ильинский и А.Л. Павлов, Радиотехника и Электроника **17** (1972), 1387–1391.
5. А.Г. Свешников и Ю.А. Еремин, Журн. Вычисл. Мат. и Мат. Физ. **14** (1974), 499–504.
6. А.Н. Тихонов и В.Я. Арсенин, *Методы решения некорректных задач*, "Наука", Москва, 1974.
7. А.Н. Тихонов и В.И. Дмитриев, *Вычислительные методы и программирование*, Издат. Москов. Гос. ун-та, вып. XIII (1969), 209–215.
8. А.Г. Свешников и А.С. Ильинский, Доклады АН СССР **204** (1972), 1077–1080.
9. А.Г. Свешников и В.М. Репин, *Вычислительные методы и программирование*, Издат. Москов. Гос. ун-та, вып. XX (1973), 12–21.
10. А.Н. Тихонов и В.И. Дмитриев, *Вычислительные методы и программирование*, Издат. Москов. Гос. ун-та, вып. X (1968), 3–8.
11. В.М. Репин, Журн. Вычисл. Мат. и Мат. Физ. **11** (1971), 152–163.
12. А.С. Ильинский, В.М. Репин и Л.И. Старостин, *Вычислительные методы и программирование*, Издат. Москов. Гос. ун-та, вып. XX (1973), 38–49.
13. А.С. Ильинский, Журн. Вычисл. Мат. и Мат. Физ. **13** (1973), 119–126.
14. ———, Журн. Вычисл. Мат. и Мат. Физ. **14** (1974).

Московский Государственный Университет
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Invariant Subspaces

J. H. Wilkinson

1. Introduction. In the last twenty years great advances have been made in the solution of the standard eigenvalue problem $Ax = \lambda x$, and as far as dense matrices are concerned one of the few outstanding problems is the development of a satisfactory general purpose algorithm to deal with nonnormal matrices having ill-conditioned eigensystems. Perhaps the major difficulty is to decide precisely what such an algorithm should attempt to do.

The most widely used algorithm for real nonnormal matrices is that based on a preliminary orthogonal reduction to Hessenberg form followed by the use of the double Francis QR algorithm [1]. The combined process gives an orthogonal reduction of the original A to triangular form T (or rather to quasi-triangular form since complex conjugate pairs of eigenvalues give rise to 2×2 blocks on the diagonal) and is therefore a practical realisation of the Schur canonical form.

This algorithm is in one sense almost best possible since the computed T is exactly orthogonally similar to $A + E$ where $\|E\|_2/\|A\|_2$ is of the order of the machine precision. However, the output of algorithms based on this reduction consists of n eigenvalues and n eigenvectors notwithstanding the fact that if A is defective it will not have n independent eigenvectors. This decision was not taken in ignorance of the facts of life; it was recognised to be an interim measure. The state of the art has now been reached when serious efforts might well be made to deal with this shortcoming.

2. Ill-conditioned eigensystems. It is natural to ask "Why not compute the Jordan canonical form J and the associated X such that $AX = XJ$?" The columns of X would then give the *n principal vectors*. The difficulty is a purely practical one. It is not generally known in advance whether a matrix is defective, and one has to depend on *computed* eigenvalues for making this decision. A matrix can be de-

fective only if it has multiple eigenvalues, but defective eigenvalues are, in general, very sensitive to perturbations in the elements of A. If, e.g., A has an eigenvalue λ_i of multiplicity r associated with an elementary divisor of degree r then the eigenvalues $\lambda_i(\varepsilon)$ of $A+\varepsilon B$ include values such that $\lambda_i(\varepsilon) - \lambda_i = O(\varepsilon^{1/r})$ as $\varepsilon \to 0$. Hence even if one has exact eigenvalues of some $A+E$ the computed values may not include any that appear unduly close. How are we to recognize that they 'belong' to a set of r equal eigenvalues?

However, it is a mistake to concentrate too heavily on defective matrices. It is more instructive to study the problems presented by a simple ill-conditioned eigenvalue, λ_i. In this case $A + \varepsilon B$ (with $\|B\|_2 = 1$) has an eigenvalue $\lambda_i(\varepsilon)$ such that

$$\lambda_i(\varepsilon) - \lambda_i \sim \varepsilon(y_i^H Bx_i)/(y_i^H x_i) \quad \text{as } \varepsilon \to 0,$$

where y_i and x_i are (normalised) left-hand and right-hand eigenvectors. Hence we have $|\lambda_i(\varepsilon) - \lambda_i| \leqq \varepsilon s_i$ where $s_i = y_i^H x_i$ is the cosine of the angle between y_i and x_i. The bound is attained when $B = y_i x_i^H$. Not surprisingly a matrix having a small s_i is close to a matrix having a multiple eigenvalue. Indeed there is a matrix $A + E$ with $\|E\|_2 \leqq \|A\|_2 s_i$ having λ_i as an eigenvalue of multiplicity not less than two. This is rather a weak result; in general there is a much nearer matrix having a double eigenvalue at a point other than λ_i.

A further problem associated with ill-conditioned nondefective matrices is that their eigenvectors are almost linearly dependent. In fact if X is the matrix of right-hand eigenvectors, $y_i^H X = [0, 0, \cdots, s_i, \cdots, 0]$ and hence $\sigma_n(X) \leqq s_i$ showing that if s_i is small X is almost singular. This means that even if, in spite of the practical difficulties, one obtains quite accurate eigenvectors one cannot compute from them an accurate orthogonal basis for the corresponding invariant subspace. This is not surprising because when A is truly defective, some of the eigenvectors are coincident and a complete basis is constructed by adding principal vectors.

There have been then two main strands in the research on ill-conditioned eigenvalue problems. In the first a direct attempt is made to find the Jordan canonical form and an associated basis of a truly defective matrix in the neighbourhood of the given A. In the second, eigenvalues are grouped together in some appropriate manner and orthogonal bases are determined for the corresponding invariant subspaces.

3. The Jordan canonical form (J.c.f.). Three essentially different algorithms have been developed for computing the J.c.f. Suppose λ is a multiple eigenvalue; then $B = A - \lambda I$ has a corresponding multiple eigenvalue equal to zero.

Method 1. This is based on the observation that the space spanned by the principal vectors of grades 1, 2, ..., r is the null space of B^r. Since one is concerned with rank determination, the singular value decomposition (S.V.D.) is used throughout. The general step is as follows. Suppose we have derived the relation

$$B^r W_r = [A_r | 0]$$

where W_r is an orthogonal matrix and the last m_r columns of the R.H.S. are null.

Then the last m_r columns of V_r give an orthogonal basis of vectors of grades $1, \cdots, r$. We now have

$$B^{r+1}W_r = [BA_r|0] \equiv [C_{r+1}|0]$$

and if the S.V.D. of C_{r+1} is given by $C_{r+1} = U_{r+1}\Sigma_{r+1}V_{r+1}^H$ then Σ_{r+1} has n_{r+1} zero singular values. Writing

$$\tilde{V}_{r+1} = \left[\begin{array}{c|c} V_{r+1} & 0 \\ \hline 0 & I \end{array}\right], \qquad W_{r+1} = W_r\tilde{V}_{r+1},$$

then

$$B^{r+1}W_{r+1} = [U_{r+1}\Sigma_{r+1}|0] = \left[\begin{array}{c|c|c} A_{r+1} & 0 & 0 \\ & m_{r+1} & n_r \end{array}\right].$$

The last m_r columns of W_{r+1} are those of W_r. It is now clear that $m_r = n_1 + \cdots + n_r$, and in general the last m_{r+1} columns of W_{r+1} give principal vectors of grades $r+1, \cdots, 1$. The process comes to an end when C_{r+1} has no zero singular values. An outstanding difficulty of the method is providing a satisfactory criterion for determining whether elements of Σ_{r+1} may be regarded as zero.

Method 2. This involves a systematic use of orthogonal deflation. The matrix $B \equiv B^{(0)}$ is reduced successively to $B^{(1)}$, $B^{(2)}$, $\cdots$ by orthogonal similarities. The form of $B^{(r)}$ and its relationship to B is adequately illustrated by the fact that

$$B^{(2)} = (W^{(2)})^TBW^{(2)} = \left[\begin{array}{c|c|c} B_{11}^{(2)} & 0 & 0 \\ \hline B_{21}^{(2)} & 0 & 0 \\ \hline B_{31}^{(2)} & B_{32}^{(2)} & 0 \end{array}\right] \begin{array}{l} n-m_2 \\ n_2 \\ n_1 \end{array}$$
$$\qquad\quad n-m_2 \quad n_2 \quad n_1$$

where $W^{(2)}$ is orthogonal. If the multiplicity of the zero eigenvalue is $m_2 = n_1 + n_2$ the process is complete. Otherwise since $B_{11}^{(2)}$ has the remaining eigenvalues of B it must be singular. If the S.V.D. of $B_{11}^{(2)}$ is $B_{11}^{(2)} = U^{(3)}\Sigma^{(3)}(V^{(3)})^T$ then

$$(V^{(3)})^TB_{11}^{(2)}V^{(3)} = (V^{(3)})^TU^{(3)}\Sigma^{(3)} \equiv Z^{(3)}\Sigma^{(3)} = \left[\begin{array}{c|c} B_{11}^{(3)} & 0 \\ \hline B_{21}^{(3)} & 0 \end{array}\right] \begin{array}{l} n-m_3 \\ n_3 \end{array}$$
$$\qquad\qquad n-m \quad n_3$$

where n_3 is the number of zero singular values of $B_{11}^{(2)}$. Hence writing

$$\tilde{V}^{(3)} = \left[\begin{array}{c|c} V^{(3)} & 0 \\ \hline 0 & I \end{array}\right], \qquad W^{(3)} = W^{(2)}\tilde{V}^{(3)},$$

we have

$$B^{(3)} = (W^{(3)})^TBW^{(3)} = \left[\begin{array}{c|c|c|c} B_{11}^{(3)} & 0 & 0 & 0 \\ \hline B_{21}^{(3)} & 0 & 0 & 0 \\ \hline B_{31}^{(3)} & B_{32}^{(3)} & 0 & 0 \\ \hline B_{41}^{(3)} & B_{42}^{(3)} & B_{43}^{(3)} & 0 \end{array}\right] \begin{array}{l} n-m_3 \\ n_3 \\ n_2 \\ n_1 \end{array}$$
$$\qquad\quad n-m_3 \quad n_3 \quad n_2 \quad n_1$$

From the definition of the S.V.D. it follows that $B_{i+1,i}^{(k)}$ is of full rank for $i \leqq k$ and from this it is easy to deduce that the trailing columns of $W^{(3)}$ give orthogonal bases for vectors of grades 3, 2, 1.

Decisions on zero elements of the $\Sigma^{(r)}$ are much simpler than with Method 1. Both methods are essentially due to Kublanovskaya though she did not use the S.V.D. and the alogrithms were not described in these terms. An outstanding weakness of both methods is that in general they require the computation of the S.V.D.'s of matrices of high order.

Method 3. This is based on the observation that if x is a principal vector of grade $r+1$ then Bx is a nonnull vector of grade r and hence lies in the subspace spanned by the principal vectors of grades 1, 2, $\cdots$, r. However if we were to obtain all such nonnull x they would, of course, include the previously computed vectors of grades 2, $\cdots$, r. A technique is needed which will exclude them. If $B = U\Sigma V^H$ then $U^HBU = \Sigma V^HU \equiv \Sigma W^H$ (say) and it is simpler to work with ΣW^H rather than with B itself. The step in which vectors of grade $k+1$ are computed may then be described as follows.

Let the vectors of grade s be denoted by $u_1^{(s)}, \cdots, u_{n_s}^{(s)}$ $(s = 1, \cdots, k)$. Then we require solutions of

$$\Sigma(W^Hx) \equiv \Sigma y = \left[p_1^{(k)}p_2^{(k)}\cdots p_{n_1-n_s}^{(k)}u_1^{(k)}\cdots u_{n_s}^{(k)}\right]\alpha \equiv W^{(k)}\alpha$$

where α is a vector of order n_1 which is to be chosen so that solutions exist. The source of the $p_i^{(k)}$ will become evident at the end of this step. The diagonal matrix Σ has zero elements in the last n_1 positions and solutions y are given for any α which gives zero components in the last n_1 positions of the right-hand side. To determine such α we write

$$W^{(k)} = \left[\begin{matrix} W_1^{(k)} \\ \hline W_2^{(k)} \\ n_1 \end{matrix}\right] \begin{matrix} n - n_1 \\ n_1 \end{matrix}$$

and denoting the S.V.D. of $W_2^{(k)}$ by $W_2^{(k)} = U_k\Sigma_kV_k^H$ we have

$$\Sigma y = \left[\frac{W_1^{(k)}V_k}{U_k\Sigma_k}\right](V^H\alpha) \equiv \left[Z^{(k)}\right]\beta,$$

and if Σ_k has n_{k+1} zero elements, the last n_{k+1} columns of $Z^{(k)}$ give appropriate right-hand sides. Hence we obtain n_{k+1} vectors y and the n_{k+1} corresponding x give $u_1^{(k+1)}, \cdots, u_{n_{k+1}}^{(k+1)}$, the vectors of grade $k+1$. The remaining $n_1 - n_{k+1}$ columns of $Z^{(k)}$ are the vectors $p_1^{(k+1)}, \cdots, p_{n_1-n_{k+1}}^{(k+1)}$ of the next stage. This method is much more economical than the previous two, since we have only to compute the S.V.D. of an $n_1 \times n_1$ matrix at each stage.

4. Orthogonal bases of invariant subspaces. Even when using the S.V.D. the methods of the previous section are not particularly satisfactory from the point of view of numerical stability. Moreover they force one to treat matrices which are not truly defective as though they were. There is a good deal to be said for grouping eigenvalues together "appropriately" and finding an orthogonal basis for the

corresponding invariant subspace. Naturally the criterion should be such that when A really is defective the corresponding eigenvalues will be grouped together. A good starting point is the triangular matrix T given by the orthogonal reduction described in §1. This matrix is exactly orthogonal to $A+E$ with $\|E\|_2/\|A\|_2$ of the order of the computer precision. From the triangular form, the left-hand and right-hand vectors corresponding to each of the eigenvalues are easily and accurately computable and thence the values of the s_i for the triangle. It is clear that if $\|E\|[1/s_i + 1/s_j] > |\lambda_i - \lambda_j|$ the eigenvalues λ_i and λ_j may well correspond to coincident eigenvalues in the original matrix A and should certainly be grouped together. In practice one needs a less severe criterion than this; otherwise eigenvalues will not be grouped together even when they are associated with an almost defective matrix. A useful criterion in practice has been to group together eigenvalues for which

$$|\lambda_i - \lambda_j| \max(s_i, s_j) \leqq 2^{t_1-t}\|A\|_F$$

where computation is on a t binary digit computer and subspaces are required which are correct to roughly $t_1 (\leqq t)$ binary places. Notice that if we take $t_1 = t$ we will usually be forced to group all the eigenvalues together.

Having decided which eigenvalues to group together probably the simplest method of determining an orthogonal basis for the corresponding invariant subspace is to subject T to an orthogonal similarity which retains triangularity and brings the grouped eigenvalues into the leading diagonal positions. At this stage we have (ignoring rounding errors)

$$Q^TAQ = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$$

where the r grouped eigenvalues are the diagonal elements of the triangular matrix T_{11} of order r. The required orthogonal basis is then the first r columns Q_1 of Q. Ruhe [3] has described a simple method for bringing the required eigenvalues into the leading positions by a series of plane rotations. This can be generalised to cover the case when there are 2×2 blocks on the diagonal of T corresponding to complex conjugate pairs of eigenvalues.

Having obtained an approximate orthogonal basis in this way it may be improved by a process of iterative refinement analogous to that used for linear systems. For this it is necessary to be able to compute the residual matrix R defined by $R = AQ_1 - Q_1T_{11}$ and this must be done working to higher accuracy, preferably by accumulating inner-products in double precision.

References

1. J. G. F. Francis, *The QR transformation: A unitary analogue to the LR transformation.* I, II, Comput. J. **4** (1961/62), 265–271, 332–345. MR **23** #B3143; **25** #744.

2. V. N. Kublanovskaya, *A method for solving the complete problem of eigenvalues of a degenerate matrix*, Ž. Vyčisl. Mat. i Mat. Fiz. **6** (1966), 611–620 = U.S.S.R. Comput. Math. and Math. Phys. **6** (1968), no. 4, 1–14. MR **34** #943.

3. A. Ruhe, *An algorithm for numerical determination of the structure of a general matrix*, BIT **10** (1970), 196–216.

4. ———, *Perturbation bounds for means of eigenvalues and invariant subspaces*, BIT **10** (1970), 343–354. MR **42** #8678.

5. J. M. Varah, *Rigorous machine bounds for the eigensystem of a general complex matrix*, Math. Comp. **22** (1968), 793–801. MR **39** #5052.

6. ———, *Computing invariant subspaces of a general matrix when the eigensystem is poorly conditioned*, Math. Comp **24** (1970), 137–149. MR **41** #9434.

7. W. Kahan, *Conserving confluence curbs ill-condition*, Computer Science Technical Report 6 of the University of California, 1972.

8. J. H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon Press, Oxford, 1965. MR **32** #1894.

9. G. H. Golus and J. H. Wilkinson, *Ill-conditioned eigensystems* (to appear).

NATIONAL PHYSICAL LABORATORY
TEDDINGTON, MIDDLESEX, ENGLAND

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Difficulty in Problems of Optimization

Philip Wolfe

1. Introduction. The concept of *difficulty*, or *computational complexity*, is central to the mathematization of the art of computation. A *problem* (short for *class of problem*) is a function from a certain set of mathematical objects to others: e.g., (a) from pairs of integers to their sum, (b) from systems of linear equations to their solutions, (c) from analytic functions to their sets of zeroes. Given a vocabulary of operations on objects of the domain, an *algorithm* is a sequence of operations yielding the value of the function for any member of its domain. The number of operations—or time, or other measure of work—an algorithm takes in evaluation, for the worst case in its domain, is the *difficulty of the problem-algorithm pair;* and the minimum of problem-algorithm difficulty over all algorithms solving the problem is the *difficulty of the problem.*

2. Combinatorial problems. These notions have their sharpest formulation in the first of three kinds of problem we discuss, the "combinatorial" type like (a), whose domain and value is always a finite set of integers. Results of fundamental importance for computation are typified by the work of Winograd [1, Bibliography, 1965A, 1967A],[1] who determined the difficulty of addition and multiplication for pairs of integers in terms of the gating operations used in real computers: e.g., $\log_2 \log_2 N$ gatings for adding two integers not greater than N, no matter how they are represented.

In combinatorial problems of optimization we know the difficulty of many problem-algorithm pairs, but few problem difficulties. One tries to summarize that information as follows: For a specific problem, let L be the "length of the input string", the number of bits required to encode a point in the domain of the algo-

[1] For compactness, applicable references are made to the bibliography of our [1, pp. 187–122]. [1] constitutes an excellent survey of the subject of computational complexity.

rithm A. One attempts to determine a simple function $d_A(L)$ bounding the number of steps needed by A to solve the worst problem of length L, and represent problem difficulty by $D(L) = \text{Min}_A\ d_A(L)$. The most closely studied question has been whether particular problems are solvable in *polynomial time* ($D(L)$ has polynomial growth, or slower) or *exponential time* (there is no polynomial bound for D). Many problems are known to be "polynomial" by exhibition of a suitable A; finding the shortest paths between two points of a graph, or between all pairs, and the "assignment" problem [2] of finding that permutation P of $(1, 2, \cdots, n)$ minimizing $\sum_i M_{i,P(i)}$ for a given n-order matrix M, are famous examples. It is not known whether any interesting combinatorial optimization problem is exponential, but there is good reason to think that many of them are. Karp [1, Bibliography, 1972B] has listed 21 problems of importance and shown them equivalent with respect to polynomial-time solvability. His problems, on some of which much effort has been expended for the discovery of efficient algorithms, include

(i) integer programming: given an integer-valued matrix C and integer vector d determine the existence of a vector x of zeroes and ones such that $Cx = d$;

(ii) Hamiltonian circuits: given a finite graph, determine whether it has a cycle containing each node exactly once;

(iii) three-dimensional matching: given T finite and $U \subseteq T \times T \times T$, determine the existence of $W \subseteq U$ such that no two elements of W agree in any coordinate and W has the cardinality of T. (If T^3 of (iii) is replaced by T^2, the problem is polynomial. Defining $M_{ij} = -1$ if $(i, j) \in U$, $= 0$ otherwise, it asks whether the minimum in the assignment problem above is $-|T|$.)

3. Finite numerical problems. The assignment problem connects combinatorial problems with finite numerical problems: problems whose domains are sets of real numbers, and algorithms whose operations are those of real arithmetic (addition, division, comparison, $\cdots$), which solve the problems exactly using finitely many operations. The convex hull of the set of permutation matrices is the set of all doubly-stochastic matrices of order n, i.e., such that $X_{ij} \geqq 0$, $\sum_i X_{ij} = 1$, $\sum_j X_{ij} = 1$ for all i, j, and each matrix in any set of permutation matrices is an extreme point of the convex hull of the set. The assignment problem can thus be solved by solving the linear programming problem of minimizing $\sum_{i,j} M_{ij} X_{ij}$ under the above constraints. This formulation leads to a variety of special algorithms requiring only $O(n^2)$ operations, owing to the simple structure of the constraints, although its solution implicitly makes a choice among $n!$ objects. (Although any permutation-sum problem has similarly an equivalent linear program, it is usually hard to find and of impractical size. The "traveling-salesman problem" is that of finding the *cyclic* permutation P minimizing $\sum_i M_{i,P(i)}$. Kuhn [3] found for $n = 5$ (24 permutations) the convex hull, of dimension 11, had 390 faces. Since this problem is at least as difficult as (ii) above, it is hard to believe that it has a polynomial-time algorithm.)

The general linear programming problem has the form Min $\sum_j c_j x_j$ subject to the constraints $x_j \geqq 0$ $(j = 1, \cdots, n)$, $\sum_j a_{ij} x_j = b_i$ $(i = 1, \cdots, m)$. The only useful

algorithm for it is the simplex method of Dantzig [4], whose arithmetic operations per step can be closely estimated (and are less than mn); attention has focused on the number $S(m, n)$ of steps required to solve the problem. Empirically, S/m lies between 1 and 3 for most problems, and plausible arguments that $S/m \approx \log(n/m)$ when $n \gg m$ have been substantiated by experiment [5]. Nevertheless, it has exponential complexity. Klee and Minty [6] show the existence for any m of a problem with $n = 3m$, defining geometrically a deformation of the m-cube, for which the simplex method visits every vertex: $S = 2^m - 1$. (Their example assumes the "usual" decision rule for the simplex method; it is not known whether such examples exist for other rules also used in practice.) It seems very likely that the linear programming problem has exponential difficulty.

4. Nonfinite numerical problems. A real-variable optimization problem $\operatorname{Min}\{f(x): x \in S \subseteq R^n\}$ more complicated than that of "quadratic programming"—the minimization of a quadratic function under linear inequality constraints—must be "solved" by the generation of a sequence of points $\{X_k\}$ for which $f_k = f(X_k) \to \operatorname{Min} f$. The useful definition of efficiency has been that of order of convergence (explored in detail by Ortega and Rheinboldt [7]): In one version,

$$p = \operatorname{up}\{q: (f_k - \operatorname{Min} f)/(f_{k-1} - \operatorname{Min} f)^q \text{ is bounded}\}.$$

Standard hypotheses are that f is convex and twice differentiable, or even analytic, and that the functions defining the constraint set S, if any, also are. In that case much is known about the order of convergence for certain important algorithms for optimization in the absence of constraints [8], [9], and something for constrained problems [10].

The richness of possibilities for algorithms in this area makes determining problem difficulty tremendously hard, and there is not general agreement on the important concepts. We outline the approach we think most promising. We cannot suppose the problem domain consists of finite strings, but suppose that as the algorithm proceeds it may call for the computation of f or of some derivatives at certain points. The number of such calls will be the work done by the algorithm, whatever computation done between such calls being ignored. An algorithm is then a sequence g_k of functions, generating X_k by the recursion $X_{k+1} = g_k(X_k, X_{k-1}, \cdots; f(X_k), f(X_{k-1}), \cdots; f'(X_k), f'(X_{k-1}), \cdots; \cdots)$. Algorithms are classified by the information actually used: e.g., a single-point second-derivative algorithm has

$$g_k(\cdots) = h_k(X_k, f(X_k), f'(X_k), f''(X_k)).$$

The central question is: Among all algorithms of a certain class which converge to a solution when started in a sufficiently small neighborhood of the solution, what gives the highest order of convergence?

Answers have been given in two extreme cases. For single-point second-derivative algorithms for minimizing strictly convex, thrice-differentiable functions of n variables, Makurov [11] has shown the limiting order to be 2 which is, of course, achieved by Newton's method: $g_k = [f''(X_k)]^{-1} f'(X_k)$. Results are given for general g_k only for $n = 1$, when much is known about the almost identical problem of find-

ing a simple root of $f(x) = 0$. Brent, Winograd, and Wolfe [12] have shown that if d is the order of the highest derivative used in g_k, and the algorithm converges when started in the neighborhood of a simple root of any analytic function, then the order of convergence cannot be higher than $d + 2$. That order can be approximated as closely as desired by the usual scheme which interpolates all the data with a polynomial and finds a root of that. It is to be expected that polynomial interpolation is likewise optimal even when information is limited, but the problem is subtle, and has not been resolved. The absence of results for $n > 1$ in other than the single-point case can be attributed to the absence of a satisfactory theory of interpolation for functions of several variables. Such results, when they arrive, will be fascinating, for it is hard to believe that our present methods, practically excellent though they be, approach the ultimate in power: and the ability to answer that kind of question will convert optimization from an art into a science.

References

1. R. E. Miller and J. W. Thatcher (editors), *Complexity of computer computations*, Plenum Press, New York, 1972.

2. J. von Neumann, *A certain zero-sum two-person game equivalent to the optimal assignment problem*, Contributions to the Theory of Games, vol. 2, Ann. of Math. Studies, no. 28, Princeton Univ. Press, Princeton, N. J., 1953, pp. 5–12. MR **14**, 998.

3. H. W. Kuhn, *On certain convex polyhedra*, Bull. Amer. Math. Soc. **61** (1955), 557. Abstract #799t.

4. G. B. Dantzig, *Linear programming and extensions*, Princeton Univ. Press, Princeton, N. J., 1963. MR **34** #1073.

5. B. Mercier, *Statistical study of the number of iterations needed in the simplex method in linear progamming*, IBM Research Center, Yorktown Heights, N.Y., September 1970 (manuscript).

6. V. Klee and G. Minty, *How good is the simplex algorithm*? Inequalities III (edited by O. Shisha), Academic Press, New York, 1972, pp. 159–175.

7. J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970. MR **42** #8686.

8. D. M. Himmelblau, *Applied nonlinear programming*, McGraw-Hill, New York, 1972.

9. D. G. Luenberger, *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading, Mass., 1973.

10. P. Wolfe, *Convergence theory in nonlinear programming*, Integer and Nonlinear Programming (edited by J. Abadie), North-Holland, Amsterdam, 1970.

11. V. Ju. Makurov, *On the optimality of the method of Newton-Kantorovič*, J. Computational Math. Phys. **13** (1970), 814–819.

12. R. Brent, S. Winograd and P. Wolfe, *Optimal iterative processes for root-finding*, Numer. Math. **20** (1973), 327–341. MR **47** #6079.

IBM Research Center
Yorktown Heights, New York 10598, U.S.A.

Section 18

Discrete Mathematics and Theory of Computation

Труды Международного Конгресса Математиков
Ванкувер, 1974

Индуктивный Вывод Автоматов, Функций и Программ

Я. М. Барздинь

1°. Настоящий доклад посвящен изложению некоторых результатов, полученных в последние годы по теории индуктивного вывода. В нем рассматривается восстановление описаний (номеров) рекурсивных функций, в том числе автоматов и программ, по их работе на отдельных примерах. Сразу отметим, что мы здесь не будем касаться традиционной теории экспериментов с автоматами, когда о рассматриваемом автомате заранее известна верхняя оценка числа состояний.

2°. Остановимся сначала на одном общем принципе индукции, высказанном еще Лапласом. Согласно этому принципу наиболее простые соотношения (гипотезы) считаются наиболее возможными. Некоторый подход к формализации такого принципа и его применению при экстраполяции последовательностей содержится в работе Соломонова [1]. Другое использование этого принципа было рассмотрено автором [2] при идентификации конечных автоматов.

Автоматы рассматриваются как "черные ящики", с которыми можно проводить кратные эксперименты. Входной и выходной алфавиты считаются известными. Никакая другая информация об автоматах не дана. Уже Муром [3] было замечено, что не возможен эксперимент, идентифицирующий любой такой автомат. Автором был рассмотрен следующий алгоритм идентификации Σ_g, g—неубывающая вычислимая функция. Работа алгоритма Σ_g, примененного к автомату $\mathfrak{A}$, состоит из последовательных шагов $1, 2, 3\cdots$. На каждом шаге порождается определенная гипотеза; A_0—начальная гипотеза—произвольным образом фиксированный автомат. Опишем i-ый шаг. Берется гипотеза A_{i-1}, порожденная на предыдущем шаге ($|A_{i-1}|$—число состояний), и проверяется, совместна ли она с "черным ящиком" $\mathfrak{A}$ на входных словах длины

$g(|A_{i-1}|)$. Если да, то алгоритм останавливается и его результатом $\Sigma_g(\mathfrak{A})$ объявляется гипотеза A_{i-1}. Если нет, то в качестве новой гипотезы A_i рассматривается автомат с минимальным числом состояний ("наиболее простая гипотеза"), который совпадает с $\mathfrak{A}$ на входных словах длины $g(|A_{i-1}|)$, и осуществляется переход к следующему шагу. Алгоритм Σ_g идентифицирует $\mathfrak{A}$, если он останавливается и $\Sigma_g(\mathfrak{A})$ совпадает с $\mathfrak{A}$ на всех входных словах.

Ясно, что описанный алгоритм нигде не использует информацию о верхней оценке числа состояний автомата $\mathfrak{A}$. Тем не менее, оказывается [2], что для любого $\varepsilon<1$ существует алгоритм Σ_g, который идентифицирует ε-ую долю всех автоматов (т.е. для любого k доля автоматов, которые Σ_g идентифицирует, среди всех автоматов с k состояниями, не меньше ε). При этом общая длина входных слов, используемых указанным алгоритмом на автомате $\mathfrak{A}$, не превосходит $|\mathfrak{A}|^c$, c—некоторая константа.

В то же время ясно, что в случае более общих классов функций возможности такой идентификации (будем ее называть прямой) весьма ограничены.

3°. Дальнейшая часть доклада будет посвящена более общему понятию идентификации—предельной идентификации. В данном случае процесс индуктивного вывода рассматривается как предельный процесс, когда сначала выдается одна гипотеза k_1, затем вторая гипотеза k_2, затем k_3 и т.д. Мы говорим, что этот процесс в пределе дает искомый результат, если начиная с некоторого i все гипотезы k_i истинны (хотя мы сами это i можем и не знать и продолжать проверку). По такой схеме происходит, например, отладка программ. Число вариантов программ k_1, k_2, k_3, ···, которые программист перебирает, иногда бывает довольно большим. По такой же схеме осуществляются и различные процессы обучения.

В качестве средств вывода используются произвольные эффективные правила, называемые стратегиями. Формально под стратегиями будем понимать произвольные 1-местные общерекурсивные функции G. В дальнейшем, если не оговорено противное, под φ будем понимать фиксированную допустимую геделевскую нумерацию класса всех 1-местных частично рекурсивных функций. R—класс всех общерекурсивных функций.

Пусть f—некоторая общерекурсивная функция, которая рассматривается как "черный ящик". С ней можно экспериментировать, т.е. для любого x можно узнать, чему равно $f(x)$. Процесс вывода, ассоциированный со стратегией G, определяется как последовательное вычисление значений

$$G(\langle 0, f(0)\rangle),\ G(\langle 0, f(0), 1, f(1)\rangle),\ \cdots,\ G(\langle 0, f(0), \cdots, i, f(i)\rangle),\ \cdots.$$

Эти значения будем обозначать соответственно через

$$k_0, k_1, \cdots, k_i, \cdots$$

и будем называть гипотезами. Гипотеза k_i считается истинной, если $\varphi_{k_i} = f$. Стратегия G предельно идентифицирует f, если начиная с некоторого i все гипотезы k_i истинны и равны между собой. Класс U предельно идентифицируем, если существует стратегия, которая идентифицирует каждую функцию из U.

Понятие предельной идентификации было введено Голдом [4]. Им же были получены и первые результаты.

4°. Первый цикл результатов относится к выяснению принципиальных возможностей предельной идентификации. Ясно, что ее возможности значительно шире по сравнению с прямой идентификацией. Но тем не менее, как показал Голд [4], класс R не может быть предельно идентифицирован. Этот результат был усилен Подниексом [10]. Он показал, что если $\varepsilon > 0$, то класс R не может быть идентифицирован также и с частотой ε. С другой стороны, любой класс общерекурсивных функций, содержащийся в эффективно перечислимом классе общерекурсивных функций, предельно идентифицируем. Блюмом [5] и автором [6], [11] было замечено, что существуют и такие классы общерекурсивных функций, которые идентифицируемы (даже в прямом смысле), но которые не содержатся ни в каком эффективно перечислимом классе общерекурсивных функций.

Блюмом [5] была найдена интересная характеристика классов общерекурсивных функций, предельно идентифицируемых с помощью так называемых строгих стратегий, характеризующихся тем, что для любой $f \in R$, если G не идентифицирует f, то последовательность гипотез $k_0, k_1, \cdots, k_i, \cdots$ не сходится. Это классы, для которых в некотором смысле существует наилучший способ вычисления. Однако вопрос о характеристике классов функций, идентифицируемых с помощью произвольных стратегий, остается открытым.

Автором [11] было показано, что если в определении предельной идентификации отбросить требование, что начиная с некоторого места все гипотезы должны быть равными, то это приводит к увеличению множества предельно идентифицируемых классов общерекурсивных функций.

Стратегию G назовем регулярной на U, если для любой $f \in U$ и любого i гипотеза $k_i = G(\langle 0, f(0), \cdots, i, f(i)\rangle)$ обладает свойством: $\varphi_{k_i}(x) = f(x)$ для $x = 0,1,\cdots,i$. Представляется, что всякая "разумная" стратегия должна обладать свойством регулярности. Вместе с тем можно указать класс $U \subset R$, который в обычном смысле идентифицируем, но с помощью регулярных на U стратегий не идентифицируем.

Фрейвалдом [7] недавно был рассмотрен более общий случай, когда φ— произвольная вычислимая нумерация. Им было показано, что только конечные классы общерекурсивных функций идентифицируемы относительно любой вычислимой нумерации φ. Более того, существует фиксированная вычислимая нумерация φ, относительно которой идентифицируемы только конечные классы общерекурсивных функций.

5°. Следующий цикл результатов относится к оценкам числа изменений гипотез в процессе идентификации. Рассматривается идентификация на произвольной последовательности натуральных чисел

$$\omega = \{x_0, x_1, \cdots, x_i, \cdots\}.$$

Гипотеза

$$k_i = G(\langle x_0, f(x_0), \cdots, x_i, f(x_i)\rangle)$$

называется истинной, если $\varphi_{k_i}(x) = f(x)$ для $x \in \omega$ (φ—допустимая геделевская нумерация). Естественным образом уточняется понятие: G предельно идентифицирует f на ω. Рассмотренное ранее понятие предельной идентификации является частным случаем, когда $\omega = \omega_0 = \{0, 1, \cdots, i, \cdots\}$. Понятие регулярности теперь определяется так: G регулярна на U, если для любых ω, $f \in U$ и i, $\varphi_{k_i}(x) = f(x)$, $x = x_0, x_1, \cdots, x_i$. Положим

$$G^*(w, f) = \begin{cases} \text{число изменений гипотез, если } G \text{ идентифицирует } f \text{ на } \omega, \\ \infty \text{ в противном случае.} \end{cases}$$

Рассматривается идентификация нумерованных классов (U, τ) общерекурсивных функций (τ—вычислимая нумерация класса U). Изучается максимальное число изменений гипотез при идентификации первых n функций класса U:

$$G^*(\omega, n) = \max_{0 \leqq j < n} G^*(\omega, \tau_j).$$

Фрейвалдом и автором [8], [12] доказана теорема:

Для любого нумерованного класса (U, τ) можно построить регулярную на U стратегию G такую, что для любой последовательности ω

$$G^*(\omega, n) \lesssim \log_2 n.$$

Существует нумерованный класс (U, τ) что для любой стратегии G

$$G^*(\omega_0, n) \gtrsim \log_2 n.$$

Для сравнения отметим, что стратегия, которая каждый раз в качестве гипотезы выбирает функцию с наименьшим τ-номером (среди оставшихся функций), дает только оценку вида $G^*(\omega, n) \leqq n$ и порядок этой оценки для таких стратегий в общем случае не может быть понижен. Идея построения стратегии G, дающей оценку $G^*(\omega, n) \lesssim \log_2 n$, опирается на другой принцип. А именно, пусть U—произвольный перечислимый класс всюду определенных функций. Пусть каждой $f \in U$ приписана некоторая "вероятность" $p(f)$ так, что $\Sigma p(f) \leqq 1$. Функцию h назовем функцией наибольшей вероятности для класса $U' \subseteq U$, если для любого x значение $h(x)$ равно наименьшему y такому, что для любого $z \neq y$:

$$\sum_{f \in U' \& f(x) = y} p(f) \geqq \sum_{f \in U' \& f(x) = z} p(f).$$

Рассмотрим следующую, вообще говоря, неэффективную стратегию G: для любых f и i значение $G(\langle x_0, f(x_0), \cdots, x_i, f(x_i)\rangle)$ равно номеру функции наибольшей вероятности для класса $U' = \{g \mid g \in U \,\&\, g(x_0) = f(x_0) \,\&\cdots\&\, g(x_i) = f(x_i)\}$. Легко убедиться, что для любой $f \in U$ имеет место $G^*(\omega, f) \leqq \log_2(1/p(f))$. Отсюда, сопоставляя функции τ_n "вероятность" $p(\tau_n) = c/n(\log_2 n)^2$, получаем $G^*(\omega, \tau_n) \leqq \log_2 n + O(\log_2 \log_2 n)$. Можно показать, что стратегию G можно сделать эффективной, не меняя асимптотику полученной оценки.

Из приведенной теоремы следует, что существуют нумерованные классы

(U, τ), для которых имеются асимптотические наилучшие стратегии идентификации (в смысле величины $G^*(\omega_0, n)$). Подниексом, Кинбером и автором [13] показано, что существуют и такие нумерованные классы (U, τ), для которых возможно любое общерекурсивное уменьшение величины $G^*(\omega_0, n)$. Это верно как для произвольных стратегий, так и для регулярных на U стратегий. В то же время представляется правдоподобным, что для "естественных" классов (U, τ) среди регулярных на U стратегий существуют асимптотически наилучшие. Более подробно эти вопросы еще не исследованы.

6°. Более трудной задачей для нумерованных классов (U, τ) является идентификация τ-номеров. В данном случае гипотеза k_i считается истинной, если $\tau_{k_i}(x) = f(x)$ для $x \in \omega$. $G^\tau(\omega, n)$—соответствующая функция числа изменений гипотез. Оказывается, тривиальная верхняя оценка $G^\tau(\omega, n) \leqq n$ в общем случае не может быть существенно понижена [14]: существует нумерованный класс (U, τ) такой, что для любой стратегии G, $G^\tau(\omega_0, n) \gtrsim n$.

Приведем одно достаточное условие, при котором верхняя оценка может быть понижена до $\log_2 n$. Пусть (U, τ)—нумерованный класс. Для любого кортежа $\langle i_1, \cdots, i_k\rangle$ натуральных чисел рассмотрим набор функций $\tau_{i_1}, \cdots, \tau_{i_k}$. Определим функцию $h_{\langle i_1,\cdots,i_k\rangle}$: для любого x ее значение $h_{\langle i_1,\cdots,i_k\rangle}(x)$ равно наименьшему y такому, что для любого $z \neq y$: $\text{Card}\,\{v | \tau_{i_s}(x) = y\} \geqq \text{Card}\,\{v | \tau_{i_s}(x) = z\}$. Класс (U, τ) назовем полным, если для любого кортежа $\langle i_1, \cdots, i_k\rangle$ эффективно можно построить такое s, что $h_{\langle i_1,\cdots,i_k\rangle} = \tau_s$. Имеет место теорема: Для любого полного нумерованного класса (U, τ) можно построить регулярную на U стратегию G такую, что для любой последовательности ω имеет место $G^\tau(\omega, n) \lesssim \log_2 n$.

Отсюда следует, например, что для класса всех примитивно рекурсивных функций при обычной его нумерации (обозначим ее через τ) имеет место оценка $G^\tau(\omega, n) \lesssim \log_2 n$. Отсюда следует также аналогичная оценка для конечных автоматов. Сформулируем это следствие, выбирая в качестве параметра число состояний автомата. Пусть $U_{a,b}$—класс автоматов с фиксированным a-буквенным входным алфавитом и b-буквенным выходным алфавитом. Пусть ω—произвольная последовательность входных слов, $G^*(\omega, A)$—число изменений гипотез при идентификации автомата A на последовательности ω и $G^*(\omega, U_{a,b}, k) = \max G^*(\omega, A)$, где max берется по всем $A \in U_{a,b}$, имеющим не более k состояний. Тогда существует регулярная на U стратегия G такая, что для любой последовательности ω, $G^*(\omega, U_{a,b}, k) \lesssim (a-1)k \log_2 k$. В [15] показано, что эта оценка асимптотически неулучшаема. В [15] также показано, что аналогичную верхнюю оценку можно получить и в случае, когда выходные алфавиты рассматриваемых автоматов произвольные. Отсюда вытекает одно важное следствие для синтеза программ по историям их работы [16]: существует стратегия G, которая синтезирует программы P по операционно-логическим историям их работы на отдельных примерах, совершая при этом асимптотически не более, чем $\|P\| \log_2 \|P\|$ изменений гипотез (ошибок), $\|P\|$—число логических команд, входящих в P. Это число

относительно небольшое и почти соизмеримое с числом ошибок, которые делает средний программист при программировании традиционными методами. Вопрос о том, имеет ли место аналогичная оценка, когда вместо операционно-логических историй рассматриваются операционные истории, остается открытым. Также открытым остается вопрос, можно ли получить аналогичную оценку с помощью стратегии, которая каждый раз в качестве гипотезы выдает программу с минимальным числом команд ("наиболее простую гипотезу").

В заключение отметим, что за пределами настоящего доклада остались более практические вопросы, касающиеся синтеза и проверки корректности программ по их работе на отдельных примерах ([9], [17], и др.).

Литература

1. R. J. Solomonoff, *A formal theory of inductive inference*, Information and Control **7** (1964), 1–22, 224–254.

2. Б.А. Трахтенброт и Я.М. Барздинь, *Конечные автоматы (поведение и синтез)*, "Наука", Москва, 1970.

3. E. F. Moore, *Gedanken-experiments on sequential machines*, Automata Studies, Princeton Univ. Press, 1956, pp. 129–153.

4. E. M Gold, *Limiting recursion*, J. Symbolic Logic **30** (1965), 28–48.

5. L. Blum and M. Blum, *Inductive inference: a recursion theoretic approach*, Memorandum No. ERL-M386, University of California, Berkeley, 1973.

6. J. M. Barzdin, *Prognostication of automata and functions*. Information Processing 71, North-Holland, Amsterdam, 1972, pp. 81–84.

7. Р.В. Фрейвалд, *О предельном синтезе номеров общерекурсивных функций в различных вычислимых нумерациях*, Доклады АН СССР **219** (1974), 812–814.

8. Я.М. Барздинь и Р.В. Фрейвалд, *О прогнозировании общерекурсивных функций*, Доклады АН СССР **206** (1972), 521–524.

9. A. W. Bierman, R. Baum, R. Krishnaswamy and E. Petry, *Automatic program synthesis*, Computer and Information Science Research Centre, Technical Report 73–6, Ohio State University, Columbus, Ohio, 1973.

10. К.М. Подниекс, *Сравнение различных типов предельного синтеза и прогнозирования функций*, Теория алгоритмов и программ, I, Учен. зап. Латв. гос. унив. **210** (1974), 68–81.

11. Я.М. Барздинь, *Две теоремы о предельном синтезе функций*, там же, 82–88.

12. Я.М. Барздинь и Р.В. Фрейвалд, *Прогнозирование и предельный синтез эффективно перечислимых классов функций*, там же, 110–116.

13. Я.М. Барздинь, Е.Б. Кинбер и К.М. Подниекс, *Об ускорении синтеза и прогнозирования функций*, там же, 117–128.

14. Я.М. Барздинь, *Предельный синтез τ-номеров*, там же, 112–116.

15. ———, *Прогнозирование и предельный синтез конечных автоматов*, там же, 129–144.

16. ———, *Замечание о синтезе программ по историям их работы*, там же, 145–151.

17. Я.М. Барздинь, Я.Я. Бичевский и А. А. Калниньш, *Построение полной системы примеров для проверки корректности программ*, там же, 152–187.

Латвийский Государственный Университет
Рига, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Spectral Functions of Graphs*

Alan J. Hoffman

1. Introduction. Let G be a graph, and let its adjacency matrix $A(G)$ be the symmetric (0, 1) matrix with 0 diagonal whose (i, j)th entry is 1 if and only if vertices i and j of G are adjacent. Denote the eigenvalues of $A(G)$ by $\lambda_1 \geqq \lambda_2 \geqq \cdots$ or $\lambda^1 \leqq \lambda^2 \leqq \cdots$. The careful examination of the eigenvalues of $A(G)$, and of the eigenvalues of some linear combinations of $A(G)$, J, and I, plays a central role in many combinatorial investigations (see the excellent survey in [7]). Despite these achievements there are very few functions of a graph whose magnitude, even roughly, we know to depend on the spectrum of its adjacency matrix. To focus attention on this question, we first define precisely what we mean by such a "spectral function" and survey briefly what spectral functions are now known.

An unbounded numerical function $f(G)$ on the set of all graphs is said to be spectral if there do not exist two sequences of graphs $G_1, G_2, \cdots$ and $G_1', G_2', \cdots$ such that (i) for all i, the spectrum of $A(G_i)$ is the same as the spectrum of $A(G_i')$, and (ii) $\{f(G_i)\}$ is bounded, $\{f(G_i')\}$ is unbounded. An auxiliary question, given a spectral function, is to determine the precise part of the spectrum on which it depends.

(1) The maximum degree $d(G)$ of the vertices of a graph is a spectral function, depending on λ_1. By the Perron-Frobenius theory of nonnegative matrices, it is easy to see that $d((G))^{1/2} \leqq \lambda_1 \leqq d$; indeed [5], if $\delta(G)$, the minimum degree of the vertices of G, exceeds 2, then $d(G)/(d(G) - 1)^{1/2} < \lambda_1$.

(2) Let $\gamma(G)$ be the smallest number of parts into which the set of edges of a graph G can be partitioned so that each part is a complete bipartite graph on the vertices touched. It can be shown [2] that $\gamma(G)$ is a spectral function, depending on the number of eigenvalues each of which is at most -1. (On the other hand, if each

*This work was supported (in part) by the U. S. Army under contract DAHC04–72-C-0023.

part is a clique, or if each part is a complete multipartite graph, the corresponding function is not spectral.)

(3) Call two vertices of G equivalent if every other vertex is either adjacent to both or adjacent to neither. Then the number of equivalence classes is a spectral function [3], and depends on the number of eigenvalues each of which is at most $(-1-5^{1/2})/2$ plus the number of eigenvalues each of which is at least 1.

One way of discovering spectral functions, depending just on one particular eigenvalue, is to look for nested families of graphs such that no large representative of any family is an induced subgraph of a graph with that eigenvalue of modest size. For example, let $m_1(G)$ be the smallest integer m such that neither K_m nor $K_{1,m}$ is an induced subgraph of G. Then (1) above implies, by Ramsey's theorem, that $m_1(G)$ is a spectral function of G, depending on λ_1.

(4) Let H_n be the graph formed by adjoining to K_{2n} one additional vertex adjacent to exactly n vertices of K_{2n}. Let $m^1(G)$ be the smallest integer m such that neither $K_{1,m}$ nor H_m is an induced subgraph of G. Then [1] $m^1(G)$ is a spectral function of G, depending on λ^1. A similar result [6] holds yielding a spectral function $m_2(G)$ depending on λ_2.

These suggest the following

CONJECTURE. For each i, there exists a finite set of sequences of graphs $\{G_n^j\}$, $n = 1, 2, \cdots$, j, belonging to a finite set of indices, such that (i) G_n^j is an induced subgraph of G_{n+1}^j for all j, (ii) if $m_i(G)$ is the smallest m such that, for all j, G_m^j is not an induced subgraph of G, then $m_i(G)$ is a spectral function, depending on λ_i.

(Of course, one should make the corresponding conjecture for λ^i.) At the moment, we seem to have the tools for exploring this conjecture for small i, which involves discovering the specific families $\{G_n^j\}$, but not for considering the general case.

Results analogous to (4) can be established for the eigenvalues of the matrix $J - I - 2A(G)$, and have various interesting interpretations and corollaries.

(5) Let G and H be graphs, on the same set of vertices, and let $d(G, H)$ be the maximum valence in the graph whose set of edges is $(E(G) - E(H)) \cup (E(H) - E(G))$. Let $d_B(G)$ be minimum of $d(G, H)$ where H ranges over all complete bipartite graphs on all vertices. Then [4] $d_B(G)$ is a spectral function (of the eigenvalues of $J - I - 2A(G)$), depending on $\lambda^1(J - I - 2A(G))$.

COROLLARY. *If R is any symmetric matrix every entry of which is ± 1, then every eigenvalue of R other than the largest is bounded by a function just of $\lambda^1(R)$.*

(6) Results on the second eigenvalue of symmetric matrices, every entry of which is ± 1, have recently been proved by J. J. Seidel and the author. Together with (5), they establish that the following is true for $i = 1$ and 2:

CONJECTURE. For each i, there exists a function f_i such that, if R is any symmetric matrix every entry of which is ± 1, then

$$\min \lambda_i(DRD + J) \leqq f_i(\lambda_i(R)), \qquad D \text{ diagonal}, d_{ii} = \pm 1.$$

References

1. A. J. Hoffman, *On spectrally bounded graphs*, A Survey of Combinatorial Theory, N. Srivastava, et al., editors, North-Holland, Amsterdam, 1973, pp. 277–283.

2. ———, *Eigenvalues and partitionings of the edges of a graph*, Linear Algebra and Appl. **5** (1972), 137–146. MR **46** #97.

3. ———, *Applications of Ramsey-style theorems to the eigenvalues of a graph*, Combinatorics, Math. Centre Tracts, 57, Amsterdam, 1974, pp. 43–57.

4. ———, *On eigenvalues of symmetric* $(+1, -1)$ *matrices*, Israel J. Math. **17** (1974), 69–75.

5. A. J. Hoffman and J. H. Smith, *On the spectral radii of topologically equivalent graphs*, Proc. Second Czechoslovakian Sympos. on Graph Theory (to appear).

6. Leonard Howes, *On subdominantly bounded graphs*, J. Combinatorial Theory (to appear).

7. J. J. Seidel, *A survey of two-graphs*, Proc. Internat. Colloq. Theoric Combinatoric, Accad. Naz. Lincei, Roma, 1973 (to appear).

IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Extremal Properties on Partial Orders

D. J. Kleitman

This paper contains a brief description of recent developments in three areas related to extremal properties of subsets of finite partially ordered sets. The three areas concern generalizations of Dilworth's theorem; a study of partial orders obeying inequalities analogous to that of Lubell, Meshalkin and Yamamoto for the lattice of subsets of a finite set; and extensions of the Littlewood-Offord problem on the distribution of linear combinations of vectors. Fuller descriptions of the first two problems appear in the Proceedings of the Nyenrode Conference (July 1974).

1. Generalizations of Dilworth's theorem. A finite partially ordered set is a finite set on which a binary "order" relation (one that is reflexive, asymmetric and transitive) is defined. A "chain" is a subset of such an order every pair of elements of which is ordered. An antichain is a subset of an order no two distinct elements of which are ordered. A well-known theorem of Dilworth states that the maximal size of an antichain in a partially ordered set S is equal to the minimal number of blocks in a partition of S into chains. If A is the set of all antichains of S, C the set of all chains of S, P and Q the sets of all partitions of S into chains and antichains respectively, this theorem can be written as

$$f_1 \equiv \operatorname*{Max}_{B\in A} \sum_{G\in B} 1 = \operatorname*{Min}_{D\in P} \sum_{H\in D} 1.$$

The dual relation (interchanging chain and antichain) also holds:

$$f_1^* \equiv \operatorname*{Max}_{B\in C} \sum_{G\in B} 1 = \operatorname*{Min}_{D\in Q} \sum_{H\in D} 1.$$

We define a k-family to be the union of k antichains, and a k^*-family to be the union of k chains. Let Q_k be the set of k-families and P_k the set of k^*-families.

C. Greene and the present author considered the question: What can be said about the maximum size of a k- (or k^*-) family?

Since each chain can intersect each antichain at most once, the right-hand sides in the relations above must obviously be greater than the left-hand sides. The analogous obvious inequalities for k-families can also be proven to be equalities. The results obtained by C. G. and D. J. K. are

$$(1) \qquad f_k \equiv \operatorname*{Max}_{B \in Q_k} \left(\sum_{G \in B} 1 \right) = \operatorname*{Min}_{D \in P} \left(\sum_{H \in D} \operatorname{Min}(k, |H|) \right)$$

where $|H|$ is the number of elements of H.

(2) If a k-saturated partition D into chains is one for which $f_k = \sum_{H \in D} \operatorname{Min}(k, |H|)$ then there always exist simultaneously k and $k+1$ saturated partitions. (Similar results for k and $k+j$ fail for $j \neq 1$.)

Greene has obtained the following additional results: The same results hold with chain and antichain reversed. Moreover, if one constructs an array in which the length of the kth row is $f_k - f_{k-1}$, and rows are left justified, then the length of the kth column is $f_k^* - f_{k-1}^*$, with f_k^* the maximal k^* family size.

A. J. Hoffman has noted that these relations may all be proven using a formulation of the properties of partial orders in terms of linear inequalities using duality in linear programming, in particular an extension of the formulation used by Dantzig and himself to prove Dilworth's theorem.

2. The LYM property. The Boolean algebra of subsets of a finite set S forms a commonly encountered example of a partial order. For this example there is a "rank function" defined on the order; each element is a subset of T and has rank given by the number of elements it contains. For our purposes a rank function is an order preserving function defined on the elements of the order with values in Z. Let the number of elements of rank k be N_k.

For the subset order, Sperner, in 1928, proved that $f_1 = \operatorname{Max}_k N_k$. This result has been proven for the divisors of an integer under "divides" as order relation, for the subspaces of a finite vector space over a finite field; it may or may not hold for partitions of a finite set or finite integer with refinement as order relation.

Lubell, Meshalkin, and Yamamoto (and apparently others as well) independently noticed that a stronger inequality holds. If x is any k-family and x_j is the number of its elements of rank j, then

$$(3) \qquad \sum (x_j / N_j) \leqq k.$$

(The proof is easy—each summand on the left measures the proportion of all possible maximal chains met by members of x of rank j; the inequality can be interpreted as saying that a k-family meets a chain at most k times.) This relation implies Sperner's inequality as well as many other results.

The question considered in the work announced here is: For what partial orders does the same result hold?

The central result is the following characterization: The LYM inequality (3) will hold on a partial order if and only if the order S obeys the "normalized matching condition".

This condition states that for each k there exists a 1-1 mapping from N_{k+1} copies of the kth rank of S to N_k copies of the $(k+1)$st rank of S such that each element is ordered with respect to its image.

As it is known that the divisors of an integer under divisibility as order relation satisfies normalized matching, this order must therefore satisfy the LYM property, a result previously obtained by I. Anderson.

The order of partitions of n fail to satisfy the LYM property, as the lattice of partitions of a finite set (noted by J. Spencer).

A number of consequences and implications of the LYM property can be described. The reader is referred to the reference for details.

3. The Littlewood-Offord problem. Given n vectors, each of unit magnitude or more, consider all linear combinations of them with coefficient 0 or 1. (Other coefficient sets may be considered with conclusions to most of the results described below.) Erdös, generalizing some questions of Littlewood and Offord, raised the question, how many of these linear combinations can lie in the interior of the union of k unit diameter regions? Katona introduced the question, if one deals with α diameter regions, what analogous conclusions follow?

Erdös completely solved the problem for vectors in one dimension by noticing that for that case the linear combinations lying in the union of k unit diameter regions must be a k-family in the sense of the first section above.

Katona and myself independently found a generalization of the Sperner property that solved the problem for a single unit diameter region for vectors in two dimensions. Katona obtained a solution in the same case for a $\sqrt{2}$ diameter region. A somewhat different approach of Kleitman yields the same diameter 1 and diameter $\sqrt{2}$ upper bounds (namely $\binom{n}{[n/2]}$ and $\binom{n+1}{[(n+1)/2]}$) in arbitrary dimensional space.

The k-family result in one dimension implies that with suitable definition of rank, a LYM inequality can be obtained restricting the number of linear combinations for this case.

It is the aim of the present research to examine the questions

1. What larger diameter results exist?
2. What strengthening of the bounds may be obtained?
3. What analogues of the LYM property may be obtained in higher dimensions than one?

A number of results have been found in all three of these directions. A lengthy paper describing all these will appear in Advances in Mathematics.

In this article we confine ourselves to stating some results and open questions of the first kind, and to proving one simple result that to some extent is a mixture of the two latter categories. A number of other results are described in the reference.

The best bounds have been obtained for vectors in two dimensions with diameters $\sqrt{3}$ and $\sqrt{5}$. They are

$$2\binom{n}{[n/2]}, \quad \text{and} \quad \binom{n}{[n/2]} + \binom{n}{[n/2]+1} + \binom{n}{[n/2]-1}$$

for n even, respectively. These bounds are conjectured for all dimensions, the latter

for even n, and for odd $n \geqq 5$. It is not difficult to conjecture upper bounds in terms of geometrically simple structures for any sufficiently small diameter compared to n. We prove the following result.

THEOREM. *Suppose two of the vectors have angle between them having cosine* $\leqq \frac{1}{2}$ *in magnitude. Then the number of linear combinations in a unit diameter region cannot exceed* z;

$$z = \binom{n+1}{[(n+1)/2]} - 2\binom{n-1}{[(n-1)/2]}.$$

(*This is a tighter bound than* $\binom{n}{[n/2]}$.) *This bound is not best possible.*

If all pairs of vectors have $|\cos\theta| > \frac{1}{2}$ for θ the angle between them, the problem reduces to the one-dimensional case and LYM inequalities can be obtained for it.

We conclude with a proof of this theorem. It is extremely simple—unfortunately simpler than most results on this area.

PROOF. We obtain our bound by partitioning the set L of all linear combinations into blocks such that no two in a block are within a distance one of each other. Since the linear combinations in the interior of a diameter one region can have at most one l.c. per block, if the number of blocks is z our conclusion follows. Now we can partition L into $\binom{n}{[n/2]}$ such blocks by imitating a standard partition of the subset of a set into that number of chains.

Such a partition for subsets can be constructed inductively as follows. Consider each block B in a partition of the subsets of $(1 \cdots k)$ into chains. If a new element $k+1$ is introduced one can define two chains of subsets of $(1, \cdots, k+1)$ corresponding to B—B itself and the addition of $(k+1)$ to each member of B. One obtains the desired partition by removing the maximal member of the latter and adding it to B. It is easily verified that this partition is j-saturated for every j.

Thus for linear combinations, if for each k we can likewise transfer one member of one of the "daughter" blocks of a block B on a partition of l.c.'s up to vector k to the other, we obtain a partition for $k+1$ vectors having the same number of blocks of each size as the chain partition just described. We can always do this, since if the vector a_{k+1} is introduced we can transfer the linear combination in B having minimum component in the direction of a_{k+1} from B to the translate of B by a_{k+1}. We thereby obtain two daughter blocks from B obeying the distance requirements on blocks and having sizes one less than and one greater than B.

This argument yields the $\binom{n}{[n/2]}$ bound in any dimension. To obtain the improved bound of the theorem consider the two vectors having cosine of their included angle $\leqq \frac{1}{2}$ in magnitude. All four of their linear combinations have distance $\geqq 1$ from one another, so that for $k = 2$ they form one 4-element block instead of a 3- and a 1-element block as in the chain case. Obtaining daughters from one 4-element block for $k = 2$ by induction as indicated yields a partition into

$$\binom{n+1}{[(n+1)/2]} - 2\binom{n-1}{[(n-1)/2]}.$$

blocks as can easily be shown.

References

1. C. Greene, *Sperner families and partitions of a partially ordered set*, Proc. Nyenrode Conf. 1974, vol. 2, Math. Centrum Tract 86, Amsterdam, 1974, pp. 91–106.

2. D. J. Kleitman, *On an extremal property of antichains in partial orders*, Proc. Nyenrode Conf. 1974, vol. 2, Amsterdam, 1974, pp. 77–90.

3. ———, *Some new results on the Littlewood-Offord problem*, J. Combinatorial Theory (1975) (to appear).

Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

On the Theory of Inference Operators*

Rolf Lindner

It is well known that the theory of inductive inference contributes to a theory of interpolation in the discrete case: One investigates algorithms, which, from given values of a function, compute new ones (see **[Go]**, **[Ba]**, **[Fe]**, **[Thi]**, **[Blu]** et al.). In this discrete interpolation theory, such algorithms assign, to certain local informations, constructive, global informations, i.e., one realizes mappings of the form $\Phi: \mathfrak{P}N^2 \to N$; the values in range Φ (hypotheses) are interpreted in a suitable effective numbering $\beta: N \to Pa^1$, where Pa^1 denotes the set of partially recursive functions. Detailed studies of such inference operators are to be found, e.g., in **[KlLi]**.

At the same time, it is well known that this inductive recognition is a very special case of the one used in mathematical statistics. However, on the one hand, in mathematical statistics one finds hardly any constructive elements (global information, i.e., inductively determined laws mean here parameters of distributions), on the other hand, inductive inference restricts itself a priori to recursive objects (laws mean here computing procedures, Gödel numbers etc.).

Discrete mathematics, however, has already had for quite a while a much more far reaching concept of law, namely that of a constructively described null set.

In this paper, we want to discuss some aspects of a theory of inference operators, in which the used hypotheses are constructively described null sets. For the sake of simplicity, we explain our ideas using the example of the set of indicator functions on N, which we shall identify with the set X^ω ($X = \{0, 1\}$) of infinite sequences with values in X; we assume further $\mu(0) = \mu(1) = \frac{1}{2}$ and denote the product measure on X^ω by $\bar{\mu}$.

Following **[Ma]**, **[Ja]**, **[Sch]**, we call a set $\mathfrak{N} \subseteq X^\omega$ a recursive null set, if there

*Translated from German at the author's request.

exists a recursive sequential test for $\mathfrak{N}$, i.e., a Y with the following properties:

(R1) $Y \subseteq \boldsymbol{N} \times X^*$ and Y is enumerable.

(R2) If we let $[Y_i] =_{\text{def}} \{p: [i, p] \in Y\} \cdot X^\omega$, then $\bar{\mu}([Y_i]) \leqq 2^{-i}$.

(R3) If we let $\mathfrak{N}_Y =_{\text{def}} \bigcap_i [Y_i]$, then $\mathfrak{N}_Y = \mathfrak{N}$.

Moreover, following [**Sch**], we shall call $\mathfrak{N}$ a totally recursive null set, if there exists a recursive sequential test Y for $\mathfrak{N}$, such that the following holds:

(T4) The function f with $f(i) =_{\text{def}} \bar{\mu}([Y_i])$ is computable in the sense of recursive analysis.

What are the general implications, if one pursues inductive recognition by recursive null sets, i.e., if one considers inference operators $\Phi: X^\omega \to \boldsymbol{N}$, for which $h \in \text{range } \Phi$ is interpreted by an effective numbering β of recursive sequential tests (null sets)?

For this, we let $AKZ(\Phi, \beta) =_{\text{def}} \{\xi: \xi \in \text{domain } \Phi \wedge \xi \in \beta \circ \Phi(\xi)\}$ ("$\in$" is here shorthand notation for: ξ is an element of the null set, described by the sequential test $\beta \circ \Phi(\xi)$).

(I) The results of the theory of inductive recognition carry over. (According to [**Sch**], given any recursive $\xi \in X^\omega$, there is a totally recursive sequential test Y with $\mathfrak{N}_Y = \{\xi\}$.)

(II) The types of realizations of Φ are very similar to those discussed in inductive inference. (Identification by numbering, limit recursion, etc.; see [**So**], [**Go**].)

(III) The weaker concept of law allows also the nontrivial recognition of certain nonrandom and nonrecursive objects.

(IV) The fusion of statistics and constructive mathematics, seen as a goal by [**Sch**], is completed.

In order to illustrate (I)–(IV) we formulate here some statements about a certain type of recognition, using regular null sets. As in [**McN**], we call $M \subseteq X^\omega$ regular, if there exist regular W_i, $V_i \subseteq X^*$ $(1 \leqq i \leqq n)$ such that $M = \bigcup_{i=1}^n W_i V_i^\omega$. Since recursive null sets are, topologically speaking, G_δ sets (i.e., countable intersections of open sets), we shall consider from now on regular G_δ sets. From [**StWa**] we know

LEMMA 1. *The following statements are equivalent:*

(a) *M is a regular G_δ set;*

(b) *there are regular W_i, $V_i \subseteq X^*$ $(1 \leqq i \leqq n)$, which are totally disordered with respect to the initial word relation such that $M = \bigcup_{i=1}^n W_i V_i^\omega$;*

(c) *there is a finitely determined acceptor $\mathfrak{A} = [Z, X, f, z_0]$ and a $\mathfrak{Z} \subset \mathfrak{P}Z$, such that*

$$M = \{\xi: \exists Z' (Z' \in \mathfrak{Z} \wedge U(\Phi_{\mathfrak{A}}(\xi)) \cap Z' \neq \emptyset)\}$$

where $\Phi_{\mathfrak{A}}(\xi) =_{\text{def}} f(z_0, \xi(1)) \cdot \cdots \cdot f(z_0, \xi(1) \cdots \xi(n))$ and $U(\zeta) = \{z: \text{card}\{n: \zeta(n) = z\} = \aleph_0\}$.

This immediately implies

THEOREM 2. *Every regular G_δ null set is totally recursive.*

Let R be the family of all regular G_δ null sets and $\mathfrak{N}_R =_{\text{def}} \bigcup R$; let Δ be an effec-

tive numbering of all recursive null sets (which exists according to [**Ma**], [**Sch**]). For any generally recursive function F: $X^* \to \boldsymbol{N}$, we shall consider the set

$$(*) \qquad \Delta N(F) \underset{\text{def}}{=} \left\{\xi : s = \lim_{p \to \xi} F(p) \wedge \Delta(s) \in R \wedge \xi \in \mathfrak{C}(\Delta(s))\right\}$$

of all "laws", recognized by F, by a limit procedure. (Here $\mathfrak{C}(M)$ denotes the closure of M.) This definition was made, following the method "GN" in [**Ba**], for example.

Definition $(*)$ does not force us to leave the regular G_δ null sets because of

LEMMA 3 (STAIGER). *$\mathfrak{C}(M) \in R$ for $M \in R$.*

Further let $\Delta N =_{\text{def}} \{\Delta N(F) : F \text{ generally recursive}\}$ and $\mathfrak{N}_{\Delta N} =_{\text{def}} \bigcup \Delta N$.

COROLLARY 4. $\mathfrak{N}_R = \mathfrak{N}_{\Delta N}$.

Through identification by enumeration one has

THEOREM 5. *Given any effective β: $\boldsymbol{N} \to R$, one can find effectively a generally recursive F such that $\mathfrak{N}_\beta \subseteq \Delta N(F)$ for $\mathfrak{N}_\beta =_{\text{def}} \bigcup_n \beta_n$.*

In the proof one uses the fact that the property $p \cdot X^\omega \cap M \neq \emptyset$, for $p \in X^*$ and regular $M \subseteq X^\omega$ can be decided. This, too, explains why we restrict ourselves here to regular null sets.

Let $\mathfrak{N}_T$ be the union of all totally recursive null sets. A comparison with $\mathfrak{N}_{\Delta N}$ gives

THEOREM 6. $\mathfrak{N}_T \supset \mathfrak{N}_{\Delta N}$.

One can prove the stronger result: If $\xi = 01 \cdots 0^n 1^n \cdots$ is an element of a regular set $M \subseteq X^\omega$, then $\bar{\mu}(M) > 0$.

Because of Corollary 4 we have, therefore, for the set $Al \subseteq X^\omega$ of generally recursive sequences

THEOREM 7. *$Al \backslash \mathfrak{N}_{\Delta N} \neq \emptyset$, since it is well known that $Al \subset \mathfrak{N}_T$ (see* [**Sch**]).

Thus, on the one hand, one has recognitions of type ΔN for nonrecursive, nonrandom sequences; on the other hand, not all recursive sequences are ΔN-recognizable.

We ask ourselves now whether there exists a universal ΔN-recognition for $\mathfrak{N}_R$. To this effect, we prove

THEOREM 8. *The property of a regular G_δ set to be a null set is decidable.*

By Lemma 1, it suffices to decide the property $\bar{\mu}([V]) < 1$ for $V \subseteq X^*$ regular and totally disordered.

To this end we use

LEMMA 9 (PAZ, WECHSUNG). *If $V \subseteq X^*$ is totally disordered and if V is accepted by the finite determined automaton $\mathfrak{A} = [X, Z, f, z_0, F]$ then one can effectively find polynomials g and h of degree* card Z *with rational coefficients, such that $\bar{\mu}([V]) = g(2^{-1}) \cdot (h(2^{-1}))^{-1}$.*

Thus one has

COROLLARY 10. *One can construct an effective* $\beta : \boldsymbol{N} \to R$ *such that* $\beta(\boldsymbol{N}) = R$.

Hence we conclude from Theorem 5 and Lemma 3

THEOREM 11. *One can construct a universal generally recursive* F: $X^* \to \boldsymbol{N}$ *such that* $\Delta N(F) = \mathfrak{N}_R$.

At the same time one has

THEOREM 12. *One can construct a totally recursive sequential test* Y *such that* $\mathfrak{N}_R \subseteq \mathfrak{N}_Y$.

For the proof one uses the following lemma, which goes back to [Sch], because of Theorem 2.

LEMMA 13. *Given any effective* $\beta : \boldsymbol{N} \to R$, *one can find a totally recursive sequential test* Y *such that* $\mathfrak{N}_\beta \subseteq \mathfrak{N}_Y$.

This is another confirmation of Theorem 6, if one uses

THEOREM 14 (SCHNORR). *Given any totally recursive test* Y, *one can effectively construct a* $\xi \in Al \backslash \mathfrak{N}_Y$.

Since there exists an effective numbering of all acceptors of regular G_δ null sets such that the state numbers of the numbered automata increase, Theorem 8 implies the following sharpening of Corollary 10.

THEOREM 15. *One can give an effective Occam numbering* $\hat{\beta} : \boldsymbol{N} \to R$ *for* R.

For the strategy F: $X^* \to \boldsymbol{N}$, constructed in Theorem 11, we even get now that $\Delta \circ \Phi(\xi)$ is a weight-minimal regular G_δ null set for every $\xi \in \mathfrak{N}_R$. By this, complexity in the discovered laws becomes accessible to investigation.

The theorems formulated here for the case of ΔN-realizations of inference operators confirm sufficiently (I)—(IV) and suggest how to continue the investigations for other types of realizations (GH', GN^∞, etc.; see [Ba]) and for other classes of null sets constructively described. This will be done in a subsequent paper which is being prepared.

References

1 [Ba]. Ja. M. Barzdin' and K. M. Podnieks, *On the theory of inductive inference*, Mathematical Foundations of Computer Science, High Tatras, Czechoslovakia, 1974.

2 [Blu]. L. Blum and M. Blum, *Inductive inference: A recursion theoretic approach*, Memorandum ERL-M 386, University of California, Berkeley, Calif., March 1973.

3 [Fe]. J. Feldman, *Some decidability results on grammatical inference and complexity*, Information and Control **20** (1972).

4 [Go]. E. M. Gold, *Language identification in the limit*, Information and Control **10** (1967).

5 [Ja]. K. C. Jacobs, *Turing-Maschinen und zufällige* 0-1-*Folgen*, Selecta Mathematica, II, Springer, Berlin, 1970, pp. 141–167. MR **44** #6413.

6 [KlLi]. R. Klette and R. Lindner, *Recognition of operators and essential sets*, Sb. Dishret. Mat., Vyčisl. Centr Akad. Nauk SSSR, Moscow (in press). (Russian)

7 [Ma]. P. Martin-Löf, *The definition of random sequences*, Information and Control **9** (1966), 602–617. MR **16** #6228.

8 [McN]. R. McNaughton, *Testing and generating infinite sequences by a finite automaton*, Information and Control **9** (1966), 521–530. MR **35** #4105.

9 [Sch]. C. P. Schnorr, *Zufälligkeit und Wahrscheinlichkeit*, Lecture Notes in Math., vol. 218, Springer-Verlag, Berlin and New York, 1971.

10 [So]. R. Solomonoff, *A formal theory of inductive inference.* I, Information and Control **7** (1964), 1–22. MR **30** #2963.

11 [StWa]. L. Staiger and K. Wagner, *Automatentheoretische und automatenfreie Charakterisierung topologischer Klassen regulärer Folgenmengen*, EIK (Elektronische Informationsverarbeitung und Kybernetik) **10** (1974), no. 7, Akademie-Verlag, Berlin.

12 [Th]. H. Thiele, *Lernverfahren zur Erkennung formaler Sprachen*, in Lernende Systeme, VEB Deutscher Verlag der Wissenschaften, Berlin, 1973.

13 L. Staiger and G. Wechsung, Sektion Mathematik der Friedrich-Schiller-Universität Jena (persönliche Informationen).

UNIVERSITY OF JENA
JENA, GERMAN DEMOCRATIC REPUBLIC

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Inherent Computational Complexity of Theories of Ordered Sets*

Albert R. Meyer

The significance of the theoretical distinctions between problems which are effectively decidable and those which are not can be challenged by objections of at least two kinds:

(1) Only a *finite* collection of sentences about arithmetic, for example, are of human concern, so the undecidability of the infinite collection of true sentences of arithmetic is immaterial.

(2) An *efficient* decision procedure for the monadic predicate calculus, for example, would have important practical applications, but the mere fact that it is effectively decidable is immaterial.

We first consider the second objection. Clearly more time and effort are required to prove the truth or validity of long sentences simply because it may require a long time to read them. The "efficiency" of a decision procedure must thus be measured relative to the size of the sentences to which the procedure is applied. A hint that many decision problems in logic cannot have efficient decision procedures follows from the observation that short sentences can define relatively large sets.

For example, consider the pure predicate calculus with monadic (i.e., one argument) predicate letters $P_1, P_2, \cdots, P_n$. If we choose some interpretation of $P_1, P_2, \cdots, P_n$ as predicates on some domain, then any element x in the domain can be identified with the vector of truth-values $\langle P_1(x), P_2(x), \cdots, P_n(x)\rangle$. Then the formula

$$D_i(x, y) := (P_i(x) \equiv \neg P_i(y)) \wedge \bigwedge_{j \neq i} (P_j(x) \equiv P_j(y))$$

means that x and y differ precisely in their ith component. Hence, the sentence

*This research was supported by the National Science Foundation under research grant GJ-43634X.

$$S := \forall x \left(\bigwedge_{i=1}^{n} \exists y D_i(x, y) \right)$$

means that for every vector x and every component of x, there is another vector y which differs from x precisely at that component. Clearly S is satisfied only by interpretations in which each of the 2^n vectors occurs in the domain. Notice, however, that S, even with the abbreviated conjunctions fully expanded, contains only proportional to n^2 connectives and is of length proportional to $n^2 \cdot \log n$. (By the length of S we mean the number of occurrences of all symbols including connectives, variables, predicate letters and parentheses. The factor $\log n$ appears because the alphabet of symbols is assumed to be fixed and $n \cdot \log n$ subscript digits are then required to represent the n distinct predicate letters $P_1, \cdots, P_n$.)

The following theorem implies that any decision procedure for satisfiability of sentences of monadic predicate calculus with n predicate letters requires essentially the same effort as exhaustively testing all possible interpretations on domains of size up to 2^n. (The usual proofs of the decidability of monadic predicate calculus imply that testing domains up to this size is sufficient.)

THEOREM (MEYER [**MR75**]). *Any Turing machine which, given any sentence of monadic predicate calculus, decides whether the sentence is satisfiable, requires a number of steps exceeding* $2^{\varepsilon \cdot \text{length}(S)/\log(\text{length}(S))}$ *for some* $\varepsilon > 0$ *and infinitely many sentences* S.

It should be apparent that if a Turing machine requires an exponentially growing number of steps, then so will any other reasonable model of a computer. Moreover, the lower bound of the theorem applies even to nondeterministic Turing machines, which implies that the shortest *proofs* of satisfiability or validity of such sentences will also be exponential (cf. [**FR74**] for further discussion of proof-length).

The preceding theorem and others to follow can in retrospect be seen as a natural extension of Gödel's first incompleteness theorem and Turing's and Church's proofs of undecidability: One "arithmetizes" or codes the computations of Turing machines into a domain and constructs sentences which assert that the coded computation halts in an accepting state. The technical flavor of the proofs differs from undecidability proofs in that emphasis rests on efficiently (relative to the size of the Turing machine) constructing short sentences which describe computations which eventually halt after a long time. We shall not attempt to describe the proofs further.

The first proof that a decidable theory, namely the weak monadic second-order theory of the successor function on the nonnegative integers (WS1S), was discovered in May, 1972 [**Me72**], [**Me73**]. Since then reasonably close upper and lower bounds on the inherent computational complexity of most of the classical examples of decidable theories have been obtained.

THEOREM. *Any Turing machine which decides membership in the set* $\mathscr{A}$ *requires a number of steps exceeding*

$$2^{2^{\cdot^{\cdot^{2^{\text{length}(S)}}}}} \varepsilon \cdot \log (\text{length}(S))$$

for some $\varepsilon > 0$ and infinitely many $S \in \mathscr{A}$, where $\mathscr{A}$ is any of the following:

1. *WS1S* (*Meyer* [**Me73**]),

2. *star-free expressions (from automata theory) for the empty set* (*Stockmeyer* [**St74**], [**SM75**]),

3. *the theory of linear orders* (*Meyer* [**St74**], [**SM75**]),

4. *the theory of any nonempty family of infinite linear orders with a single monadic predicate* (*Stockmeyer* [**St74**], [**SM75**]),

5. *the theory of two successors and prefix* (*Meyer and Stockmeyer* [**St74**], [**SM75**]),

6. *the theory of a single unary function* (*M. Fischer and Meyer* [**FM75**]),

7. *the theory of pure finite types* (*M. Fischer and Meyer* [**FM75**]),

8. *the theory of addition on the nonnegative integers with the predicate "x is a power of 2 and x divides y"* (*Meyer* [**Me73**]),

9. *the theory of any nonempty family of pairing functions* (*Rackoff* [**Rac74b**]).

For each of these examples, decision procedures are known which require at most

$$2^{2^{\cdot^{\cdot^{2}}}} \} n$$

Turing machine steps on inputs of length n. It is a curious empirical observation that all natural decision problems known to be decidable require at most this many steps. (Of course it is not hard to contrive examples of decidable theories which are not even primitive recursive (cf. [**Rac74b**]).)

An exponential lower bound for the computational complexity of the theory of essentially any algebraic structure follows from the following theorem. A family of semigroups is of *unbounded order* if for every $k > 0$ there is a semigroup in the family and an element s in the semigroup such that $s^i \neq s^j$ for all $1 \leqq i < j \leqq k$.

THEOREM (M. FISCHER [**FR74**]). *The first order theory of any family of semigroups of unbounded order requires time* $2^{\varepsilon \cdot \text{length}(S)}$.

An immediate corollary is that exponentially many steps are required to decide sentences in the theory of the real numbers under addition, and, a fortiori, efficient implementations of Tarski's celebrated decision procedure for the real field do not exist. Decision procedures for sentences of length n in the first order theory of the real field which require at most $2^{2^{kn}}$ steps for some constant k have recently been announced by Collins [**Co174**] and independently by Monk [**Mo74**]and Solovay [**So74**].

The decision problem for Presburger's arithmetic (i.e., the first order theory of addition of integers) which admits a very simple proof of decidability compared to real closed fields is computationally more difficult.

THEOREM (M. FISCHER-RABIN [**FR74**]). *Presburger's arithmetic requires time* $2^{2^{\varepsilon \cdot \text{length}(S)}}$

THEOREM (OPPEN [**Op73**], FERRANTE-RACKOFF [**FeRa75**], [**Rac74**]). *Moreover time* $2^{2^{\varepsilon \cdot \text{length}(S)}}$ *is sufficient*.

Fischer-Rabin [**FR74**] also have shown that *three* exponentials of steps are re-

quired to decide the theory of multiplication of positive integers, and Rackoff [**Rac74a, b**] has developed a general theorem relating the complexity of theories of structures to theories of powers of structures (the positive integers under multiplication being the weak direct power of the nonnegative integers under addition) which yields an upper bound of *four* exponentials for the theory of integer multiplication.

We note that since the lower bounds apply to nondeterministic as well as deterministic Turing machines, while the upper bounds are always deterministic, upper and lower bounds which differ by only one exponential are well matched (cf. [**St74**] for further explanation of this remark).

Apparently similar theories may have quite different complexities.

THEOREM (FERRANTE [**Fe74**]). *The first order theory of $\mathscr{B}$ requires for sentences of length n time $L(n)$ and can be decided in time $U(n)$ where*

$\mathscr{B} =$	$L =$	$U =$
1. *Integers with successor*	?	$2^{\varepsilon \cdot n}$
2. *Integers with order*	?	$2^{\varepsilon \cdot n}$
3. *Integers with successor and a single monadic predicate*	$2^{2^{\varepsilon \cdot n}}$	$2^{2^{2^{\varepsilon \cdot n}}}$
4. *Integers with order and a single monadic predicate*	$2^{2^{\cdot^{\cdot^{2}}}\}\varepsilon \cdot n}$	$2^{2^{\cdot^{\cdot^{2}}}\}n}$

(The lower bound in the last line follows from the result of Stockmeyer cited earlier, and the upper bound is implicit in the decision procedures of Büchi-Elgot [**Bu60**], [**El61**] and Rabin [**Rab69**].)

Thus the second objection raised in the first paragraph seems cogent. Mere decidability of a problem cannot be taken even to suggest that the problem admits feasible, practical decision procedures. Indeed nearly all the known decidability results of logic are inherently impractical in that exponential or more time is required by any possible decision procedure. (A notable exception is the decision problem for the propositional calculus. No subexponential time procedure is known, but neither have any nontrivial lower bounds been proved. We regard the determination of the computational complexity of the decision problem for the propositional calculus as the most important open problem in the theory of computation. Cook [**Coo71**] and Karp [**Ka72**] show that dozens of classical problems of combinatorial optimization are computationally equivalent to the decision problem for the propositional calculus.)

To the first objection, however, the proofs of the above theorems as well as classical undecidability theorems provide an implicit answer; the theorems about infinite problems often contain information from which one can estimate the difficulty of finite problems.

Consider Boolean functions of n variables and programs (or logical networks) which compute them by successively applying binary Boolean operations to the variables and to previously computed results. For example, the sequence of operations

$$a_1 := x_1 \wedge x_2, \qquad a_4 := a_1 \vee a_2,$$
$$a_2 := x_1 \wedge x_3, \qquad a_5 := a_4 \vee a_3,$$
$$a_3 := x_2 \wedge x_3, \qquad a_6 := \neg(a_1 \wedge x_3),$$
$$a_7 := a_6 \wedge a_5,$$

comprises a program for the function

$$f(x_1, x_2, x_3) = [(x_1 \wedge x_2) \vee (x_1 \wedge x_3) \vee (x_2 \wedge x_3)] \wedge \neg(x_1 \wedge x_2 \wedge x_3).$$

Thus $f(x_1, x_2, x_3) = 1$ if and only if exactly two of the zero-one valued variables x_1, x_2, x_3 have the value one.

We choose sentences of length n in some formal theory and code the symbols in these sentences as binary sequences. In the particular example below, six binary digits are sufficient to code all the symbols required, so sentences of length n correspond to binary sequences of length $6n$. We then inquire about the minimum number of binary operations required in a network computing the function of $6n$ variables which equals one if and only if its $6n$ inputs are the code of a well-formed true sentence.

The previous proofs show that short sentences can describe large networks as well as large Turing machine computations, and from this one can deduce that the size of networks which decide sentences of length n must grow exponentially with n. (This was first proved in 1967 by Ehrenfeucht **[Eh72]** for the sentences of "bounded arithmetic" with explicit use of constants in exponential notation, e.g., $3^{29'}$, allowed in the sentences.) In particular,

THEOREM (STOCKMEYER-MEYER **[St74]**, **[SM75]**). *If we choose sentences of length* 616 *in the decidable theory of WS1S and code these sentences into* $6 \times 616 = 3696$ *binary digits, then any logical network with* 3696 *inputs which decides truth of these sentences contains at least* 10^{123} *operations.*

We remind the reader that the radius of a proton is approximately 10^{-13} cm, and the radius of the known universe is approximately 10^{28} cm. Thus for sentences of length 616, a network whose atomic operations were performed by transistors the size of a proton connected by infinitely thin wires would densely fill the entire universe.

References

[Bu60] J. R. Büchi, *Weak second-order arithmetic and finite automata*, Z. Math. Logik Grundlagen Math. **6** (1960), 66–92. MR **23** #A2317.

[Col74] G. E. Collins, *Quantifier eliminations for real closed fields by cylindrical algebraic decomposition*, preliminary report, Proc. EUROSAM 74, ACM SIGSAM Bull. **8** (1974), 80–90.

[Coo71] S. A. Cook, *The complexity of theorem proving procedures*, Proc. Third ACM Sympos. on Theory of Computing, 1971, pp. 151–158.

[Eh72] A. Ehrenfeucht, *Practical decidability*, Report CU-CS-008-72. Dept. of Computer Science, Univ. of Colorado, 1972.

[El61] C. C. Elgot, *Decision problems of finite automata design and related arithmetics*, Trans. Amer. Math. Soc. **98** (1961), 21–51. MR **25** #2962.

[Fe74] J. Ferrante, *Some upper and lower bounds on decision procedures in logic*, Doctoral Thesis, Dept. of Mathematics, M. I. T., Cambridge, Mass., 1974.

[FeRa73] J. Ferrante and C. Rackoff, *A decision procedure for the first order theory of real addition with order*, SIAM J. Computing **4** (1975), 69–76.

[FM75] M. J. Fischer and A. R. Meyer, (1975) (in preparation).

[FR74] M. J. Fischer and M. O. Rabin, *Super-exponential complexity of Presburger arithmetic*, Complexity of Computation, edited by R. Karp, SIAM-AMS Proc., vol. **7**, Amer. Math. Soc. (1974), 27–41.

[Ka72] R. M. Karp, *Reducibility among combinatorial problems*, Complexity of Computer Computations, edited by R. E. Miller and J. W. Thatcher, Plenum Press, New York, 1972, pp. 85–104.

[Me72] A. R. Meyer, *Weak* S1S *cannot be decided*, Notices Amer. Math. Soc. **19** (1972), A-598. Abstract #72T-E67.

[Me73] ———, *Weak monadic second order theory of successor is not elementary-recursive*, Boston Univ. Logic Colloq. Proc. 1975 (to appear); also: Project MAC Technical Memo 38, M.I.T., Cambridge, Mass., 1973.

[MR75] A. R. Meyer and C. Rackoff, (1975) (in preparation).

[Mo74] L. Monk, *An elementary-recursive decision procecure for* $TH(R, +, \cdot)$, Dept. of Math., Univ. of California, 1974 (manuscript).

[Op73] D. C. Oppen, *Elementary bounds for Presburger arithmetic*, Proc. Fifth ACM Sympos. on Theory of Computing, Austin, Texas, 1973, pp. 34–37.

[Rab69] M. O. Rabin, *Decidability of second-order theories and automata on infinite trees*, Trans. Amer. Math. Soc. **141** (1969), 1–35. MR **40** #30.

[Rac74a] C. Rackoff, *On the complexity of the theories of weak direct products*, Project MAC Technical Memo 42, M.I.T., Cambridge, Mass., 1974; J. Symbolic. Logic (to appear).

[Rac74b] ———, *Complexity of some logical theories*, Doctoral Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., 1974.

[So74] R. Solovay, Correspondence with A. R. Meyer, September 1974.

[St74] L. J. Stockmeyer, *The complexity of decision problems in automata theory and logic*, Doctoral Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass., 1974; Project MAC Technical Report 133, M.I.T., Cambridge, Mass., 1974.

[SM75] L. J. Stockmeyer and A. R. Meyer, *Inherent computational complexity of decision problems in logic and automata theory*, Lecture Notes in Comp. Sci., Springer-Verlag, 1975 (to appear).

M. I. T. PROJECT MAC
CAMBRIDGE, MASSACHUSETTS 02139, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Complexity of Product and Closure Algorithms for Matrices†

M. S. Paterson

0. Abstract. Some algorithms for matrix product and for "transitive closure" of matrices over various structures are described, and relationships among the complexities of these operations are demonstrated. Determining the connected components of directed graphs, finding shortest distances in a network, and recognizing context-free languages are all applications of closure algorithms in different structures.

1. Matrix product over rings. If A, B are $n \times n$ matrices over a ring, then their *product* and *sum* are given by

$$(A \cdot B)_{ik} = \sum_j a_{ij} \times b_{jk}, \qquad (A + B)_{ik} = a_{ik} + b_{ik}.$$

Counting just the elementary ring operations used by the obvious algorithm, we have n^3 multiplications and $n^3 - n^2$ additions, and so $O(n^3)$ operations for the product and n^2 operations for the sum. We shall refer loosely to these counts as the *complexity* of the corresponding algorithm.

If A, B are partitioned in a suitably consistent way as shown below, we have the identity:

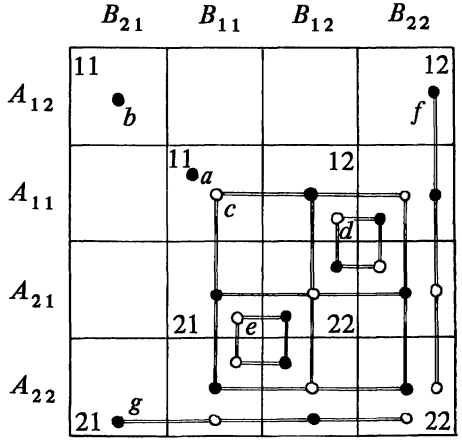
$$\left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \cdot \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right) = \left(\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right).$$

This suggests a recursive algorithm which forms the product of two $n \times n$ matrices ($n > 1$) by computing the 8 indicated products of $n/2 \times n/2$ matrices and then summing them as shown. These smaller products are computed by the same algorithm

†This research was supported in part by the National Science Foundation under grant GJ-43634X.

applied recursively. This algorithm has a complexity $P(n)$, which satisfies $P(n) \leqq 8 \cdot P(\lceil n/2 \rceil) + O(n^2)$. It is easily seen, however, that $P(n)$ is proportional to n^3.

2. Strassen's algorithm. The remarkable algorithm discovered by Strassen [**Str**] provides a key motivation for the results presented in this paper. He showed that the four submatrices $A_{i1}B_{1j} + A_{i2}B_{2j}$ for $i, j = 1, 2$, can be computed using only 7 rather than 8 small matrix products. The method is most easily presented in diagrammatic form.



In this figure a ● (○) in cell (A_{ij}, B_{kl}) represents the term $+ (-) A_{ij}B_{kl}$. The 7 connected groups each represent a set of terms which can be produced by a single product. For example, the large group of 9 terms comes from the product $(-A_{11} + A_{21} + A_{22})(B_{11} - B_{12} + B_{22})$. The four quadrants of the product matrix C are given by

$$C_{11} = a + b, \qquad C_{12} = a + c + e + f,$$
$$C_{21} = a + c + d + g, \qquad C_{22} = a + c + d + e.$$

These identities are a variant of those in [**Str**], obtained by applying a simple linear transformation. The complexity of Strassen's algorithm satisfies $P(n) \leqq 7 \cdot P(n\lceil n/2 \rceil) + O(n^2)$ which yields $P(n) \leqq O(n^\theta)$ where $\theta = \log_2 7 \approx 2 \cdot 8$.

Since recurrence relations of the above type will occur frequently, we present here a more general solution.

LEMMA. *Let $G(n) = n^\beta \cdot h(n)$ where $\beta \geqq 0$ and h is a positive, nondecreasing function satisfying $h(n) = o(n^\varepsilon)$ for all $\varepsilon > 0$. If $a \geqq 1$, $b > 1$ and F is a nonnegative function satisfying:*

$$F(n) \leqq a \cdot F(\lceil n/b \rceil) + O(G(n))$$

then, if $\alpha = \log_b a$,

$$\begin{aligned} F(n) &\leqq O(n^{\alpha}) && \text{if } \alpha > \beta, \\ &\leqq O(G(n)) && \text{if } \alpha < \beta, \\ &\leqq O(G(n)\cdot \log n) && \text{if } \alpha = \beta. \end{aligned}$$

3. Matrix inversion and determinants. For fields these immediate applications of Strassen's product algorithm are given in [**Str**]. Since

$$\left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)^{-1} = \left(\begin{array}{c|c} A_{11}^{-1} + A_{11}^{-1}A_{12}E^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}E^{-1} \\ \hline -E^{-1}A_{21}A_{11}^{-1} & E^{-1} \end{array}\right)$$

where $E = A_{22} - A_{21}A_{11}^{-1}A_{12}$ and we assume A_{11} and E to be nonsingular, we get a recursive algorithm for matrix inverse with complexity $I(n)$ satisfying

$$I(n) \leqq 2\cdot I(\lceil n/2 \rceil) + O(P(n)) + O(n^2) \leqq O(n^{\theta})$$

where $\theta = \log_2 7$ by the general solution given above. Similarly, we have for determinants that

$$\det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \det(A_{11})\cdot \det(E)$$

and so the complexity $D(n)$ satisfies

$$D(n) \leqq 2\cdot D(\lceil n/2 \rceil) + I(\lceil n/2 \rceil) + O(P(\lceil n/2 \rceil))$$

and so $D(n) \leqq O(n^{\theta})$. A pivotal algorithm for inversion which does not require the nonsingularity assumption above and which has the same order of complexity is given in [**Bun**].

A converse reduction shows that inversion cannot be appreciably easier than product,

$$\begin{pmatrix} I & A & 0 \\ 0 & I & B \\ 0 & 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} I & -A & A\cdot B \\ 0 & I & -B \\ 0 & 0 & I \end{pmatrix}$$

where I is the identity matrix and 0 is the zero matrix. Therefore $P(n) \leqq I(3n)$.

4. Boolean matrices. Matrices with domain $\{0, 1\}$ (for {false, true}) are used to represent relations on a finite set, where, e.g., $A_{ij} = 1$ iff (i, j) is in the relation. Composition of relations corresponds to matrix product with the Boolean operations $\wedge$, $\vee$. We write $A \boxtimes B$ for this product when we want to make the operations explicit.

$$(A \boxtimes B)_{ik} = \bigvee_j a_{ij} \wedge b_{jk}.$$

Also, $A \vee B$ corresponds to matrix sum where $(A \vee B)_{ik} = a_{ik} \vee b_{ik}$, and we write $A \subseteq B$ for $A \vee B = B$. A relation R is reflexive and transitive if R contains the identity relation and also $R \circ R$.

Thus the (*reflexive and*) *transitive closure* A^* of a square matrix A is the unique minimal matrix satisfying

$$X \supseteq A \vee I \vee (X \boxtimes X)$$

where I is the identity matrix. If we define $A^0 = I$, $A^{n+1} = A \boxtimes A^n$ for $n \geqq 0$, then

since the Boolean product is associative, it is easily shown that $A^* = \bigvee_{n \geqq 0} A^n$.

The following useful formula for the closure of a partitioned matrix is given and proved for regular algebras in [**Con**]:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^* = \begin{pmatrix} A_{11}^* \vee A_{11}^* A_{12} E^* A_{21} A_{11}^* & A_{11}^* A_{12} E^* \\ E^* A_{21} A_{11}^* & E^* \end{pmatrix}$$

where $E = A_{22} \vee A_{21} A_{11}^* A_{12}$.

If $P(n)$ is the complexity of Boolean product and $C(n)$ the complexity of closure, then based on the above formula we have $C(n) \leqq 2 \cdot C(\lceil n/2 \rceil) + O(P(n)) + O(n^2)$ and so if P satisfies the conditions on G in the general solution we have $C(n) \leqq O(P(n))$. Also, since

$$\begin{pmatrix} 0 & A & 0 \\ 0 & 0 & B \\ 0 & 0 & 0 \end{pmatrix}^* = \begin{pmatrix} I & A & A \cdot B \\ 0 & I & B \\ 0 & 0 & I \end{pmatrix}$$

we have $P(n) \leqq C(3n) \leqq O(P(n))$.

The obvious similarities with the results for matrix inversion over a field stem from the analogy between $(I - A)^{-1} \approx \sum_{n \geqq 0} A^n$ in a field and $A^* = \bigvee_{n \geqq 0} A^n$ for the Boolean operations.

5. Reduction of Boolean product to ring product. Strassen's method provides a subcubic algorithm for matrix products over a ring but it is not directly applicable to Boolean product. However, we can use the homomorphism $h: \boldsymbol{Z} \to \{0, 1\}$ defined by

$$\begin{aligned} h(n) &= 1 \quad \text{if } n \neq 0, \\ &= 0 \quad \text{if } n = 0, \end{aligned}$$

under which $\times$, $+$ correspond to $\wedge$, $\vee$, respectively. If h is extended to matrices in the obvious way, we have $A \boxtimes B = h(A \boxtimes B)$ for all 0, 1-matrices. Better still is to use Strassen's algorithm on $\boldsymbol{Z}_{n+1}$, the ring of integers modulo $n + 1$, since the ring product of two $n \times n$ matrices over $\{0, 1\}$ in $\boldsymbol{Z}$ has entries no greater than n. In this case the basic ring operations can be expressed as sequences of Boolean operations acting on binary representations of the ring elements. In this way we obtain a Boolean product algorithm, and hence a closure algorithm of order $O(n^\theta \cdot (\log n)^2)$. For more details of these reductions, see [**Fis**], [**Fur**] and [**Mun**].

If we regard Boolean matrices as representing directed graphs ($A_{ij} = 1$ iff there is an arc from i to j) then the closure matrix gives the connectivity relation ($A_{ij}^* = 1$ iff there is a path from i to j). When the domain is taken to be {nonnegative reals} $\cup$ $\{+\infty\}$, a matrix represents a directed network with distances on the arcs. The closure of such a matrix with respect to the operations $+$, minimum, i.e., "product" is $\underset{\min}{+}$, "sum" is min, gives the lengths of the shortest path between each pair of nodes. The recursive formula for closure in terms of product is still valid with $+$, min in place of $\wedge$, $\vee$, and hence the close relation between the complexity of product and closure. In this case, however, no subcubic algorithm for

product is known. Different partitions of the matrix yield a family of $O(n^3)$ algorithms, including the well-known Floyd algorithm given in **[Flo]**.

6. Context-free language recognition [Val]. Our final example which introduces some new difficulties concerns context-free languages. With no essential loss of generality we can define a *context-free grammar* (cfg) G as follows. G consists of a finite alphabet $\Sigma = \{A_1, \cdots, A_k\}$ and a finite set P of *productions*, each of the form "$A_i \to A_jA_k$". We write $w_1A_iw_2 \Rightarrow w_1A_jA_kw_2$ where w_1, w_2 are arbitrary strings over Σ and $A_i \to A_jA_k$ is in P. Let "$\Rightarrow^*$" be the reflexive and transitive closure of the "$\Rightarrow$" relation. For each $A_i \in \Sigma$ we define the *language generated by* A_i as $L_{A_i} = \{w \mid A_i \Rightarrow^* w\}$. The *recognition problem* for G is to determine for a string $w \in \Sigma^*$ whether or not $w \in L_{A_1}$. For a cfg G we define the binary operator $\otimes$ over the domain of subsets of Σ by

$$S_1 \otimes S_2 = \{A_i \mid \exists\, A_j \in S_1, \exists\, A_k \in S_2 \text{ and } A_i \to A_jA_k \in P\}.$$

We can then define a $\boxtimes$-product and $\bigcup$-sum for matrices with subsets of Σ as elements. Then the *transitive closure* A^+ of such a matrix A is the minimal matrix satisfying: $X \supseteq A \cup (X \boxtimes X)$. The $\otimes$-operator and hence the corresponding matrix product is *not* associative and so a new algorithm for transitive closure is necessary. The correspondence with the recognition problem is the following. If w is the string $X_1 \cdots X_{n-1} \in \Sigma^{n-1}$, let M_w be the $n \times n$ matrix defined by

$$M_{i,i+1} = \{X_i\} \quad \text{for } i = 1, \cdots, n-1,$$
$$M_{i,j} = \emptyset \quad \text{if } j \neq i+1.$$

LEMMA. *For any* w, *and* $1 \leqq u < v \leqq n$,

$$(M_w^+)_{uv} = \{A_i \mid A_i \Rightarrow^* X_u \cdots X_{v-1}\}.$$

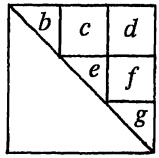
PROOF. By induction on $v - u$.

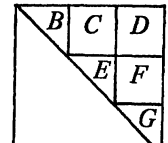
Thus, to solve the recognition problem it is sufficient to compute M_w^+. Younger's algorithm **[You]** computes this matrix in a straightforward way in $O(n^3)$ operations.

Valiant **[Val]** obtains a subcubic algorithm by a reduction to Strassen's method. He first shows that the $\boxtimes$-product for any cfg is of the same order as Boolean product. For a cfg with only one production the product is isomorphic with Boolean product. The matrix product for an arbitrary cfg is obtainable as the union of the products corresponding to each production as a singleton. The remaining step is to show that transitive closure is of the same order as product for *strictly upper triangular* (s.u.t) matrices, i.e., $A_{ij} = \emptyset$ if $j \leqq i$.

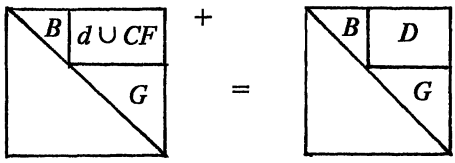
LEMMA *If* A *is s.u.t. then* A^+ *is the unique s.u.t. matrix satisfying* $X = A \cup (X \boxtimes X)$.

If

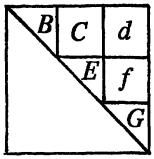
$A =$ 

is s.u.t. and $A^+ =$ 

then, by the last lemma, $D = d \cup BD \cup CF \cup DG$. Thus



again by the lemma. We therefore have a recursive algorithm $\boldsymbol{Q}$ for closing a s.u.t. matrix of the form



where the submatrices (B, C, E) and (G) are already closed. We apply $\boldsymbol{Q}$ to the submatrix (E, f, G) to compute F, then compute the product CF, and finally compute D by applying $\boldsymbol{Q}$ to the matrix $(B, d \cup CF, G)$ as above. If the linear dimensions of the partition are in the ratio $b:e:g = 1:1:\sqrt{2}$ and $G(n)$ is the complexity of $\boldsymbol{Q}$, we have the recurrence relation

$$Q(n) \leqq 2 \cdot Q(\lceil n/\sqrt{2} \rceil) + O(P(n)).$$

For a closure algorithm $\boldsymbol{C}$ we close matrices (b, c, e) and (g) recursively and then apply $\boldsymbol{Q}$. Hence

$$C(n) \leqq C(\sqrt{2} \cdot n/(1 + \sqrt{2})) + C(n/(1 + \sqrt{2})) + Q(n) \leqq 2 \cdot C(2n/3) + Q(n).$$

Since $P(n) \leqq O(n^{\theta}(\log n)^2)$, we derive $C(n) \leqq O(n^{\theta}(\log n)^2)$ with $\theta = \log_2 7$ also, thus achieving a subcubic recognition algorithm.

7. Conclusion. We have shown how the computational complexity for different matrix operations over various structures can be related, and in particular how any fast matrix product algorithms for rings can be applied indirectly to give correspondingly fast algorithm for finding connected components of directed graphs and for recognizing context-free languages. Strassen's $O(n^{\log_2 7})$ algorithm for matrix product provides a nontrivial application of these reductions.

References

[Bun] J. Bunch and J. Hopcroft, *Triangular factorization and inversion by fast matrix multiplication*, Math. Comp. **28** (1974), 231–236.

[Con] J. Conway, *Regular algebra and finite machines*, Chapman and Hall, London, 1971.

[Fis] M. Fischer and A. Meyer, *Boolean matrix multiplication and transitive closure*, IEEE Conf. Record of Twelfth Annual Sympos. on Switching and Automata Theory, 1971, pp. 129–131.

[Flo] R. Floyd, *Algorithm 97, shortest path*, CACM **5** (1962), 345.

[Fur] M. E. Furman, *Application of a method of rapid multiplication of matrices to the problem of finding the transitive closure of a graph*, Dokl. Akad. Nauk SSSR **194** (1970), 524 = Soviet Math. Dokl. **11** (1970), 1252. MR **42** #5833.

[**Mun**] I. Munro, *Efficient determination of the transitive closure of a directed graph*, Information Processing Letters **1** (1971), 56–58.

[**Str**] V. Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356. MR **40** # 2223.

[**Val**] L. Valiant, *General context-free recognition in less than cubic time*, J. Comput. and System Sci. **10** (1975), 308–315.

[**You**] D. Younger, *Recognition and parsing of context-free languages in time n^3*, Information and Control **10** (1967), 189–208.

UNIVERSITY OF WARWICK
COVENTRY, ENGLAND, CV4 7AL

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Families of Sets

Richard Rado

1. Notation, terminology. Let I be a set, fixed throughout this article. Its elements will be used to label the objects we shall consider. The symbol $(x_\nu : \nu \in I)$ denotes the *family* whose members are the objects x_ν. The same system of objects gives rise to the *set* $\{x_\nu : \nu \in I\}$. If I is finite we use the simpler notation (x, y, z), $\{x, y\}$, etc. To illustrate the notation for sets and families, I note that $\{3, 7, 8\} = \{7, 3, 8, 3\}$, $(3, 7, 8) \neq (7, 3, 8, 3) \neq (7, 3, 8) \neq (3, 7, 8)$. We put $(x_\nu : \nu \in I) = (y_\lambda : \lambda \in J)$ only if $I = J$ and $x_\nu = y_\nu$ for all $\nu \in I$. If n is an ordinal number and $I = \{\nu : \nu < n\}$ we also write

$$I = \underline{n} = \{0, \cdots, \hat{n}\},$$

where $\hat{}$ denotes the *obliterator*, an operator whose effect consists in deleting from a well-ordered sequence the term above which it is placed. If $I = \underline{n}$ we also write $(x_\nu : \nu \in I)$ as $(x_0, \cdots, \hat{x}_n)$. In a family $F = (x_\nu : \nu \in I)$ the x_ν are the *members* of F. For every family $F = (A_\nu : \nu \in I)$ whose members are sets, we put $F_M = \bigcup_\nu (\nu \in M) A_\nu$ for all $M \subseteq I$, and

$$F_{[M]} = \bigcap_\nu (\nu \in M) A_\nu \quad \text{if} \quad M \neq \emptyset, \qquad A_{[\emptyset]} = \emptyset.$$

Throughout, from now on, we use the notation $F = (A_\nu : \nu \in I)$; $G = (B_\nu : \nu \in I)$, where I, A_ν, B_ν are sets.

We put $F \cong G$ (*strong isomorphism*) whenever there is a bijection $f : A_I \to B_I$ such that[1] $f(A_\nu) = B_\nu$ for all $\nu \in I$.

Put $F \simeq G$ (*weak isomorphism*) whenever there is a bijection $\pi : I \to I$ such that $F \cong (B_{\pi(\nu)} : \nu \in I)$. The *atoms*

[1] I do not hesitate to risk the well-known ambiguity inherent in the notation $f(A_\nu)$.

$$A_{M,I} := A_{[M]} \setminus A_{I \setminus M}$$

of the family F are pairwise disjoint and, apart from this restriction, can be prescribed arbitrarily. We have $F \cong G$ if and only if $|A_{M,I}| = |B_{M,I}|$ for all $M \subseteq I$.

The relation $A \subset\subset B$ means, by definition, that $A \subseteq B$ and $|A| < \aleph_0$.

2. The reconstruction problem. Let $I \neq \emptyset$. Put, for all $\nu_0 \in I$,

$$F_{\nu_0} = (A_\nu : \nu \in I \setminus \{\nu_0\}),$$

and similarly for G. Suppose that, for some family F, the isomorphism class of each of the familes F_{ν_0} is known. Does this imply a knowledge of the isomorphism class of F? To express this question more precisely: does $F_{\nu_0} \cong G_{\nu_0}$ for all $\nu_0 \in I$ imply $F \cong G$? The following theorem [1] shows that the answer is in the negative.

THEOREM 1. *Let $I \neq \emptyset$. Then there are families F, G such that $F_{\nu_0} \cong G_{\nu_0}$ for all $\nu_0 \in I$, but $F \ncong G$ and hence, $F \ncong G$.*

In fact, assume that $D_\alpha \cap D_\beta = \emptyset$ for $\alpha \neq \beta$, and, for some $\nu_1 \in I$,

$$0 < |D_{\nu_1}| \leqq |D_\nu| \geqq \aleph_0 \quad \text{for all } \nu \neq \nu_1.$$

Put $A_{\nu_1} = D_I$ and $B_{\nu_1} = D_I \setminus D_{\nu_1}$, whereas $A_\nu = B_\nu = D_\nu$ for all $\nu \neq \nu_1$. Then the required relations hold.

Here is a positive theorem.

THEOREM 2. *Let $|A_\nu| < \aleph_0$ for all $\nu \in I$ and $(A_\nu : \nu \in J) \cong (B_\nu : \nu \in J)$ for all $J \subset\subset I$. Then $F \cong G$.*

It is worth mentioning that this result does not seem to follow from general compactness considerations.

Problem 1. Is there an analogue to Theorem 2 in which $\aleph_0$ is replaced by some other cardinal?

3. $\bigcup$-and $\bigcap$-representation. Let, for every $M \subseteq I$, $f(M)$ be a cardinal. Under what conditions does there exists a family F such that

Case 1. $|A_M| = f(M)$ for all $M \subseteq I$ ($\bigcup$-*representability of* f),

Case 2. $|A_{[M]}| = f(M)$ for all M ($\bigcap$-*representability*)?

There are trivial necessary conditions, such as

in Case 1: $f(\emptyset) = 0$; $f(M) \leqq f(N)$ for $M \subset N$;

in Case 2: $f(\emptyset) = 0$; $f(M) \geqq f(N)$ for $\emptyset \subset M \subset N$.

The next theorem [2] settles the two representability problems in the cases when either (i) $|I| < \aleph_0$ and f arbitrary or (ii) I arbitrary and $f(M) < \aleph_0$ for all M.

THEOREM 3. (1) *Let $|I| < \aleph_0$ and $f(M)$ be a cardinal for all $M \subseteq I$. Then f is $\bigcup$-representable if and only if: $f(\emptyset) = 0$; $f(M) \leqq f(N)$ for $M \subset N$; whenever $f(M_0) \geqq \aleph_0$, then there is $\nu_0 \in M_0$ with $f(M_0) = f(\{\nu_0\})$.*

(1′) *Let $f(M) < \aleph_0$ for all M. Then f is $\bigcup$-representable if and only if: For every $J \subset\subset I$, the function $(f(M) : M \subseteq J)$ is $\bigcup$-representable; whenever $M_0 \subseteq I$, there is $M_1 \subset\subset M_0$ with $f(M_0) = f(M_1)$; if $M \subset N$, then $f(M) \leqq f(N)$.*

(2) *Let* $|I| < \aleph_0$ *and* $f(M)$ *be a cardinal for all* M. *Then* f *is* $\bigcap$*-representable if and only if*: $f(\emptyset) = 0$; $f(M) \geqq f(N)$ *for* $\emptyset \subset M \subset N$; *whenever* $\emptyset \subset M_0 \subseteq I$, *then*

$$\sum_N (M_0 \subseteq N; |N \setminus M_0| \text{ even}) f(N) \geqq \sum_N (M_0 \subset N; |N \setminus M_0| \text{ odd}) f(N).$$

(2′) *Let* $f(M) < \aleph_0$ *for all* M. *Then* f *is* $\bigcap$*-representable if and only if*: $f(\emptyset) = 0$; *whenever* $\emptyset \subset M_0 \subseteq I$ *and* $N_0 \subset\subset I \setminus M_0$, *then*

$$\sigma(M_0, N_0) := \sum_P (P \subseteq N_0)(-1)^{|P|} f(M_0 \cup P) \geqq 0;$$

whenever $\emptyset \subset M_0 \subseteq I$, *then, for all* $M_0 \supset \emptyset$,

$$\sum_M (M_0 \subseteq M \subseteq I) \inf\{\sigma(M, N): N \subset\subset I \setminus M\} = f(M_0).$$

Problem 2. Let $|I| = \aleph_0$; $f(M) \leqq \aleph_0$ for all $M \subseteq I$; $f(M_0) = \aleph_0$ for some M_0. Find conditions for $\bigcup$-representability (or $\bigcap$-representability) of f.

Problem 3. Let I, f, g be given, where $f(M)$ and $g(M)$ are cardinals, for all $M \subseteq I$. Find conditions for the existence of a family F satisfying, for all M, the equations $|A_M| = f(M)$ and $|A_{[M]}| = g(M)$. If $I, f(M), g(M)$ are finite it is easy to state such conditions.

There are some results which are closely linked to the representation problem. I mention the following, which is due to P. Erdös, A. Hajnal, E. C. Milner and R. Rado.

Let $\aleph_\alpha$ be regular and let, for $\mu < \nu < \omega_{\alpha+1}$, $f(\mu, \nu) \in \{0, 1\}$. Then there are sets $A_0, \cdots, \hat{A}_{\omega_\alpha} \in \omega_\alpha^3$ such that the order types tp satisfy the conditions: $\operatorname{tp} A_\nu = \omega_\alpha^2$ for all ν; if $\mu < \nu$, then

$$\begin{aligned} \operatorname{tp}(A_\mu \cap A_\nu) &< \omega_\alpha \quad \text{if } f(\mu, \nu) = 0, \\ &= \omega_\alpha \quad \text{if } f(\mu, \nu) = 1. \end{aligned}$$

4. Families with exactly one transversal. A *transversal* of F is, by definition, a family $(x_\nu: \nu \in I)_=$ such that $x_\nu \in A_\nu$ for all $\nu \in I$. Thus the family $(\{1, 2\}, \{1, 2\})$ has exactly two transversals, viz: (1, 2), (2, 1). Much work has been done on answering the question whether a given family has at least one transversal. I consider here the question whether F has exactly one transversal.

For every ordinal $\rho \geqq 0$, define, inductively, a set I_ρ by putting

$$I_\rho = \{\nu \in I: |A_\nu \setminus A_{I_\rho}| = 1\} \quad \text{for all } \rho \geqq 0.$$

I recall that $A_{I_\rho} = \bigcup_\mu (\mu \in I_0 \cup \cdots \cup \hat{I}_\rho) A_\mu$. It easily follows that $I_\sigma \cap I_\rho = \emptyset$ if $\sigma < \rho$. Hence there is $\rho_* = \min\{\rho: I_\rho = \emptyset\}$. Put $I_* = I_{\rho_*}$. For each $\nu \in I_*$, there is a unique pair (ρ, a_ν) such that $\nu \in I_\rho$ and $A_\nu \setminus A_{I_\rho} = \{a_\nu\}$. We have [2]

THEOREM 4. *For every family* F, *the statements* (i), (ii), (iii) *are equivalent, where*

(i) F *has exactly one transversal.*

(ii) $I_* = I$. *If* $\rho < \rho_*$, $\mu, \nu \in I_\rho$ *and* $\mu \neq \nu$, *then* $a_\mu \neq a_\nu$.

(iii) *There is an ordinal* k *and a representation* $I = \{\nu_0, \cdots, \hat{\nu}_k\}_{\neq}$ *and a sequence* $(x_0, \cdots, \hat{x}_k)_{\neq}$ *such that*

$$x_i \in A_{\nu_i} \cap \{x_0, \cdots, x_i\} \quad \textit{for all } i < k.$$

The equivalence of (i) and (ii) has been proved independently by K. P. Steffens.

5. The cardinal module. How many mathematicians have not wished at one time or another that the notion of a set embodied something like multiplicity of occurrence! Also, in some way, two families such as (x, y, z, y) and (y, z, x, y) should not really be considered as essentially different from one another; how could our notation take account of this? I propose to offer one way of accomplishing this and, so I believe, much more by introducing a structure I call the cardinal module [2].

Let us use the following notation:

V = the universe = the class of all objects (nothing but a convenient notation),

Σ = the class of all sets (we use the Zermelo-Fränkel axioms with axiom of choice),

Γ = the class of all cardinals, $\Sigma^+ = \Sigma \setminus \{\varnothing\}$; $\Gamma^+ = \Gamma \setminus \{0\}$.

Let Ω denote the class of all mappings $\phi : V \to \Gamma$ which are such that $\phi^{-1}(\Gamma^+) \in \Sigma$. I call Ω the *cardinal module* because, in a certain sense, Ω is a module over Γ. The elements of Ω are called m-sets (multisets). I will now describe operations and relations on Ω.

Let $\phi_\nu \in \Omega$ for $\nu \in I$. I define

$$\sum_\nu (\nu \in I)\phi_\nu := \psi,$$

where $\psi(x) = \sum_\nu \phi_\nu(x)$ for all $x \in V$.

Since $\psi^{-1}(\Gamma^+) = \bigcup_\nu \phi_\nu^{-1}(\Gamma^+) \in \Sigma$, we have $\sum_\nu \phi_\nu \in \Omega$. For finite I one also writes $\phi_0 + \phi_1$, etc. This addition is associative and commutative. If, moreover, $I \in \Sigma^+$ then I put

$$\prod_\nu \phi_\nu := \psi',$$

where $\psi'(x) = \prod_\nu \phi_\nu(x)$ for all $x \in V$.

For $a \in \Gamma$ and $\phi \in \Omega$ I put $a\phi := \phi'$, where $\phi'(x) = a\phi(x)$ for all $x \in V$.

I introduce a valuation on Ω by putting, for $\phi \in \Omega$,

$$|\phi| = \sum_x (x \in \phi^{-1}(\Gamma^+))\, \phi(x).$$

This valuation is strongly additive. If $a_\nu \in \Gamma$ and $\phi_\nu \in \Omega$ for all $\nu \in I$, then

$$\Big|\sum_\nu a_\nu \phi_\nu\Big| = \sum_\nu a_\nu |\phi_\nu|.$$

If, in addition, $I \in \Sigma^+$ then

$$\Big|\prod_\nu \phi_\nu\Big| \leqq \prod_\nu |\phi_\nu|.$$

Here the inequality sign may hold. Thus, if $|I| \geqq 2$, $A_\nu \in \Sigma^+$ for $\nu \in I$; $\phi_\nu(x) = 1$ for $x \in A_\nu$; $\phi_\nu(x) = 0$ for $x \in V \setminus A_\nu$; $A_\mu \cap A_\nu = \varnothing$ for $\mu \neq \nu$, then $\big|\prod \phi_\nu\big| = 0 < \prod |\phi_\nu|$.

Finally, I associate with every family $H = (x_\nu : \nu \in I)$ the m-set ϕ_H given by the rule:

$$\phi_H(x) = \left|\{\nu \in I : x_\nu = x\}\right| \quad \text{for all } x \in V.$$

Then $|\phi_H| = |I|$. Given any m-set ϕ, there always is a family H with $\phi = \phi_H$. If $H = (x_\nu : \nu \in I)$ and $K = (y_\lambda : \lambda \in J)$, then $\phi_H = \phi_K$ if and only if there is a bijection $f : I \to J$ such that $x_\nu = y_{f(\nu)}$ for all $\nu \in I$, i.e., if H and K differ only in the rather trivial way described at the beginning of this section. If $L \in \Sigma$ and H_λ is a family, for each $\lambda \in L$, and if, for some family K, $\sum_\lambda (\lambda \in L)\phi_{H_\lambda} = \phi_K$, then the family K is obtained by combining, with proper regard to multiplicities, the families H_λ.

I define a partial order on Ω by putting $\phi \leqq \psi$ whenever $\phi(x) \leqq \psi(x)$ for all $x \in V$, and $\phi < \psi$ whenever $\phi \leqq \psi$ and $\phi \neq \psi$. Then $(\Omega, <)$ is a complete lattice. For every $A \in \Sigma$, I define the *characteristic m-set* χ_A by putting $\chi_A(x) = 1$ for $x \in A$, and $\chi_A(x) = 0$ for $x \in V \setminus A$. Then, for $A, B \in \Sigma$, we have $|\chi_A| = |A|$; $\chi_{A \cap B} = \chi_A \chi_B$; $\chi_{A \cup B} \leqq \chi_A + \chi_B$, with equality if and only if $A \cap B = \emptyset$. Thus Σ is embedded in Ω by the mapping $A \mapsto \chi_A$, under preservation of the valuation and of multiplication and, partially, of addition. In order to obtain the full benefit from the introduction of the cardinal module, I identify χ_A with A. This creates no confusion because different symbols are used for operations and relations in Σ and in Ω.

In Σ : operations $\cup$, $\cap$, $\setminus$, $|\cdots|$, and the relation $\subset$.
In Ω : operations $+$, $\cdot$, $-$, $|\cdots|$, and the relation $<$.

Thus, $\emptyset$ (in Σ) is identified with 0 (in Ω), and we have, for all $A, B \in \Sigma$, $A \cap B = A \cdot B$; $A \cup B \leqq A + B$ with equality if and only if $A \cdot B = 0$; $A \subset B$ if and only if $A < B$, etc.

The following theorem asserts that the m-sets χ_A $(= A)$ form a basis of Ω over Γ:

THEOREM 5. *Every m-set ϕ has a unique representation $\phi = \sum_\nu (\nu < n) a_\nu A_\nu$, where n is an ordinal; $\{a_0, \cdots, \hat{a}_n\}_< \subset \Gamma^+$; $A_0, \cdots, \hat{A}_n \in \Sigma^+$, and $A_\mu \cap A_\nu = \emptyset$ for $\mu < \nu < n$.*

In conclusion, I mention some identities that hold in Ω.

$$A \cup B + A \cap B = A + B \quad \text{for all } A, B \in \Sigma. \tag{1}$$

This is a special case of the equations (2) below. Let $I \neq \emptyset$. It can be shown [1] that in the power set $\mathscr{P}(I)$ one can define a *parity relation* $\equiv$, i.e., an equivalence relation having two classes and satisfying the condition $M \not\equiv M \setminus \{x\}$ whenever $x \in M \subseteq I$. In fact, if $|I| < \aleph_0$ then the only parity relation is the relation given by $M \equiv N$ if and only if $|M| - |N|$ is even, and if $|I| \geqq \aleph_0$ then there are $2^{2^{|I|}}$ parity relations. We have, for every $I \in \Sigma^+$ and every parity relation $\equiv$ on $\mathscr{P}(I)$,

$$\begin{aligned} \sum_M (\emptyset \equiv M \subseteq I) A_M + A_{[I]} &= \sum_N (\emptyset \not\equiv N \subseteq I) A_N, \\ \sum_M (\emptyset \equiv M \subseteq I) A_{[M]} + A_I &= \sum_N (\emptyset \not\equiv N \subseteq I) A_{[N]}. \end{aligned} \tag{2}$$

These equations have been used to carry out the discussion of the $\cup$- and $\cap$-representability mentioned in §3.

Quite apart from its applications, the cardinal module seems worth studying for its own sake. At its lowest it offers a convenient vehicle for expressing operations in combinatorial set theory.

References

1. C. Berge and R. Rado, *Note on isomorphic hypergraphs and some extensions of Whitney's theorem to families of sets*, J. Combinatorial Theory Ser. B **13** (1972), 226–241. MR **47** #72.

2. R. Rado, Ann. Mat. Pura. Appl (to appear).

THE UNIVERSITY
READING, ENGLAND RG6 2AN

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Some Results in Algebraic Complexity Theory

Volker Strassen

1. Summary. The minimal number of multiplications/divisions involved in various problems of symbolic manipulation of polynomials and rational functions is investigated.

2. Definitions. A finite set of rational functions can always be computed (i.e., evaluated) without loss in efficiency on most inputs by a program not containing any branching instructions. The sequence of intermediate results produced by implementing such a program on an idealized computer is called a computation. All elements of a computation are rational functions in the input variables. Formally, we have

DEFINITION 1. Let k be an infinite field, $x_1, \cdots, x_m$ indeterminates over k. A finite sequence β from $k(x_1, \cdots, x_m)$ is called a computation in $k(x_1, \cdots, x_m)$ iff each element of β is either an indeterminate, or an element of k, or is obtained from two previous elements by applying addition, subtraction, multiplication or division. β computes a finite set $\{f_1, \cdots, f_r\}$ of rational functions iff each f_i occurs in β.

The indeterminates are interpreted as inputs; the elements of k appearing in β are thought of as being stored in the program.

The running time of a program will depend on how long it takes the computer to perform the various arithmetic operations. For mathematical convenience we will assume that k-linear operations are instantaneous. Thus we define the running time or length of a computation β as the number of elements of β which are neither indeterminates nor are k-linearly dependent on the set of previous elements. If k has characteristic 0, most of the results below remain correct when all multiplications and divisions are counted.

DEFINITION 2 (OSTROWSKI [15]). Let $f_1, \cdots, f_r \in k(x_1, \cdots, x_m)$. $L(f_1, \cdots, f_r) :=$ minimal length of computations in $k(x_1, \cdots, x_m)$ which compute $\{f_1, \cdots, f_r\}$ is

called the complexity of $\{f_1, \cdots, f_r\}$.

Upper bounds for the complexity are usually proved by exhibiting an algorithm. Thus Horner's rule implies

$$L(a_0x^n + \cdots + a_n) \leqq n, \tag{1}$$

considered in $k(a_0, \cdots, a_n, x)$.

3. Pan's method. In the above-mentioned paper Ostrowski conjectured that we have equality in (1). This was proved 12 years later by Pan [**16**] by an elementary but ingenious method, which consists in substituting for one indeterminate a linear combination of the others and looking at the effect of this substitution on the first nonlinear operation of a given computation. Pan's method can be successfully applied to a surprising number of simple computational problems "with general coefficients", such as the evaluation of several polynomials, of a polynomial in several indeterminates, of a homogeneous polynomial, of the product of a vector by a matrix, or of a continued fraction (see Winograd [**24**], [**25**], Borodin-Munro [**2**], Strassen [**20**]). Also an arbitrary single quadratic form can be treated, as long as char $k \neq 2$. Unfortunately, the lower bounds derived by Pan's method cannot exceed the number of inputs to the problem.

4. Nonlinear lower bounds. In the sequel, $f \sim g$, $f \asymp g$, and $f \prec g$ mean respectively that f and g are asymptotically equal, f and g have the same order of magnitude, $f = O(g)$. All logarithms are to the base 2. The proofs of the lower bounds in this and the next section use some algebraic geometry. They can be found in Strassen [**21**], [**22**] and [**23**].

Let us first consider the problem of computing the set of elementary symmetric functions in n variables:

$$\sigma_1 := \sum x_i, \quad \sigma_2 := \sum_{i<j} x_ix_j, \quad \cdots, \quad \sigma_n := x_1 \cdots x_n.$$

Clearly $L(\sigma_1) = 0$. Pan's method yields $L(\sigma_2) = n - 1$ for $k = \boldsymbol{R}$ and $L(\sigma_n) = n - 1$. Also $L(\sigma_1, \cdots, \sigma_n) \leqq n \log n$ (Horowitz [**8**]).

THEOREM 1. $L(\sigma_1, \cdots, \sigma_n) \sim n \log n$.

More generally one has

THEOREM 2. *Let F be a finite set of symmetric rational functions of transcendency degree t over k. Then $L(F) \geqq t \log(t/e)$. E.g., if* char $k = 0$ *and* $s_\rho := \sum_i x_i^\rho$ *then* $L(s_1, \cdots, s_n) \sim n \log n$.

Horner's rule is optimal for evaluating a general polynomial at one point. Is it also optimal for evaluating such a polynomial at many (say $n + 1$) general points? In other words, what is the complexity of $y_0, \cdots, y_n$ in $k(a_0, \cdots, a_n, x_0, \cdots, x_n)$, where

$$y_0 = a_0x_0^n + \cdots + a_n, \quad \cdots, \quad y_n = a_0x_n^n + \cdots + a_n? \tag{2}$$

Surprisingly, separate evaluation using Horner's rule is not optimal (Borodin and

Munro [1]). One even has the following drastic result

$$L(y_0, \cdots, y_n) \prec n \log n$$

(Fiduccia [7], Moenck and Borodin [14], amended by Sieveking [19], Strassen [21]; see also Kung [11] and the result of S. Cook in Knuth [9, p. 275]).

THEOREM 3. $L(y_0,\cdots, y_n) \geqq (n + 1) \log n$, *and therefore* $L(y_0,\cdots,y_n) \asymp n \log n$.

In contrast to Pan's result Theorem 3 remains true if $a_0,\cdots, a_n$ are replaced by arbitrary elements $\alpha_0,\cdots, \alpha_n \in k$, as long as $\alpha_0 \neq 0$. So, e.g., $L(x_0^n,\cdots,x_n^n) \sim n \log n$.

The inverse problem to evaluation is interpolation. Here inputs $x_0,\cdots, x_n$, $y_0, \cdots, y_n$ are given and the coefficients $a_0,\cdots, a_n$ of the unique polynomial of degree n that interpolates y_i at x_i are to be computed. Equivalently, $a_0,\cdots, a_n$ can be defined by (2), where now $x_0, \cdots, x_n$, $y_0,\cdots, y_n$ are interpreted as indeterminates. Again one has $L(a_0,\cdots, a_n) \prec n \log n$ (Horowitz [8], Moenck and Borodin [14]; see also Strassen [21]).

THEOREM 4. $L(a_0,\cdots, a_n) \geqq (n + 1) \log n$, *and therefore* $L(a_0,\cdots, a_n) \asymp n \log n$.

As it happens, several of the previous results are concerned with the computational complexity of going from one representation of a univariate polynomial to another: Computing the elementary symmetric functions means computing the coefficients from the roots; evaluation and interpolation relate the coefficient representation to the representation by a list of values (at $n + 1$ points). Our methods apply to several similar problems. Going from the set of roots to a list of values, going from one list of values to a new one, differentiating or integrating a polynomial given by a list of values all have a complexity of order of magnitude $n \log n$. On the other hand, one can expand a polynomial at a new point in linear time (Shaw and Traub [18]).

The problems discussed here belong to the field of symbolic manipulation (Collins [6]). Because of the constant use of modular algorithms in this area, evaluation and interpolation are of special importance. Usually one is interested in the case $k = Z_p$, since one has already applied modular reductions to integer coefficients (see Brown [4] for a typical situation). In many cases neither the base points for evaluation and interpolation nor the primes p to be used are known in advance. Thus apart from treating the base points as inputs (as we do in this paper) one has to look for algorithms that work over any Z_p (or at least over any Z_p with p not too small). Now it is easy to see that, roughly speaking, such algorithms are equivalent to algorithms over $\boldsymbol{Q}$. Since $\boldsymbol{Q}$ is an infinite field, the results of this paper apply (see Strassen [23] for a detailed discussion).

5. A problem involving branching. Let A_0, A_1 be univariate polynomials over a field k such that $n := \deg A_0 \geqq \deg A_1 \geqq 0$. For simplicity assume char $k = 0$ (but the remarks at the end of the last section apply here too). Euclid's algorithm

$$A_0 = Q_1A_1 + A_2, \quad A_1 = Q_2A_2 + A_3, \quad \cdots, \quad A_{t-1} = Q_tA_t,$$

with $\deg A_i > \deg A_{i+1}$ for $i \geqq 1$ yields the Euclidean representation $(Q_1,\cdots, Q_t, A_t)$

of the pair (A_0, A_1). From this representation one can read off several important items: the continued fraction of A_0/A_1, the greatest common divisor of A_0 and A_1, the resultant of A_0 and A_1 (Collins [5]), the discriminant of A_0 if $A_1 = A_0'$, the number of zeroes of A_0 in an arbitrary interval if $A_1 = A_0'$ and if k is the field of real numbers (Sturm). Improving the work of Lehmer [12] and Knuth [10], Schönhage [17] computes the coefficients of $Q_1, \cdots, Q_t, A_t$ from the coefficients of A_0, A_1 with $\prec n \log n$ multiplications and divisions (actually these papers are concerned with the analogous problem in number theory; the translation to polynomials is due to Moenck [13]).

Size and shape of the output $(Q_1, \cdots, Q_t, A_t)$ is determined by its sequence of degrees $\boldsymbol{d} := (d_1, \cdots, d_t, d_{t+1})$. Since $\boldsymbol{d}$ depends on the input polynomials A_0, A_1, every algorithm for computing the Euclidean representation has to use branching instructions, say of the form "if $f = 0$ then go to i else go to j", where f has been previously computed. Let $M_{\boldsymbol{d}}$ be the set of inputs for which the output has shape $\boldsymbol{d}$ and let $H(\boldsymbol{d})$ be the entropy of the probability vector that is obtained from $\boldsymbol{d}$ by normalization.

THEOREM 5. *There are constants $0 < c < c'$ with the following properties:*

(1) *For all $\boldsymbol{d}$ Schönhage's algorithm takes $< c'n(H(\boldsymbol{d}) + 1)$ multiplications and divisions on $M_{\boldsymbol{d}}$.*

(2) *For all $\boldsymbol{d}$ any algorithm that computes the Euclidean representation takes $> cn(H(\boldsymbol{d}) + 1)$ multiplications and divisions on some input of $M_{\boldsymbol{d}}$.*

Thus, roughly speaking, Schönhage's algorithm is uniformly optimal. We remark that although branching instructions themselves are not counted, every multiplication and division is counted, even if it serves only to prepare a branching instruction.

To a reader, who is interested in a detailed treatment of algebraic complexity theory, we suggest the book by Borodin and Munro [3].

References

1. A. Borodin and I. Munro, Information Processing Lett. **1** (1971), 66–68.
2. ———, J. Comput. System Sci. **6** (1972).
3. ———, *Computational complexity of algebraic and numeric problems*, American Elsevier, New York, 1975.
4. W. S. Brown, *On Euclid's algorithm and the computation of polynomial greatest common divisors*, J. Assoc. Comput. Mach. **18** (1971), 478–504. MR **46** #6570.
5. G. E. Collins, *Subresultants and reduced polynomial remainder sequences*, J. Assoc. Comput. Mach. **14** (1967), 128–142. MR **35** #6352.
6. ———, *Computer algebra of polynomials and rational functions*, Amer. Math. Monthly **80** (1973), 725–755. MR **48** #2106.
7. C. Fiduccia, Proc. Fourth Sympos. on the Theory of Computing, ACM, New York, 1972, pp. 88–93.
8. Ellis Horowitz, *A fast method for interpolation using preconditioning*, Information Processing Lett. **1** (1971/72), 157–163. MR **47** #4413.
9. D. E. Knuth, *The art of computer programming*. Vol. II: *Seminumerical algorithms*, Addison-Wesley, Reading, Mass., 1969. MR **44** #3531.

10. D. E. Knuth, *The analysis of algorithms*, Proc. Internat. Congress Math. (Nice, 1970), vol. 3, Gauthier-Villars, Paris, 1971, pp. 269–274.

11. H. T. Kung, *On computing reciprocals of power series*, Dept. of Comput. Sci., Carnegie-Mellon University, Sept. 1973.

12. D. H. Lehmer, Amer. Math. Monthly **45** (1937), 227–233.

13. R. Moenck, Proc. Fifth Sympos. on Theory of Computing, ACM, New York, 1973, pp. 142–151.

14. R. Moenck and A. Borodin, Proc. Thirteenth IEEE Sympos. on Switching and Automata Theory, Oct. 1972, pp. 90–96.

15. A. M. Ostrowski, *On two problems in abstract algebra connected with Horner's rule*, Studies in Math. Mech., Presented to Richard von Mises, Academic Press, New York, 1954, pp. 40–48. MR **16**, 523.

16. V. Ja. Pan, *On means of calculating values of polynomials*, Uspehi Mat. Nauk **21** (1966), no. 1 (127), 103–134 = Russian Math. Surveys **21** (1966), no. 1, 105–136.

17. A. Schoenhage, Acta Informatica **1** (1971), no. 1, 139–144.

18. M. Shaw and J. F. Traub, Proc. Thirteenth IEEE Sympos. on Switching and Automata Theory, Oct. 1972, pp. 105–107.

19. M. Sieveking, *An algorithm for division of power series*, Computing (Arch. Elektron. Rechnen) **10** (1972), 153–156. MR **47** #1257.

20. V. Strassen, *Complexity of computer computations*, Plenum Press, New York, 1972, pp. 1–10.

21. ———, Numer. Math. **20** (1973), 238–251.

22. ———, *Die Berechnungskomplexität der symbolischen Differentiation von Interpolationspolynomen*, Theoretical Computer Science, 1975.

23. ———, *The computational complexity of continued fractions* (in preparation).

24. S. Winograd, *On the number of multiplications necessary to compute certain functions*, Comm. Pure Appl. Math. **23** (1970), 160–170. MR **41** #4778.

25. ———, Proc. Internat. Congress Math. (Nice, 1970), vol. 3, Gauthier-Villars, Paris, 1971, pp. 269–274.

UNIVERSITY OF ZURICH
ZURICH, SWITZERLAND

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

On Sets of Integers Containing No k Elements in Arithmetic Progression

E. Szemerédi

In 1926 van der Waerden [13] proved the following startling theorem: If the set of integers is arbitrarily partitioned into two classes then at least one class contains arbitrarily long arithmetic progressions. It is well known and obvious that neither class must contain an infinite arithmetic progression. In fact, it is easy to see that for any sequence a_n there is another sequence b_n, with $b_n > a_n$, which contains no arithmetic progression of three terms, but which intersects every infinite arithmetic progression. The finite form of van der Waerden's theorem goes as follows: For each positive integer n, there exists a least integer $f(n)$ with the property that if the integers from 1 to $f(n)$ are arbitrarily partitioned into two classes, then at least one class contains an arithmetic progression of n terms. (For a short proof, see the note of Graham and Rothschild [5].) However, the best upper bound on $f(n)$ known at present is extremely poor. The best lower bound known, due to Berlekamp [3], asserts that $f(n) < n2^n$, for n prime, which improves previous results of Erdös, Rado and W. Schmidt.

More than 40 years ago, Erdös and Turán [4] considered the quantity $r_k(n)$, defined to be the greatest integer l for which there is a sequence of integers $0 < a_1 < a_2 < \cdots < a_l \leqq n$ which does not contain an arithmetic progression of k terms. They were led to the investigation of $r_k(n)$ by several things. First of all the problem of estimating $r_k(n)$ is clearly interesting in itself. Secondly, $r_k(n) < n/2$ would imply $f(k) < n$, i.e., they hoped to improve the poor upper bound on $f(k)$ by investigating $r_k(n)$. Finally, an old question in number theory asks if there are arbitrarily long arithmetic progressions of prime numbers. From $r_k(n) < \pi(n)$ this would follow immediately. The hope was that this problem on primes could be attacked not by

using special properties of the primes but by only using the fact that they are numerous, a method which is often successful.

Erdös and Turán observed $r_k(m+n) \leqq r_k(m) + r_k(n)$ from which it follows by a simple argument that

$$\lim_{n\to\infty} r_k(n)/n = c_k$$

exists. Erdös and Turán conjectured that $c_k = 0$ for all k. A few years later Behrend [1] proved that either $c_k = 0$ for every k, or $\lim_{k\to\infty} c_k = 1$. Erdös and Turán also conjectured $r_k(n) < n^{1-\varepsilon_k}$, which was shown to be false by Salem and Spencer [11] who proved $r_3(n) > n^{1-c/\log\log n}$. In 1946 Behrend [2] proved $r_3(n) > n^{1-c/(\log n)^{1/2}}$ which is the best lower bound for $r_3(n)$ currently known. In [6], L. Moser constructed an infinite sequence which contains no arithmetic progression of three terms and which satisfies Behrend's inequality for every n. Behrend's corresponding inequalities for $k < 4$ were improved by Rankin in [7].

The first satisfactory upper bound for $r^3(n)$ was due to Roth [8] who proved $r_3(n) < cn/\log\log n$. In 1967, I proved that $r_4(n) = o(n)$. The proof used the general theorem of van der Waerden. Roth [9], [10] later gave an analytic proof that $r_4(n) = o(n)$ which did not make use of van der Waerden's theorem (in fact, he proved a much more general theorem) and his method probably gives $r_4(n) < n/\log_l n$ where l is a large fixed integer and $\log_l$ denotes the l-fold iterated logarithm.

In this article we give a brief outline of a proof of the general conjecture of Erdös and Turán: $c_k = 0$ for all k.

The proof is rather long and complicated although it uses only elementary combinatorial arguments. Space limitations do not permit us to outline the proof here so we shall just restrict ourselves to mentioning several of the key ideas.

An important lemma used in the proof states essentially that any finite graph can be partitioned into relatively few "nearly regular" subgraphs. The basic objects with which the proof of the main theorem deals are not just arithmetic progressions themselves but rather generalizations of arithmetic progressions called *m-configurations*. Roughly speaking, a 1-configuration is just an arithmetic progression; an m-configuration is an "arithmetic progression" of $(m-1)$-configurations. In a nutshell, one can show that for any given set of integers R of positive upper density, a very long m-configuration which intersects R in a moderately regular way must always contain a shorter (but still quite long) $(m-1)$-configuration which intersects R in an even more regular way. In this way, we eventually conclude that R must contain arbitrarily long 1-configurations, i.e., arithmetic progressions, and we are done.

References

1. F. A. Behrend, *On sequences of integers containing no arithmetic progression*, Casopis Mat. Fys. Praha **67** (1938), 235–239.

2. ———, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332. MR **8**, 317.

3. E. R. Berlekamp, *A construction for partitions which avoid long arithmetic progressions*, Canad. Math. Bull. **11** (1968), 409–414. MR **38** #1066.

4. P. Erdös and P. Turán, *On some sequences of integers*, J. London Math. Soc. **11** (1936), 261–264.

5. R. L. Graham and B. L. Rothschild, *A short proof of van der Waerden's theorem on arithmetic progressions*, Proc. Amer. Math. Soc. **42** (1974), 385–386.

6. L. Moser, *On non-averaging sets of integers*, Canad. J. Math. **5** (1953), 245–252. MR **14**, 726, 1278.

7. R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh Sect. A **65** (1960/61), 332–344. MR **26** #95.

8. K. F. Roth, *On certain sets of integers*. II, J. London Math. Soc. **29** (1954), 20–26. MR **15**, 288.

9. ———, *Irregularities of sequences relative to arithmetic progressions*. III, J. Number Theory **2** (1970), 125–142. MR **41** #6810.

10. ———, *Irregularities of sequences relative to arithmetic progressions*. IV, Per. Math. Hungar. **2** (1972), 301–326.

11. R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. U.S.A. **28** (1942), 561–563. MR **4**, 131.

12. E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta. Math. Acad. Sci. Hungar. **20** (1969), 89–104. MR **39** #6861.

13. B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw. Arch. Wisk. **15** (1928), 212–216.

HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, HUNGARY

STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305, U.S.A.

Section 19

Applied Statistics, Mathematics in the Social and Biological Sciences

Труды Международного Конгресса Математиков
Ванкувер, 1974

Стохастические Динамические Модели Экономического Равновесия*

Е. Б. Дынкин

1. Многочисленные случайные факторы существенно влияют на экономическую практику. Естественно ввести их и в математические модели экономической динамики. Мы отправляемся от известной модели фон Неймана-Гейла (см., например, [7]). Её стохастический вариант был предложен в [4] —[6] и изучался далее в [8]—[9].[1] На этой основе мы построим стохастическую динамическую модель равновесия, являющуюся развитием модели Эрроу-Дебре. Как и в модели Эрроу-Дебре, мы имеем дело с конечным числом потребителей и конечным числом производителей. Главное отличие стохастического варианта состоит в том, что производители, не зная будущего, могут максимизировать лишь ожидаемую прибыль. Равновесие, которое мы строим, это равновесие между ожиданиями и их осуществлением. Точнее, строится система случайных цен и система планов, оптимальных для каждого участника при этих ценах, такие, что во всех случайных ситуациях полностью удовлетворяется спрос и равна нулю стоимость избыточных продуктов.

Возможны различные предположения о механизме формирования доходов. Мы считаем, что доходы фиксированы. Это значит, что мы не пытаемся вывести из модели размеры трудового вклада каждого участника и оплату труда, а предполагаем, что они заданы экзогенно (см. по этому поводу [11]). В модели не учитываются возможности займов и сбережений.

Другие способы введения случайных параметров в модели равновесия (в простейшем случае, когда эти параметры принимают лишь конечное множество значений) рассматривались Дебре [3], Эрроу [1] и Раднером [13].

*Delivered by D. Gale.

[1]В близком направлении работают Раднер и его сотрудники (см., например, [12]).

2. Пусть R—пространство всех m-векторов $x = (x^1, \cdots, x^m)$. Положим $x \in R_+$, если $x^k \geqq 0$ при всех k. Будем рассматривать элементы x из R_+ как наборы продуктов: x^k означает количество k-ого продукта в наборе x. Пара элементов (x, y) из R_+ описывает *производственный процесс* с затратами x и выпуском y. Для каждого $t = 1, \cdots, T$ определено множество $\mathscr{T}_t$ производственных процессов, технологически осуществимых в момент t (*технологическое множество*). Мы будем считать, что имеется несколько производителей. Возможности i-ого производителя в момент t описываются множеством $\mathscr{T}_{ti}$.

Чтобы ввести действие случайных факторов, рассмотрим вероятностное пространство $(\Omega, \mathscr{F}, P)$ с заданными в нём σ-алгебрами $\mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \cdots \subseteq \mathscr{F}_t \subseteq \cdots \subseteq \mathscr{F}$. Элементы $\mathscr{F}_t$ интерпретируются как события, наступление или ненаступление которых становится известным к моменту t. Будем предполагать, что множества $\mathscr{T}_{ti}$ зависят от ω так, что соответствия $\mathscr{T}_{ti}$ измеримы относительно $\mathscr{T}_t$:[2] технологические возможности в момент t становятся известными к этому моменту. *План i-ого производителя*—это последовательность функций $(x_t(\omega), y_t(\omega))$ $(t = 1, \cdots, T)$, удовлетворяющая условиям: (а) x_t, y_t измеримы относительно пополнения $\overline{\mathscr{F}}_t$ σ-алгебры $\mathscr{F}_t$ по мере $\boldsymbol{P}$; $(x_t(\omega), y_t(\omega)) \in \mathscr{T}_{ti}(\omega)$ при всех t, ω. Другими словами, план это последовательность $\overline{\mathscr{F}}_t$-измеримых селекторов соответствий $\mathscr{T}_{ti}(\omega)$.

Будем предполагать, что:

2.А. При любых t, i, ω множество $\mathscr{T}_{ti}(\omega)$ замкнуто, выпукло и содержит элемент $(0, 0)$.

2.Б. Найдётся компакт K, такой, что $\mathscr{T}_{ti}(\omega) \subseteq K$ при всех t, i.

3. Решение потребителя описывается вектором $c \in R_+$— набором продуктов, который он потребляет. Основными характеристиками потребителя являются множество C возможных его решений, отношение предпочтения $\succ$ на C и доход w. Предположим, что:

3.А. C — замкнутое выпуклое подмножество множества R_+, содержащее нулевой вектор.

3.Б. Отношение $\succ$ может быть описано с помощью некоторой непрерывной вогнутой функции u: $c_1 \succ c_2$ тогда и только тогда, когда $u(c_1) > u(c_2)$.

3.В. $0 < w < \infty$.

Имеется несколько потребителей. Их предпочтения и доходы меняются со временем и зависят от ω. Таким образом, имеются семейства соответствий $C_{tj}(\omega)$ и функций $u_{tj}(\omega, c)$, $w_{tj}(c)$. Предполагается, что они $\mathscr{F}_t$-измеримы относительно ω (в каждый момент известны характеристики потребителей в этот момент). *План j-ого потребителя* это последовательность функций $c_1, c_2, \cdots, c_{t+1}$, где c_t $\mathscr{F}_t$-измеримый селектор соответствия $C_{tj}(\omega)$.

---

[2]Соответствие Φ из Ω в R — это функция, сопоставляющая каждому $\omega \in \Omega$ непустое подмножество $\Phi(\omega)$ множества R. Пусть $\mathscr{F}$ σ-алгебра в Ω. Соответствие Φ мы называем $\mathscr{F}$-измеримым, если $\{\omega: \Phi(\omega) \cap A \neq \emptyset\} \in \mathscr{F}$ для любого замкнутого множества A. (Относительно свойств измеримых соответствий и их селекторов см., например, [10].)

4. Всякий элемент p из R_+ может трактоваться как вектор цен. Стоимость набора продуктов x в ценах p выражается скалярным произведением $px = \sum p^k x^k$. Для потребителя с доходом w доступны не все наборы благ $c \in C$, а только те, которые удовлетворяют неравенству $pc \leqq w$. Обозначим через $\varphi(p)$ совокупность максимальных элементов этого множества относительно $\succ$ (для строго положительного p это множество непусто). Соответствие φ, задающее спрос, зависит от t, j, ω. Доказывается, что $\varphi_{tj}(\omega, p)$ измеримо относительно $\mathscr{F}_t \times \mathscr{B}$. (Через $\mathscr{B}$ обозначается борелевская σ-алгебра в R.)

5. Последовательность функций $p_1, p_2, \cdots, p_{T+1}$ на Ω со значениями из R называется системой цен, если p_t $\mathscr{F}_t$-измерима. План c_t потребителя j назовём оптимальным при ценах p_t, если $c_t(\omega) \in \varphi_{tj}(\omega, p_t(\omega))$ при всех t, ω.

Пусть (x, y)—производственный процесс, начинающийся в момент t. Мы считаем, что затраты оплачиваются в момент t, в то время как выпуск может быть реализован лишь в момент $t+1$. Поэтому прибыль вычисляется по формуле $p_{t+1}y - p_t x$. Пусть (x_t, y_t) $(t = 1, \cdots, T)$ какой-нибудь план i-ого производителя. Соответствующая суммарная прибыль равна

$$\sum_1^T (p_{t+1}y_t - p_t x_t). \tag{1}$$

Назовём оптимальными (при ценах p_t) все планы, для которых математическое ожидание суммы (1) достигает максимума.

Пусть каждый производитель и каждый потребитель избрали какие-нибудь планы. Выражение

$$\delta_t = \sum_i y_{t-1,i} - \sum_i x_{ti} - \sum_j c_{tj} \qquad (t = 2, 3, \cdots, T+1)$$

естественно назвать *избыточным предложением в момент* t (мы полагаем $x_{T+1,i} = 0$). При $t = 1$ действует другая формула

$$\delta_t = y_0 - \sum_i x_{1i} - \sum_j c_{1j},$$

где y_0 заданный неотрицательный вектор, который называется *вектором начальных ресурсов* (или *начальным вектором*).

Система цен p_t и система планов всех участников образует *равновесие*, если план каждого участника оптимален и если $\delta_t \geqq 0$, $p_t\delta_t = 0$ при всех t, ω.

Наша цель—доказать существование равновесия при любом положительном начальном векторе y_0. Ниже будут сформулированы нужные для этого дополнительные условия и объяснены основные идеи доказательства.

6. Сделаем предварительно одно замечание, важное для доказательства и представляющее самостоятельный экономический интерес. Чтобы действовать оптимально, каждому участнику нет надобности иметь полную информацию относительно вероятностного механизма явлений и сложившейся ситуации. Потребителю достаточно знать в каждый момент свои предпочтения, доход и действующие цены. Производителю требуется знать

в момент t лишь собственные технологические возможности, действующие цены p_t и прогноз $\bar{p}_{t+1} = \boldsymbol{E}(p_{t+1} \mid \mathscr{F}_t)$ цен на один шаг вперёд.

Сказанное очевидно в случае потребителя. Чтобы доказать это в случае производителя, сопоставим произвольной паре векторов p, q из R_+ и любому ω из Ω множество $\mathscr{T}_{ti}(\omega, p, q)$ элементов (x, y) из $\mathscr{T}_{ti}(\omega)$, таких, что

$$qy - px = \sup_{(x', y') \mathscr{T}_{ti}(\omega)} (qy' - px').$$

Доказывается, что соответствие $\mathscr{T}_{ti}(\omega, p, q)$ измеримо относительно $\mathscr{F}_t \times \mathscr{B} \times \mathscr{B}$ и план (x_t, y_t) производителя i оптимален при ценах p_t тогда и только тогда, когда при каждом t,

$$(x_t, y_t) \in \mathscr{T}_{ti}(\omega, p_t, \bar{p}_{t+1}) \quad \text{(п.н.)},$$

Таким образом, чтобы действовать оптимально, производителю достаточно на каждом шаге максимизировать непосредственно ожидаемую прибыль $\bar{p}_{t+1}y - p_t x$.

7. Наиболее существенное ограничение, нужное нам для построения равновесия, можно сформулировать следующим образом.

7.А. При каждом ω выполняется условие: любой продукт в любой момент t либо необходим потребителям, либо без него технически невозможно получить в один из последующих моментов $t' = t+1, \cdots, T+1$ некоторый продукт, необходимый потребителям в этот момент.

Чтобы расшифровать это условие, введём *совокупный спрос и совокупное технологическое множество* как алгебраические суммы

$$\varphi_t(p) = \sum_j \varphi_{tj}(p), \qquad \mathscr{T}_t = \sum_i \mathscr{T}_{ti}.$$

Мы говорим, что продукт с номером k *необходим потребителям в момент* t, если при любых ценах p все векторы $c \in \varphi_t(p)$ имеют положительные k-е координаты. Продукт k' *технически невозможно получить в момент* $t' + 1$, *не имея продукта* k *в момент* t, если из соотношений

$$\begin{aligned} x_t^k &= 0, && (x_t, y_t) \in \mathscr{T}_t, \\ x_{t+1} &\leqq y_t, && (x_{t+1}, y_{t+1}) \in \mathscr{T}_{t+1}, \cdots, \\ x_{t'} &\leqq y_{t'-1}, && (x_{t'}, y_{t'}) \in \mathscr{T}_{t'} \end{aligned}$$

следует, что $y_{t'}^{k'} = 0$.

Из 7.А вытекает, что если $c_t \in \varphi_t(\omega, p_t)$ для всех t, ω и если

$$(2) \qquad (x_t, y_t) \in \mathscr{T}_t, \qquad x_t + c_t \leqq y_{t-1} \qquad (t = 1, 2, \cdots, \mathscr{T}), \qquad c_{T+1} \leqq y_T,$$

то все векторы y_t, c_t строго положительны.

Чтобы обеспечить существование нетривиальных последовательностей, удовлетворяющих соотношениям (2), введем условие:

7.Б. Существует строго положительный вектор $\hat{x}$, такой, что производ-

ственный процесс $(\hat{x}, \hat{x})$ принадлежит технологическим множествам $\mathscr{T}_t(\omega)$ при всех t, ω.

Нам понадобится также условие:

7.В. Все доходы w_{tj} имеют конечные математические ожидания.

Наконец, нужна некоторая регулярность измеримых пространств $(\Omega, \mathscr{F}_t)$. Например, достаточно предположить, что каждая σ-алгебра $\mathscr{F}_t$ порождается отображением Ω на борелевское подмножество некоторого польского пространства. Отсюда следует существование условного распределения вероятностей на $\mathscr{F}_{t+1}$ относительно $\mathscr{F}_t$, т.е. существование зависящей от ω вероятностной меры $\boldsymbol{P}_t(\omega, \cdot)$ на $\mathscr{F}_{t+1}$, такой, что $\boldsymbol{P}(A \mid \mathscr{F}_t) = P_t(\omega, A)$ (п.н.) при $A \in \mathscr{F}_{t+1}$.

8. Конструкция, с помощью которой строится равновесие, напоминает метод динамического программирования: сначала конструируется (от $t+1$ к t) некоторая последовательность соответствий, а затем с ее помощью вычисляются (от $t-1$ к t) цены и оптимальные планы участников.

Равновесие можно рассматривать как последовательность наборов

$$g_t = (x_t, y_t, c_t, p_t, q_t) \qquad (t = 1, 2, \cdots, T+1)$$

(назовем их локальными равновесиями), где $(x_t, y_t) = \sum(x_{ti}, y_{ti})$—совокупное производство, $c_t = \sum c_{tj}$—совокупное потребление. Для построения g_t существенно знать из прошлого только y_{t-1}, а из будущего только p_{t+1}. При $t = T+1$ будущее не имеет значения. Поэтому можно обозреть все локальные равновесия в момент $T+1$, отвечающие выпуску z в момент T. Обозначим множество соответствующих равновесных цен p_{T+1} через $\phi_{T+1}(\omega, z)$. Пусть мы уже знаем множество $\phi_{t+1}(\omega, z)$ всех равновесных цен в момент $t+1$, отвечающих выпуску z в момент t. Тогда множество возможных в момент t прогнозов $\bar{p}_{t+1}$ вычисляется по формуле

$$\Psi_t(\omega, z) = \int \phi_{t+1}(\omega', z)\, \boldsymbol{P}_t(\omega, d\omega'). \tag{3}$$

Подсчитаем теперь всевозможные локальные равновесия $(x_t, y_t, c_t, p_t, q_t)$ в момент t в предположении, что выпуск в момент $t-1$ равен z. Для этого надо решить следующую задачу:

$$\begin{gathered}(x_t, y_t) \in \mathscr{T}_t(\omega, p_t, q_t), \quad c_t \in \varphi_t(\omega, p_t), \quad q_t \in \Psi_t(\omega, z), \\ c_t + x_t \leqq z, \quad p_t(c_t + x_t) = p_t z.\end{gathered} \tag{4}$$

Путь к ее решению открывает теорема Какутани о неподвижной точке.[3]

Последовательность шагов можно описать следующей диаграммой

$$\begin{array}{ccccccccccc} & \phi_1 & & \phi_{t-1} & & \phi_t & & \phi_{t+1} & & \phi_{T+1} & \\ \swarrow & & \searrow\!\swarrow & & \searrow\!\swarrow & & \searrow\!\swarrow & & \searrow\!\swarrow & & \searrow \\ \Psi_1 & & \Psi_{t-1} & & \Psi_t & & \Psi_T & & \Psi_{T+1} & = & \{0\}, \end{array} \tag{5}$$

[3] При этом мы опираемся на одну лемму, доказанную В.М. Полтеровичем.

где стрелки, направленные вверх, означают решение задачи (4), а стрелки, направленные вниз, вычисление условного математического ожидания по формуле (3).

После того, как построены соответствия (5), вычисляется от $t-1$ к t последовательность $g_t = (x_t, y_t, c_t, p_t, q_t)$. Именно, мы подставляем $z = y_{t-1}$ в задачу (4) и ищем измеримое решение этой задачи, удовлетворяющее дополнительному условию $\boldsymbol{E}(p_t | \mathscr{F}_{t-1}) = q_{t-1}$.

Реализация намеченной программы требует преодоления ряда технических трудностей. Для применения теоремы Какутани необходимо, чтобы соответствие Ψ_t принимало выпуклые компактные значения. Однако множество $\Phi_{t+1}(\omega, z)$ равновесных цен не является, вообще говоря, ин выпуклым, ни компактным. Тем не менее выпуклость $\Psi_t(\omega, z)$ можно вывести из известной теоремы Ауманна (см. [2]), если все меры $\boldsymbol{P}_t(\omega, \cdot)$ безатомные. Общий случай можно исследовать разлагая эти меры на атомы и безатомную часть.

Чтобы преодолеть трудности, связанные с некомпактностью $\Phi_{t+1}(\omega, z)$, мы урезаем технологические множества, требуя, чтобы совокупный выпуск $y_t(\omega)$ был не меньше некоторой заданной положительной $\bar{\mathscr{F}}_t$-измеримой функции $\varepsilon_t(\omega)$. При определённых условиях на набор $\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T)$ и начальный вектор y_0 описанную выше программу, действительно, удается осуществить и построить некоторый суррогат равновесия—мы называем его ξ-равновесием. Затем доказывается, что ξ-равновесие является настоящим равновесием при подходящем выборе ξ. (Доказательство использует условия 7.А и 7.Б.)

Литература

1. K. J. Arrow, *The role of securities in the optimal allocation of risk-bearing*, Essays in the Theory of Risk-Bearing, Markham, Chicago, Ill., 1971, pp. 121–133.

2. R. J. Aumann, *Integrals of set-valued functions*, J. Math. Anal. Appl. **12** (1965), 1–12.

3. G. Debreu, *Theory of value: an axiomatic analysis of economic equilibrium*, Wiley, New York; Chapman & Hall, London, 1959.

4. Е. Б. Дынкин, *Некоторые вероятностные модели развивающейся экономики*, Доклады АН СССР **200** (1971), 523–525 = Soviet Math. Dokl. **12** (1971), 1422–1425.

5. ———, *Вероятностное вогнутое динамическое программирование*, Мат. сборник **87** (**129**) (1972), 490–503 = Math. USSR Sb. **16** (1972), 501–516.

6. ———, *Optimal programs and stimulating prices in the stochastic models of economic growth*, J. Łoś & M. W. Łoś, eds., Mathematical Models in Economics, PWN, Warszawa; Elsevier, New York; North-Holland, Amsterdam-London, 1974, pp. 207–218.

7. И. В. Евстигнеев. *Оптимальное экономическое планирование с учётом стационарных случайных факторов*, Доклады АН СССР **206** (1972), 1040–1042 = Soviet Math. Dokl. **13** (1972), 1357–1359.

8. I. V. Evstigneev, *On optimal stochastic programs and their stimulating prices*, J. Łoś & M. W. Łoś, eds., Mathematical Models in Economics, PWN, Warszawa; Elsevier, New York; North-Holland, Amsterdam-London, 1974, pp. 219–252.

9. D. Gale, *The closed linear model of production. Linear inequalities and related systems*, Princeton Univ. Press, Princeton, N.J, 1956, pp. 285–303.

10. W. Hildenbrand, *Core and equilibria of the large economy*, Princeton Univ. Press, Princeton, N.J., 1974.

11. В. М. Полтерович, *Экономическое равновесие и оптимум*, Экономика и математ. методы **9** (1973), 835–845.

12. R. Radner, *Optimal stationary consumption with stochastic production and resources*, J. Economic Theory **6** (1973), 68–90.

13. ———, *Existence of equilibrium of plans, prices, and price expectations in a sequence of markets*, Econometrica **40** (1972), 289–303.

Центральный Экономико-Математический Институт АН СССР
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Future of Stochastic Modelling

P. A. P. Moran

Between the subject of probability theory (i.e., the study of nonnegative σ-finite measures) and the theory of statistical inference exists the vast field of stochastic modelling. This is the study of particular scientific models or theories in which a random element enters essentially. These three fields are very closely intertwined and need each other. Here we consider the future of stochastic modelling and its past also.

Stochastic models often cannot be separated from their deterministic limits and analogues, and thus most of classical analysis may be needed to understand them. In fact the relationship between stochastic and deterministic behaviour is quite intricate, and each may shed light on the other in unexpected ways, as for example in the relationship between Brownian motion and potential theory. Probabilistic methods, in the form of numerical simulations, can in fact be used to provide approximate solutions to deterministic problems, and when Metropolis and Ulam [4] coined the phrase "Monte Carlo method" they had in mind the solution of partial differential equations in this way. It was only later that the phrase came to mean stochastic simulation in general.

Stochastic models of natural phenomena now cover an enormous field which includes all branches of statistical physics, queueing, congestion, storage, biological populations and population genetics. Of these the earliest to develop was statistical physics, beginning with gases and then moving into liquids and solids. This vast subject had no influence at all on statistical inference, nor was itself affected by the development of a rigorous mathematics for probability theory until about the late thirties. The block in communication was complete. When a statistician attempts to read any of the classical treatises on statistical mechanics written in the twenties and thirties, such as those of Tolman and Fowler, he is bewildered by the termino-

logy, the subtleties of ergodic theory, and by the habit of the physicist of calculating the "most probable" state. Since the probability of the most probable state is usually either zero or exceedingly small, it takes some time for the statistician to realise that the physicist is really attempting to show that nearly all the probability is concentrated around this state. It is only since the classical monographs [1], [2] of Khintchine that probabilists and physicists have been on speaking terms. Close and intimate relations have not yet been established.

In classical statistical mechanics we attempt to make statements about averages taken over solutions of differential equations when the number of variables is very large. The essential feature seems to be the existence of a Hamiltonian or equivalent. When this is not so, the subject is relatively unexplored. There are however a number of other probability problems in physics which are not of this type.

An interesting example is that of electric circuits containing random elements. Suppose for simplicity that these are resistances with small random variations about their nominal values. There appears to be no general theory about this situation although some isolated problems have been solved. Rayleigh pointed out in 1882 that if we take $k > 1$ resistances whose true values are distributed about an unknown value R with a small coefficient of variation C, the ratio of their resistance in series to their resistance in parallel is nearly distributed about a mean equal to k^2 with a bias and a coefficient of variation both of order C^2. This fact is used in practice for very accurate bridge circuits for the measurement of resistances. Commercial instruments using this principle with specially designed switches are now available. Another type of problem occurs when the resistances are 0/1 random variables (or more generally $0/R_i$). Here simulation is at present the most expeditious approach. This subject is related at one extreme to the theory of inhomogeneous media and at the other to percolation and reliability theories. A survey of the percolation aspects has recently been given by Kirkpatrick [3].

A problem of a different kind is that of the long-term behaviour of globular star clusters. Here we have a Hamiltonian but only a pseudo-ergodic behaviour. These remarkable symmetric objects, of which over 100 are known, may contain 10^3 to 10^5 stars. Their energy interactions cause them to settle down into a semistationary distribution, but an individual star may have a nonzero chance of acquiring enough energy to escape finally from the system which thus in the long run declines in number. A large literature exists on this subject, but clearly no exact theory is possible. This led Ulam and his colleagues to carry out some fascinating numerical experiments with 12 attracting particles, started off in a more or less random manner, and with their trajectories obtained by numerical integration. The whole subject of modelling clusters has been recently surveyed by Aarseth and is again an example where much is to be learnt by judicious numerical experimentation.

The simulation of diffusion processes also occurs in very interesting problems of navigation. One example is the long study of bird navigation (Kendall [5]). This has some similarity to another published study by Levison, Ward, and Webb of the navigation across the Pacific of the Polynesian people under the influence of random winds and ocean currents.

Time series, i.e., stationary random processes, also occur widely in physical science, often in a manner requiring much more than conventional theory. One large class of interesting problems concerns the passage of random processes through dispersive media. An example is the study of the radiation received from pulsars. These are stellar objects which emit, at regular intervals, pulses with a wide band spectrum. On their long path to the earth they pass through a dispersive medium which spreads out the different frequency bands so that on arrival at the earth the amplitude envelope is nearly constant. The problem is to recover the underlying pulse frequency. The radio-astronomers solved this problem easily without help from statisticians by using the instinctive understanding of spectral theory possessed by communication engineers, but this example shows how the probability theory needs to be extended, e.g., in the direction of studying "local spectra".

Dispersive effects with stationary processes are also illustrated by ocean waves which show astonishing and as yet not completely explained behaviour.

Until recently geometrical probability consisted mainly of a few trivial problems in elementary textbooks but is now a flourishing subject with many practical applications. The natural tool is Haar measure invariant under the Euclidean group. For rotations this is a probability measure since the group is compact (it may be pointed out that the natural way to deal with these in E^3 and E^4 is to use quaternions), but for random points, lines and planes we need Lebesgue measure and therefore conditional probabilities (a general theory has been given by Renyi).

Many natural phenomena and problems of practical measurement require geometric probability modelling. One of these is "Stereology", the estimation of the three-dimensional objects from one- and two-dimensional flat sections. This has a large literature of its own because it is of practical importance in biology and metallurgy. There are still some unsolved problems in this subject which might be soluble. Much more difficult, however, are the unsolved problems of the more or less random movement of metallurgical dislocations.

Metallurgists are also interested in the very wide variety of models for the random division of space. The simplest such model is the random division of a plane by a Poisson field of lines. Although an elaborate theory exists in this simplest, but difficult, case many questions remain unanswered, e.g., the distribution of the area of the resulting polygons. In addition to discovering most of the known theory of this problem and its higher dimensional generalisations, Dr. R.E. Miles has succeeded (in collaboration with Dr. I.K. Crain) in programming a large-scale computer to do simulations of the resulting polygons. This has opened up the exciting prospect of the computer simulation of random geometry.

Models of the coverage of one geometric element by randomly placed elements of another shape lead to the problem of estimating the probability of complete coverage on the one hand, and of the distribution of the measure of the region not covered on the other. The former problem is usually only open to heuristic theoretical discussion but simulations give good agreement with such theory. The latter problem, which sometimes leads to central limit theorems for the measures of

random sets, is analytically complicated and has some interesting applications in statistical mechanics. Here again numerical simulations have been done.

The type of statistical models which lie at the basis of statistical inference and the mathematics of the latter now has an enormous literature. The use of computers makes it easy to calculate solutions of difficult estimation equations and, more importantly, to simulate the behaviour of any estimators for values of the parameter point in the neighbourhood of the guessed value. In the future this is going to affect profoundly the practical techniques of statistical inference.

However we must consider some rather newer and more difficult problems in inference. In any statistical inference problem we need not only to be concerned with the identifiability of the parameters but with what we may call the identifiability of the model. Thus entirely different physical models may lead to the same predicted stochastic behaviour and even if they do not, the difference may be too small for discrimination in a reasonable sample (compare Cox's discussion of discrimination between logit and probit response). This may be put in another way by saying that there is a problem of the robustness of identification of a model.

An important example of this situation is the identifiability of the internal structure of a "black box". Suppose into some system (the "box") there is an observed input $X(t)$ $(-\infty < t < \infty)$, and an observed output $Y(t)$. We take $Y(\tau)$ to be a functional, $F(X(t))$, deterministic or stochastic, of $X(t)$ for $t \leqq \tau$. We seek to determine as much as possible of the structure of $F(X(t))$ from an observed record. This is the situation in studying rainfall runoff, and $F(X(t))$ may be highly nonlinear. Identifiability is then a key question. Simulation may be of great use here if only to test the effectiveness of various methods of analysis.

Simulated models are also important in testing other more well-known statistical techniques, and there is a need for much more work on "challenge statistics" where statisticians who claim to have successful methods of, say, factor analysis or cluster analysis are challenged to apply their methods to sets of artificially constructed data of known structure.

From what has been said above, it is clear that nonlinearity is an all pervading problem and here we are confronted, if not with a brick wall, at any rate with a hill of rapidly increasing slope.

Nearly all work done on time series has been from the point of view of spectral analysis which is a linear theory and is not invariant under nonlinear transformations of the observed values. Thus many questions remain unanswered. No adequate general theory exists for processes in which $X(t)$ is nonnegative but oscillatory (except for special processes in which $X(t)$ is an integer, the number in a population—see for example the theory and simulations of M.S. Bartlett for epidemics and interacting populations). Here again simulation can be used to throw much light on what happens. Of course some work has been done in electrical theory (e.g., by Rice and others) on the spectrum of the output of a nonlinear device fed by a Gaussian stationary process with a known spectrum, but this is rather a special case. Similar problems also occur in econometrics.

Another nonlinear problem in time-series which is very important practically

is to study pairs of processes $X(t)$, $Y(t)$ where $X(t)$ is driven by $Y(t)$ in such a way that large peaks in the spectrum of $X(t)$ are produced at submultiples of the frequencies of large peaks in the spectrum of $Y(t)$. This occurs, for example, in frequency dividers and is an essentially nonlinear phenomenon. An example occurs in testing the reality of the alleged 22-year terrestrial effect of the 11-year sunspot cycle (there is a possible mechanism for this).

Nonlinearity is important in biology, especially in the study of both isolated and interacting animal populations, and in population genetics. The theory of an isolated animal population usually involves both nonlinearity ("density dependence") and lagged effects. The stochastic theory is often only approachable by simulation, whilst the deterministic theory throws up very deep problems on the global behaviour of iterated functions, i.e., discrete time topological dynamics. A beginning at a general theory of this very difficult subject has been made by Smale, Friedlander, Kesten and others. Using computer simulations as an exploratory device it has been found that even the very simplest functional relations lead to quite extraordinarily complicated behaviour.

In population genetics nonlinearity is the rule and has also been systematically studied by Karlin, Kesten and others. Some of the results may be of great importance in the study of common human diseases.

From the above we see that in studying stochastic models repeated recourse has to be made to simulation. The existence of large powerful computers and, equally important, the arrival of the printed circuit programmable desk machine, makes such simulation practicable and often easy. I seem to detect some tendency for pure mathematicians to despise such work. This is a profound mistake and much of the future of both pure and applied mathematics will be the result of heuristic investigations. As an example ,we may mention that L. J. Mordell once studied a certain cyclic algebraic inequality with seven variables. He did not succeed in proving it, but Dr. K. J. Goldberg verified the inequality for 300,000 sets of pseudo-random numbers and after this the theorem was finally proved by P.H. Diananda. This illustrates the use of stochastic modelling in quite pure mathematical problems.

References

1. A. Ya. Khintchine (A. Ja. Hinčin), *Mathematical foundations of statistical mechanics*, OGIZ, Moscow, 1943; English transl., Dover, New York, 1949. MR **8**, 187; **10**, 666.

2. ———, *Analytical foundations of physical statistics*, Trudy Mat. Inst. Steklov. **33** (1950); English transl., Russian Tracts on Advanced Math. and Physics, vol. 2, Gordon and Breach, New York; Hindustan, Delhi, 1961. MR **13**, 416; **26** #7369.

3. S. Kirkpatrick, *Percolation and conduction*, Rev. Modern Phys. **45** (1973), 574–588.

4. N. C. Metropolis and S. Ulam, *The Monte Carlo method*, J. Amer. Statist. Assoc. **44** (1949), 335–341. MR **11**, 138.

5. D. G. Kendall, *Pole seeking Brownian motion and bird navigation*, J. Roy. Statist. Soc. Ser. B **36** (1974), 365–417.

AUSTRALIAN NATIONAL UNIVERSITY
CANBERRA, A.C.T. 2600, AUSTRALIA

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Mathematics and the Picturing of Data*

John W. Tukey

1. Introduction. Why am I writing on this topic? Partly because picturing of data is important. Partly because, if present trends continue, an increasing fraction of all mathematicians will touch—or come close to touching—data during the next few decades. Mathematicians have many advantages in approaching data—and one major disadvantage. Those mathematicians who might come close to data need to know their advantages from their disadvantages.

Experience and facility with clear thinking—and with varied sorts of calculi that lead step-by-step from start to conclusion—knowledge of a variety of mathematical structures—even some of the more abstract are sometimes relevant to data—these are great advantages. The habit of building one technique on another—of assembling procedures like something made of erector-set parts—can be especially useful in dealing with data. So too is looking at the same thing in many ways or many things in the same way; an ability to generalize in profitable ways and a liking for a massive search for order. Mathematicians understand how subtle assumptions can make great differences and are used to trying to trace the paths by which this occurs. The mathematician's great disadvantage in approaching data is his—or her—attitude toward the words "hypothesis" and "hypotheses".

I must diverge for a moment to tell a story, dating to about 1946. The late Walter Mayer, then a member of the School of Mathematics at the Institute for Advanced Study, and I were chatting at the A.M.S. annual meeting at Rutgers. He was surprised that I was going to stay with Bell Laboratories, as well as with Princeton University. He explained how he had become involved with applied matters in Germany during World War I, and how happy he was to get back where, and I

*Prepared in part in connection with research at Princeton University supported by the U.S. A.E.C.

quote, "If I say a g_{ik} has certain properties, it does". If you cannot occasionally modify the attitude Walter Mayer then expressed, work close to data may not be your forte.

When you come to deal with real data, formalized models for its behavior are not *hypotheses* in the mathematician's sense—in the sense that Walter Mayer so enjoyed—the language adopted by classical mathematical statistics *notwithstanding*. Instead these formalized models are *reference situations*—base points, if you like—things against which you compare the data you actually have to see how it differs. There are many challenges to all the skills of mathematicians—*except* implicit trust in hypotheses—in doing just this.

Since no model is to be believed in, no optimization for a single model can offer more than distant guidance. What is needed, and is never more than approximately at hand, is guidance about what to do in a sequence of ever more realistic situations. The analyst of data is lucky if he has some insight into a few terms of this sequence, particularly those not yet mathematized.

2. Picturing in simple cases. Picturing of data is the extreme case. Why do we use pictures? Most crucially to see behavior we had not explicitly anticipated as possible—for what pictures are best at is revealing the unanticipated; crucially, often as a way of making it easier to perceive and understand things that would otherwise be painfully complex. These are the important uses of pictures.

We can, and too often do, use picturing unimportantly, often wastefully, as a way of supporting the feeble in heart in their belief that something we have just found is really true. For this last purpose, when and if important, we usually need to look at a *summary*.

Sometimes we can summarize the data neatly with a few numbers, as when we report:

a fitted line—two numbers,

an estimated spread of the residuals around the "best" line—one more number,

a confidence interval for the slope of the "best" line—two final numbers.

When we can summarize matters this simply in numbers, we hardly need a picture and often lose by going to it. When the simplest useful summary involves many more numbers, a picture can be very helpful. To meet our major commitment of asking what lies beyond, in the example asking "*What is happening beyond what the line describes*?", a picture can be essential.

The NW corner of Figure 1 is a wasteful, unhelpful picture, except for those who must be reassured that the U.S. population increased, roughly exponentially, between 1800 and 1890. The NE corner is an effective, helpful picture, showing how the census population deviated from one exponential. ("Census population," because census errors are comparable with what we now see.) It shows us the comparison between the data and a simple, well-understood reference. The SW corner gives, numerically, the simplest relatively close fit to some data. Who among us can look at this and tell what is going on? The SE corner shows graphically exactly the same fit. All we have to do to use it is to learn to pay no attention to horizontal

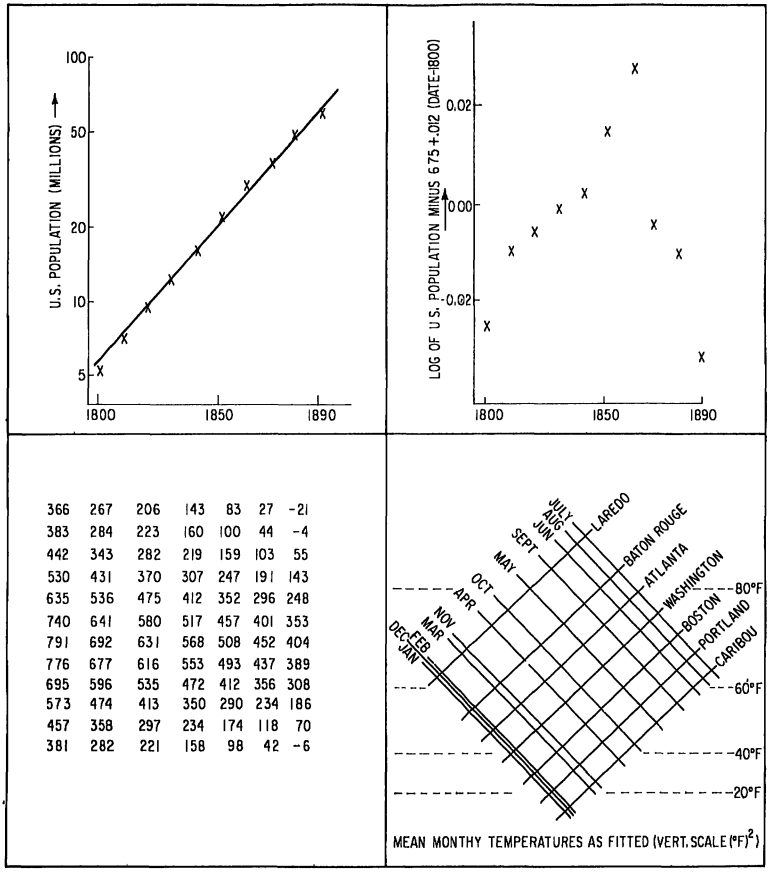


FIGURE 1

position, to forget that coordinate entirely. Such a picture makes the otherwise complicated understandable.

The main tasks of pictures are then:

to reveal the unexpected,

to make the complex easier to perceive.

Either may be effective for that which is important above all: *suggesting the next step in analysis, or offering the next insight.* In doing either of these there is much room for mathematics and novelty.

How do we decide what is a "picture" and what is not? The more we feel that we can "taste, touch, and handle" the more we are dealing with a picture. Whether it looks like a graph, or is a list of a few numbers is not important. Tangibility is important—what we strive for most.

The great geologist Chamberlain once said, in a paper recently reprinted in Science after seven decades: "Science is the holding of multiple working hypotheses". We need to go further, to the broader—prouder—maxim: "The picturing

of data must be sensitive, not only to the multiple hypotheses we hold, but to the many more we have not yet thought of, regard as unlikely or think impossible".

3. Some details. Just using residuals as a way of looking at the data against the straight line—instead of through its appearance—was not difficult. The same idea recurs, however, with ever-increasing complexity. As mathematicians we are used to taking a bit of process from one procedure and putting it in others. We need to do this widely and subtly with

$$\text{given} = \text{fit PLUS residual.}$$

Notice that we have pictured "plus" by writing it out in capital letters—thus giving it its proper emphasis instead of using a single unobtrusive symbol.

Consider a fit in the form row PLUS column. (By not writing $a_i + b_j$ I am picturing the formula—and in the process not leaving its essentials to subscripts and + signs which the nonmathematician finds it easy not to notice.) When we plot the points with coordinates (row, column) (easily found as the intersection of a family of vertical lines with a family of horizontal lines), the loci row + column = constant are parallel straight lines with slope -1, and we have only to turn our picture through 45 degrees. All very simple—but equally mathematical.

We have now made a picture where the vertical coordinate is all meaningful and the horizontal coordinate *is to be forgotten*. We have had many years' experience with graphs where both coordinates have meaning. Many find it suprisingly hard to give up the idea that ALL quantitative pictures have to involve TWO meaningful coordinates.

4. Mental picturing of matrices. The mental pictures of those concerned with data have, of necessity, to be more or less mathematical. It is important that they involve appropriate mathematics, sufficiently understood. Failure to do this is most evident in connection with finite-dimensional linear spaces.

To be useful in dealing with data, understanding of matrices needs to be both abstract and concrete, and the bridge between needs to be well trodden in both directions. An abstract matrix—whether it represents a linear transformation or a quadratic form—need not involve a coordinate system, though it must involve two vector spaces (which may coincide). A numerical matrix—whether it represents a linear transformation or a quadratic form or a change of coordinates, three interpretations that we MUST keep clearly separate, must involve not only two specific vector spaces but two specific coordinate systems (which again may coincide). To hint by saying that a matrix is "$p \times p$", that any two matrices with matching numbers of columns and rows can be freely multiplied together, is an egregious source of false understandings and error.

The linear spaces that arise in treating data are rarely finite-dimensional examples of the familiar self-conjugate Hilbert spaces. Rather they are finite-dimensional Banach spaces. (1, 3, 5) is quite different as the values of (x_1, x_2, x_3) in some given data set than it is as the coefficients $\{c_i\}$ in $c_1x_1 + c_2x_2 + c_3x_3$. Inner pro-

ducts between a vector of c's and a vector of x's are well defined, and do not need assumptions. (Quadratic norms seem always to come from precise assumptions about relative sizes of variances and covariances. If these really mattered crucially, we would almost always be in very deep trouble. Fortunately, only rough ratios matter.)

Let me remind you of the distinction between orthogonality $(x_i, x_j) = 0$, for $i \neq j$ and biorthogonality $(c_i, x_j) = 0$, for $i \neq j$.

Many have heard of orthogonality; some bow down to it as to an idol. Fewer have heard of biorthogonality, yet it is a much more important idea in handling data. The practice of solving linear equations usually involves orthogonalization, yet who teaches that this is an arithmetically convenient route to biorthogonality? (At times, I suspect this is orthogonality's chief—if not only—virtue.)

Somehow ease of writing—or concentration on linear transformations—more precisely on linear transformations from and to the same space—rather than on quadratic forms—has made many find $|A - \lambda I| = 0$ the natural equation for eigenvalues and eigenvectors—which are then said to be "of A". When we deal with data, we are much more often concerned with $|B - \lambda C| = 0$ and we usually need to be concerned with the result as being "of B compared with C".

Some may think of these as small points. They are not tremendous, but their neglect or misinterpretation has kept many from an adequate understanding of what they are doing.

Mathematicians should, I believe, see that their students:

understand that simple matrix operations, like inversions, finding eigenvalues, or finding eigenvectors really exist, and can be used—by a computer-and-programs system when available, by hand in case of need,

understand the simple abstract characteristics of linear spaces, their conjugate spaces, linear transformations, and quadratic forms,

understand what the effect of change of coordinates is on such objects,

understand how to tie the concrete numerical and the abstract algebraic together, when thinking about and working with real examples,

understand that every numerical matrix implies two coordinate systems.

If people are to throw letters for matrices around as freely as they do letters for numbers—a noble goal—they need to be equally willing to throw in numerical values in the two cases—and as willing to introduce matrix coordinate changes in the one as to introduce scalar coordinate changes, perhaps feet to inches, in the other. Matrix algebra needs to have the same reality as scalar algebra.

5. Cumulations. How do we present information about distribution—whether the distribution is a mathematical entity, a Lebesgue measure, or a batch of points? By saying "information about" we are admitting the need to give only a summary. By imbedding this question in a paper whose title includes the word "data" we are admitting that such questions as whether the mathematical distribution is discrete, singular or absolutely continuous are not to be answered.

Before we summarize, the mathematician is likely to want to use the cumulative

(cumulative distribution function), usually defined with a $\leqq$. Both he and the practitioner can gain by the redefinition

$$F(x) = \text{Probability } \{y < x\} + \tfrac{1}{2}\,\text{Probability } \{y = x\}$$

which makes the cumulative of $-y$ exactly $1 - F(x)$, any discontinuities included, and makes Fourier inversion exact at any discontinuities. (Karl Pearson did this without saying so seven decades ago.)

The empirical distribution function is often defined as *1/nth of the Count of y's* $\leqq x$ where n is the total count. Again it is better to choose

$$n \cdot F_n(x) = (\text{Count of } y\text{'s} < x) + \tfrac{1}{2}(\text{Count of } y\text{'s} = x).$$

We will soon learn to change n to $n + \frac{1}{3}$ and add $+\frac{1}{6}$ on the right.

How are we to summarize a function? This question is easier for the empirical case, where we have at most n values to begin with, where it is easy to move back from a higher-level description by a function to a lower-level description by n numbers, much more tangible as long as n is not too large.

If we draw a sample $y_1, y_2, \cdots, y_n$, of n observations from $dF(y)$ we are usually only concerned with symmetric functions of the y's. (Sometimes I think that only statisticians work with nonpolynomial symmetric functions.) The numerically least of $y_1, y_2, \cdots, y_n$ which we will label $y_{1|n}$ is surely a symmetric function of the y's. (Rearrangement surely makes no difference to its value.) The same is true of $y_{2|n}$, the next-to-smallest, and so on. The order statistics $y_{1|n} \leqq y_{2|n} \leqq \cdots \leqq y_{n|n}$ obtained by rearranging the y's in increasing order are actually the most general symmetric functions of $y_1, y_2, \cdots, y_n$.

Assume now that $y_i, \cdots, y_n$ are a random sample from a distribution with a continuous cumulative $F(x)$. This is a reference situation, not a real one, but that should not interfere with using it for guidance.

It is easy to see that the distribution of $F(y_{i|n})$ depends only on n and i, not at all on F. There are many ways to typify the location of each of these distributions with a number. The one that is most useful is by its median, especially since the operation of taking medians commutes with monotone re-expressions. Using the available tables—either of the beta distribution or of Snedecor's F distribution—we can learn that

$$\text{median } \{F(y_{i|n})\} \lessapprox \frac{i - \frac{1}{3}}{n + \frac{1}{3}}$$

where "$<$" means "is less than" and "$\approx$" means "is close to". Indeed for $n \geqq 5$ and all i, the reversed inequality holds when $(i - 0.3)/(n + 0.4)$ replaces $(i - \frac{1}{3})/(n + \frac{1}{3})$.

By the commutativity noted above, $F(\text{median}\{y_{i|n}\})$ satisfies the same condition, so that

$$\text{median } \{y_{i|n}\} \lessapprox F^{-1}\left[\frac{i - \frac{1}{3}}{n + \frac{1}{3}}\right].$$

Thus the natural fraction to associate with the ith of n is neither of the traditional

choices—neither $(i - \frac{1}{2})/n$ nor $i/(n + 1)$—but rather $(i - \frac{1}{3})/(n + \frac{1}{3})$.

What does this tell us about empirical cumulatives? Almost certainly that, if we must use them,

$$F_n(x) = \frac{(\text{Count of } y\text{'s} < x) + \frac{1}{2}(\text{Count of } y\text{'s} = x) + \frac{1}{6}}{(\text{Total count of } y\text{'s}) + \frac{1}{3}}$$

is a still better choice. At the ith order statistic, the numerator is $(i - 1) + (1/2) + (1/6) = i - (1/3)$. So much for getting our ideas and definitions a little clearer.

6. Summarizing. The n order statistics are a natural beginning for summarization. We want to pick out some of them to stand for the rest. We can do this fairly well, because we know roughly what their correlation structure is: roughly equal correlations for equal values of i_1/i_2.

If we start with $i = 1$, and double, getting, successively, $i = 2, 4, 8, 16, \cdots$ or go up by half-octaves, getting, after rounding to integers, $i = 1$, blank, 2, 3, 4, 6, 8, $11, \cdots$ we will have sequences of roughly equally correlated order statistics. We can use such sequences, from above and from below, quite effectively as summaries.

This works well for some purposes. For others, where we would like to make use of the approximate stability of $(i - \frac{1}{3})/(n + \frac{1}{3})$ as n changes, we do better to begin in the middle and make our way toward both ends using:

depth of $y_{i|n}$ = the lesser of i or $n + 1 - i$ = result of counting in from the nearest end,

depth of median = $\frac{1}{2}$ (1 + total count),

depth of hinge = $\frac{1}{2}$ (1 + depth* of median),

depth of eighth = $\frac{1}{2}$ (1 + depth* of hinge),

...

where we interpret depth* as "the integer part of the depth of" whenever, as is usual, we wish to confine i to integers or half-integers. (Otherwise we might solve $(i - \frac{1}{3})/(n + \frac{1}{3}) = 2^{-i}$ getting $i = (1/3) + (2^{-i}/3) + n/2^i$, to which the above is a reasonable approximation and which is more delicately precise than data warrants.)

As a result, we summarize our distribution information with a sequence of about $2 \log_2 n$ values (about $2 \log_2 n + 2$ if we go all the way, including depths $1h$ and 1).

We have come a long way toward tangibility in going from $F_n(x)$ to $2 \log_2 n$ selected order statistics but this is only part way. We have not yet brought in a reference situation, so we are not in the favorable position of explicitly comparing what we have with a standard.

One approach to bringing in the Gaussian reference will be noted here, another later. If, for mathematicians' use only, we label the selected order statistics $L_k, \cdots, L_3, L_2, M, U_2, U_3, \cdots, U_k$ where L_j and U_j share a depth, we can replace each pair by the corresponding "mid" and "spread"

$$M_j = (L_j + U_j)/2 \quad \text{and} \quad S_j = U_j - L_j$$

and if we then divide the spreads by the values of the corresponding spreads for the unit Gaussian distribution, we obtain two sequences of numbers such that the

median values would be μ for each mid and σ for each divided spread *provided* we were sampling from a Gaussian distribution with mean μ and variance σ^2. Further investigation then suggests that we plot both sequences against the square of the divisor used for the corresponding spread. Thus, when we want to separate the issues of skewness from those of tail elongation, rather than separating what happens in the lower tail from what happens in the upper tail, we can reach a respectable pair of pictures comparing any batch of n observations with a Gaussian standard. (We can convert very easily to a logistic standard, or a Cauchy standard, or the standard provided by any other symmetric distribution. If a seriously asymmetric distribution is a natural standard, the separation into mids and spreads is likely to be unnatural.)

7. And what of two dimensions? How can we generalize all this to the plane? Specifically to the affine plane? Direct generalizations of order statistics fail miserably. But if we regard an order statistic $y_{i|n}$ as an oriented (up or down across the real line) point, with $\geqq i$ of the n values to its "left" or on it and $\leqq i - 1$ strictly

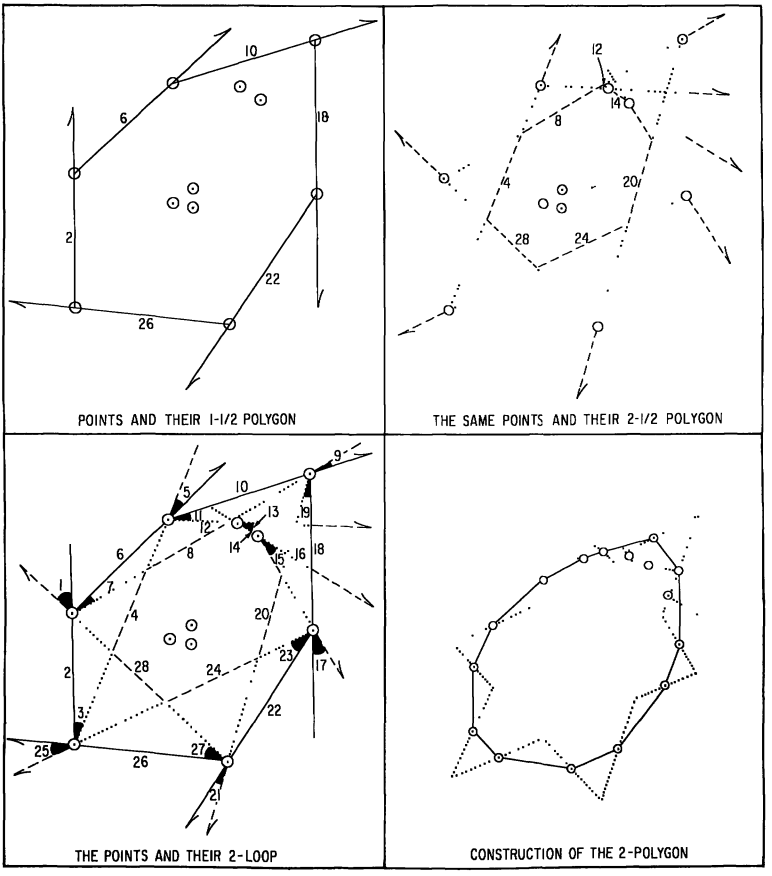


FIGURE 2

to its "left", we can generalize easily. In the plane, an (i, j) line will be any directed line with $\geqq i$ points to its left or on it and $\leqq j$ points strictly to its left. For any (i, j) the set of (i, j) lines is closed. For any $i < n$, there is one and only one $(i, i - 1)$ line in a given direction, one line of depth i. Thus the $(i, i - 1)$ lines form a closed curve of lines of depth i, the i-loop. If $j < i - 1$, the set of (i, j) lines is finite. The set of $(i, i - 2)$ lines forms a closed polygon, the $(i - \frac{1}{2})$-polygon (all its sides belong to both the $(i - 1)$-loop and the i-loop.

The NW corner of Figure 2 shows 11 points and the $1\frac{1}{2}$ polygon; the NE corner the same points and the $2\frac{1}{2}$ polygon; the SW corner the two polygons and the sectors defining the segments of pencils of lines that complete the 2-loop.

(Each filled-in sector represents all directed lines through the vertex which pass out through the sector. If a segment of a pencil is a "stub", is each such loop a "polystub"?) Clearly the 2-loop is too complex to be a satisfactory generalization of a pair of order statistics of matched depth. The midpoints of the segments cut off by the $(i - \frac{1}{2})$-polygon from the extensions of the sides of the $(i + \frac{1}{2})$-polygon define a new, intermediate polygon, the i-polygon, which seems to be a satisfactory generalization. The SE corner of Figure 2 shows the 2-polygon for the 11-point example.

8. Probability plots. A naive answer to "How should we picture a distribution based on n values?"—essentially the question asked in §6—would be to say "make a probability plot". (There are many weaknesses of such plots.)

Today, Luis Nanni and I are hard at work developing a plot that meets these difficulties as well as we can see how to do this. Why did it take so many decades to attack this problem? Presumably because too few of us have tried to make more useful pictures.

PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08540, U.S.A.

BELL LABORATORIES
MURRAY HILL, NEW JERSEY 07974, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Levels of Structure in Catastrophe Theory Illustrated by Applications in the Social and Biological Sciences

E. C. Zeeman

Catastrophe theory is a method discovered by Thom [14] of using singularities of smooth maps to model nature.

In such models there are often several levels of structure, just as in a geometry problem there can be several levels of structure, for instance the topological, differential, algebraic, and affine, etc. And, just as in geometry the topological level is generally the deepest and may impose limitations upon the higher levels, so in applied mathematics, if there is a catastrophe level, then it is generally the deepest and likely to impose limitations upon any higher levels, such as the differential equations involved, the asymptotic behaviour, etc. Again, in geometry the complexity of the higher levels may render them inaccessible, so that they can only be handled implicitly rather than explicitly, while at the same time the underlying topological invariants may even be computable. Similarly in applied mathematics the complexity of the differential equations may sometimes render them inaccessible (even to computers), so that they can only be handled implicitly rather than explicitly, while the underlying catastrophe can be modeled, possibly even to the extent of providing quantitative prediction.

Therefore catastrophe theory offers two attractions: On the one hand it sometimes provides the deepest level of insight and lends a simplicity of understanding. On the other hand, in very complex systems such as occur in biology and the social sciences, it can sometimes provide a model where none was previously thought possible. In this paper we discuss various levels of structure that can be superimposed upon an underlying catastrophe and illustrate them with an assortment of examples. For convenience we shall mostly use the familiar cusp catastrophe (see [5], [13], [14], [24]).

Level 1. Singularities.
Level 2. Fast dynamic (homeostasis).
Level 3. Slow dynamic (development).
Level 4. Feedback.
Level 5. Noise.
Level 6. Diffusion.

Thom's classification of elementary catastrophes belongs to Level 1. Levels 2,3,4 refer to ordinary differential equations, and Level 6 refers to partial differential equations.

Level 1. Singulairties. We begin by recalling the main classification theorem. Let C, X be manifolds with dim $C \leqq 5$, and let $f \in C^\infty(C \times X)$. Suppose that f is generic in the sense that the related map $C \to C^\infty(X)$ is transverse to the orbits of the group Diff $(X) \times$ Diff (R) acting on $C^\infty(X)$. (Genericity is open-dense in the Whitney C^∞-topology.) Let $M \subset C \times X$ be given by $\nabla_X f = 0$, and let $\chi : M \to C$ be induced by projection $C \times X \to C$.

THEOREM (THOM). (a) *M is a manifold of the same dimension as C.*
(b) *Any singularity of χ is equivalent to an elementary catastrophe.*
(c) *χ is stable under small perturbations of f.*

The number of elementary catastrophes depends only upon the dimension of C (and not on X):

dim C	1	2	3	4	5	6
elememtary catastrophes	1	2	5	7	11	(∞)

For details of the elementary catastrophes the reader is referred to **[10,]** **[14]**. The first complete proof was given by Mather **[8]**, and other references are **[1]**, **[17]**, **[18]**.

REMARK. The classification of singularities goes infinite for dim $C \geqq 6$, but the above table can be extended with finite entries, provided the concept of elementary catastrophes is suitably modified, as follows. The singularities correspond to orbits of a group acting on a space of germs (see Level 4 below). In particular the (∞) appears at dim $C = 6$ because there is a stratum of codimension 6 foliated by orbits of codimension 7. Arnold **[1]** calls the codimension of the foliation, which is 1 in this case, the modality of the stratum. More generally the orbits form a foliated stratification P, which Arnold has shown to be locally finite. The finite numbers of strata of each codimension give the desired extension of the above table.

The reason that catastrophe theory exists is that by a happy accident P is 5-simple, in other words each stratum of P of codim $\leqq 5$ is simple, that is trivially foliated by a single leaf. These strata correspond to the elementary catastrophes of dim $\leqq 5$, and hence the latter are finitely classified differential invariants. For most applications it suffices to have dim $C \leqq 5$, and so there is no need to worry about the foliation of the higher strata.

Application. Suppose we have some set of objects or events about which we want to test a hypothesis of cause and effect. One of the first things to do is to plot them

in cause-effect space, and see if they form a graph. Here C will describe the cause, X the effect, and $f(c, x)$ the probability that cause c will produce effect x. The most likely effects are given by the peaks of probability, where both the gradient vanishes, $\nabla_X f = 0$, and the Hessian is negative definite, $\nabla_X^2 f < 0$. This determines a submanifold G of M (of the same) dimension. Then G will be the desired cause-effect graph in $C \times X$. The events will be represented by a cloud of points clustering near G, with density of clustering depending upon the deviation of the probability distributions.

Consider the first two elementary catastrophes, which occur when dim $C = 2$. The fold-catastrophe occurs at the boundary of G, but since there is no dynamic in Level 1 there may not be any catastrophic jump here—all we can say is that the cloud of points appears to terminate.

The cusp-catastrophe occurs when a probability distribution goes bimodal. In this case observers may implicitly recognise the phenomenon, and capture part of it by either naming the two modes, or alternatively framing some form of words, such as a proverb or a belief. However the cusp-catastrophe can often reveal other facets to the phenomenon, and give a new synthesis of understanding. We illustrate the two alternatives by a couple of examples.

EXAMPLE 1. AGGRESSION [22]. According to Konrad Lorenz [7] fear and rage are conflicting drives influencing aggression. Here the two extreme behaviour modes are attack/flight, and X represents a 1-dimensional spectrum of behaviour varying from neutral to the two extremes at either end. The cause C is 2-dimensional, representing the strengths of the fear and rage drives present in the animal at that moment. Lorenz observes that in the case of dogs the coordinates of fear and rage can be read from the facial expression [7, p. 81]. Rage only causes attack, fear only causes flight, and when both are present the effect is one of two extremes but unpredictable. Therefore the probability goes bimodal, and as a first approximation we might expect our cloud of points to cluster around a cause-effect graph equivalent to the cusp-catastrophe as shown in Figure 1. We shall return to this example, and its uses, in Level 2.

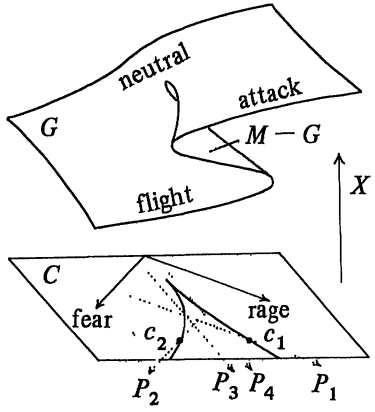


FIGURE 1

Other familiar examples of bimodality which can be modeled by the cusp-catastrophe are (i) liquid/gas [3], [11], [15], (ii) diastole/systole [23], (iii) manic/depressive [25], (iv) dove/hawk [5], or (v) bull/bear ([26], and Example 7 below).

In each case the bimodality is caused either by conflicting factors such as temperature and pressure in (i), or by a splitting factor such as tension in (ii), disease in (iii), cost in (iv) or speculation in (v). Let us now give an example of a proverb.

EXAMPLE 2. MORE HASTE LESS SPEED. This proverb is very familiar in England, although almost unknown in America. Its attraction lies in its brevity and contrariness—it is the opposite of what one would normally expect, especially if the operator is skilled at his task. And this leads to the observation that speed really depends upon two factors, haste and skill, which are conflicting. For, when both factors are present, the probability goes bimodal, because either the operator's skill enables him to increase his speed, or his fumbling haste diminishes it. Therefore again we might expect our cloud of points to cluster around a cusp-catastrophe, as in Figure 2.

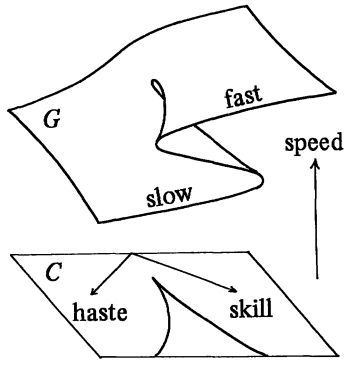


FIGURE 2

We suggest a couple of uses for such a model. Firstly in psychology it might be possible to develop it further into a predictive quantitative model for explicit skills (see Level 2 below). Secondly in sociology it might provide a prototype for reconciling conflicting theories. For, by the theorem, we should expect the cusp-catastrophe to occur in many phenomena, and, although the graph is visually simple, its subtlety is not easy to describe with any brevity in ordinary spoken or written language (see [5, §10]). Therefore although we may often recognise such a phenomenon intuitively, we tend to describe it verbally by an oversimplification, possibly by directing attention only to the unexpected mode. For instance "more haste less speed" directs attention only to the lower front sheet of Figure 2, marked "slow". Similarly two conflicting opinions in a discussion, or two conflicting sociological theories, may in fact each be referring to a single mode of an underlying bimodal phenomenon, and the conflict may sometimes be reconciled by exhibiting the two

modes as the two sheets of a cusp-catastrophe growing smoothly out of an area of common agreement.

Level 2. Dynamic. In addition to the generic function $f \in C^\infty(C \times X)$ suppose we are given a dynamic D as follows. Denote the associated function $C \to C^\infty(X)$ by $c \mapsto f_c$. Then $D = \{D_c\}$ is a family of differential equations on X, parametrised by C, such that, for each $c \in C$, f_c is a Lyapunov function for D_c. In other words, f_c increases (or decreases) along the orbits of D_c, and so the maxima (or minima) of f_c are the attractors of D_c. Therefore D_c is gradient-like, and this is the restriction that Level 1 imposes on Level 2.

The graph G now represents the attractors of D. In applications we no longer intuitively imagine a cloud of points clustering statically near G, but points flowing dynamically onto G and then staying there. The model is half dynamic and half static. It is sometimes useful to think of the parameter space C as control, and X as behaviour space. If we slowly move the control c then the behaviour x responds by moving continuously on G for as long as possible; in other words it is a theorem of Level 2 that the system obeys the delay rule of Thom [5], [14].

If c now crosses the bifurcation set, then x may cross the boundary ∂G of G. In this case the dynamic will carry x rapidly onto some other sheet of G. The word "rapidly" assumes that the movement of control is slow compared with the dynamic, and it is the sudden jump that occurs at the fold-catastrophes in Level 2 that is responsible for the name "catastrophe theory".

EXAMPLE 3. The *catastrophe machine* described in [9], [24] is a simple toy made out of a cardboard disk and two elastic bands, which exhibits the catastrophic jump well, and the uninitiated reader is recommended to make one for himself. Here the function f of Level 1 is the potential energy in the elastic given by Hooke's law of elasticity, and the dynamic D is given by Newton's law of motion, suitably damped so as to minimise f.

EXAMPLE 1. Returning to our first example we see that it can be promoted from Level 1 to Level 2. For we may reinterpret X as the space of states of that part of the brain governing mood (perhaps the hypothalamus), and D as the associated dynamic representing neurological activity. Then the attractors of D represent the attacking/retreating frames of mind, providing the background mood against which behavioural decisions are taken. Although X must necessarily be very high dimensional, and D consequently inaccessible in the sense of being only implicit, nevertheless G will still be 2-dimensional. Therefore the cusp-catastrophe can still provide an explicit model, which for individual animals might be made quantitative and predictive. Moreover since it is a Level 2 model, even though D is only implicit, there will be catastrophic jumps of mood, resulting in sudden attacks or disengagements. For example in Figure 1 the path P_1, representing increasing rage at a fixed level of fear, as for instance in a cornered dog, will lead to a sudden attack at c_1 while path P_2 to a sudden disengagement at c_2. Meanwhile paths P_3, P_4 illustrate how nearby paths can lead to divergent behaviour. Similarly humans, when made angry and frightened, are unpredicatable and are denied access to

rational behaviour, and may jump from abuse to apology, even from hysteria to tears.

The interest of this example is that it may provide a general model for control of aggression, valid for different species under varying circumstances, and may give insight into how such controls develop and have evolved. More generally it provides a prototype for relating the neurology to the psychology of moods underlying behaviour.

EXAMPLE 2. Our second example may also be promoted from Level 1 to Level 2, because, if we consider the performance of an individual, his tendency to adjust his speed to x, say, within the limitations of his skill and assuming a given amount of haste, is another way of saying there is an implicit dynamic that moves the speed to x.

A path P_1, such as in Figure 1, here represents an increasing skill at a fixed level of haste, as for instance when learning to ride a bicycle, and at the point c_1 a catastrophe occurs when the individual is suddenly able to ride. Moreover the greater the haste—for instance the swifter reactions that are needed to ride a more unstable machine—then the greater the skill needed before the catastrophe occurs. Meanwhile a path P_2 here represents increasing haste at a fixed skill, as for instance a wireless operator trying to read faster and faster Morse code, and at the point c_2 a catastrophe occurs as the performance drops sharply. Moreover the greater the skill, the greater the haste possible before the catastrophe occurs.

In general Level 2 is much easier to test experimentally than Level 1, because the cloud of points more accurately determines G, and the catastrophes determine ∂G. Whenever a phenomenon exhibits any one of the four qualities of bimodality, divergence, catastrophic jumps or hysterisis delays, then it may be possible to model it by the cusp-catastrophe, in which case it may be possible to predict the other three qualities. Sometimes the cusp-catastrophe can also be useful in applications where the control space C is high dimensional, as shown by the following example.

EXAMPLE 4. ECOMOMIC GROWTH. Let X represent the space of states of an economy, and C the external pressures on that economy together with the controls available to the government. Let D represent the implicit response of the economy. We should expect C to be high dimensional, and so at first sight the theorem is of little use. However the evolution of the economy is in fact only a 1-dimensional path in C lifted to a 1-dimensional path in G, and the corresponding 1-dimensional catastrophes are the slumps, inflation explosions, etc.

A typical problem facing the government is the realisation that whereas its present policy is now at control point c_0, it may have to change policy in the next few months to c_1, due to external pressures, balance of payments, etc. The government's freedom of action may be limited merely to choosing the path from c_0 to c_1. However such choice may be critical as we now explain. Suppose there is a choice between two paths P_1 or P_2. For simplicity let us assume that neither path involves a catastrophe. The question that must be asked is: Does the circle $P_1 \cup P_2$ link any codimension-2 stratum Σ of the bifurcation set? For if it does, then a

2-dimensional disk E spanning $P_1 \cup P_2$ will pierce Σ, and the section of G over E will contain a cusp-catastrophe, as shown in Figure 3. The lifts Q_1, Q_2 of P_1, P_2 will exhibit divergence, which could radically affect growth, inflation, unemployment, etc.

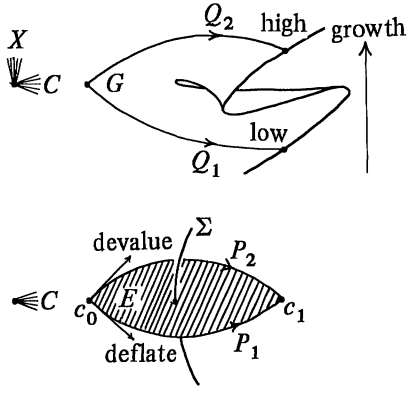


FIGURE 3

For example suppose P_1 represented deflation followed by devaluation (as in the U.K. in 1967), and P_2 the reverse order (as in France in 1968). Then Q_1 could lead to low growth because, with reduced stocks, firms would be unable to exploit the devaluation, whereas Q_2 could lead to high growth, because firms could switch sales of stock from the curtailed home market to the export market, without losing growth momentum. Therefore economists should be concerned not only with the more obvious codimension-1 problems of catastrophe, but also with the more hidden codimension-2 problems of divergence and choice.

Level 3. Development. In addition to $f \in C^\infty(C \times X)$ and the dynamic D suppose that we have time T occurring as one of the axes in the control space C. It is assumed that T is slow compared with the fast time occurring in the dynamic D.

EXAMPLE 5. EMBRYOLOGY. Level 3 occurs in Thom's main application of catastrophe theory to embryology [13], [14], [16], where the control space, $C = S \times T$, represents space-time, and X represents the states of a cell. For instance X may be a bounded open subset of R^n, with several thousand coordinates representing various chemical and physical parameters of the cell. The dynamic D represents the homeostasis of a cell returning it swiftly to equilibrium, and T the slow development of the cells.

An example of a result in this context is the following:

THEOREM [27]. *Whenever a tissue differentiates into two types, the frontier between them first forms to one side and then moves through the tissue before stabilising in its final position.*

The proof uses the cusp catastrophe as illustrated in Figure 4. S is taken to be 1-dimensional perpendicular to the frontier. Development paths of cells are given

by lifting time lines to G. The frontier first forms at c_1, then moves as a wave through S along the cusp branch c_1c_2, and then stabilises at c_2, where the cusp touches the time line c_2c_3. Such a wave is often a hidden switching on of genes, and morphogenesis may be caused after some delay by a secondary wave of physical manifestation. For example in [27] detailed models are given for the morphogenesis of gastrulation and neurulation in amphibia, and of culmination in slime-mold.

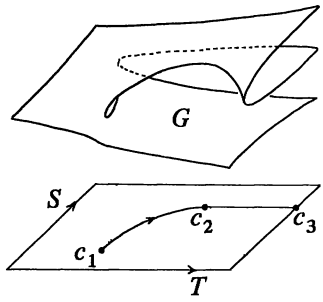


FIGURE 4

Space-catastrophes. The above result depended upon the time-axis not being tangential to the cusp-axis, which can be justified by an appeal to genericity. However to put this type of genericity on a mathematical footing requires a generalisation of the classical theory as follows, which Wassermann [19] calls space-catastrophe theory. (He also studies the dual concept of time-catastrophe theory.)

Let E_n denote the ring of germs at 0 of C^∞-functions $R^n \to R$, m_n the maximal ideal, and G_n the group of germs of C^∞-diffeomorphisms $R^n, 0 \to R^n, 0$. Then G_n acts on m_n, leaving m_n^2 invariant. Classical catastrophe theory [1], [8], [14], [17], [18] consists of analysing the foliated stratification P of m_n^2 by G_n. The elementary catastrophes of dimension s are given by the strata of P of codimension s. Since P is 5-simple, the elementary catastrophes for $s \leqq 5$ are finitely classified differential invariants, independent of n (for $n \geqq 2$).

For the generalisation we need some more definitions. We say $\alpha \in G_{n+r}$ *covers* $\beta \in G_r$ if $\pi\alpha = \beta\pi$, where $\pi : R^{n+r} \to R^r$ is the projection. Define

$$G_n^r = \{(\alpha, \beta) \in G_{n+r} \times G_{1+r};\ \exists\, \gamma \in G_r \text{ such that } \alpha, \beta \text{ cover } \gamma\}.$$

Then G_n^r acts on m_{n+r}, leaving $m_n^2 + m_rE_{n+r}$ invariant. For space-catastrophe theory we choose $r = 1$ (representing time), and analyse the foliated stratification Q of $m_n^2 + m_1E_{n+1}$ by G_n^1. The space-catastrophes of dimension s are given by the strata of Q of codimension $s + 1$.

Wassermann [19] has shown that Q is 2-simple, and hence the 1-space-catastrophes are finitely classified differential invariants, independent of $n \geqq 1$. There are exactly four, namely the beginning c_1, the middle, and the end c_2, of the wave in Figure 4, and the "silent" dual of c_2. Therefore the above theorem is valid, and exhibits them all.

However Q is not 3-simple, for Wassermann has shown that the P-strata of

swallowtails and umbilics are not only substratified by Q but also foliated. Therefore the number of singularities in 2-space goes infinite, and although the 2-space-catastrophes will still be finitely classifiable, they will no longer be differential invariants. Some will be—for example the cusp-projection of a fold-surface into 2-space (analogous to the fold-projection of a fold-curve into 1-space at c_2 in Figure 4). This example has been used to model the pattern formation of somites in amphibia [27].

However Thom [14], [16] uses the swallowtail, butterfly and umbilics extensively in embryology, and hence it is important to classify the 2- and 3-space-catastrophes. Therefore mathematically we need to analyse the strata of Q up to codimension 4, and to understand the nature of the loss in differentiability implied by their foliation.

Level 4. Feedback. Here we assume that the slow flow is not as simple as merely taking a coordinate in the control space, but may go in different directions on different sheets of G. In fact it may be conceived as a form of feedback:

$$C \underset{\text{slow feedback } F}{\overset{\text{fast dynamic } D}{\rightleftarrows}} X.$$

More precisely, in addition to f and D, suppose we are given a C^∞-map F: $C \times X \to TC$, where TC denotes the tangent bundle of C, and $F(c, x)$ is a tangent at c, for each $c \in C$, $x \in X$. Therefore D and F together form an ordinary differential equation on $C \times X$ (with the proviso that D is fast and F slow).

EXAMPLE 6. HEARTBEAT AND NERVE IMPULSE [23]. Explicit examples of differential equations in form of feedbacks on the cusp-catastrophe were taken as models. In each case the flow possessed a stable equilibrium, which if suitably disturbed by an "external agent", triggered a catastrophe via D, and a return to equilibrium via F. In the heartbeat the return involved a second catastrophe (relaxation after contraction), whereas in the nerve impulse the return was smooth (repolarisation). These models possess two interesting features. Firstly the feedback does not give a flow precisely on G, but only near G, the order of nearness depending upon the ratio K of fast/slow. If $K \to \infty$ we obtain an idealised flow on G, with instantaneous catastrophes, generalising the relaxation oscillations of electrical engineering. Secondly the words "external agent" above reveal the inadequacies of the models, in being only ordinary differential equations describing the local behaviour of heart muscle and nerve fragment; what is needed is to embed the latter in a larger partial differential equation that describes the global behaviour as waves. We return to this problem in Level 6.

EXAMPLE 7. STOCK EXCHANGES [26]. The cusp-catastrophe is used to model the behaviour of stock exchanges, as follows. The excess demand is the normal factor controlling the rate of change of index, and the speculative content of the market a splitting factor. The dynamic D represents the immediate response of index to investors, and F the somewhat slower feedback. Plausible economic hypotheses lead to a flow that exhibits periodic bull market, recession, bear market and re-

covery. However, to make this model realistic, we should promote it to Level 5 by including noise.

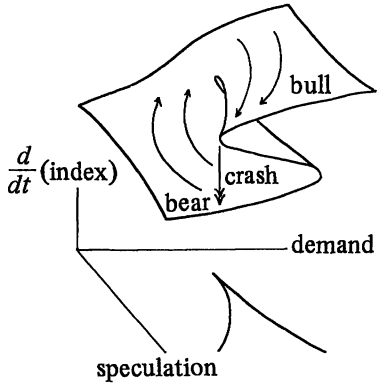


FIGURE 5

EXAMPLE 8. FUNNEL. In classifying the generic low-dimensional feedback-catastrophes, Takens [12] has recently discovered an interesting new type, the simplest of which he calls the funnel. In the associated idealised flow a 2-dimensional piece of G is funneled through a single fold-point P. Figure 6 illustrates the following explicit example:

Fast dynamic D: $\dot{x} = -K(x^2 + 2b)$, K large constant.
Slow feedback F: $\dot{a} = 1$, $\dot{b} = 3a + 4x$.

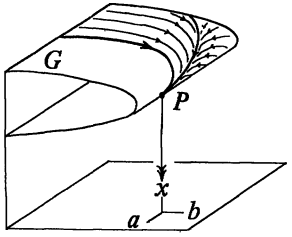


FIGURE 6

Funnels may occur in biological regulation, for instance choosing x, b to model the internal self-regulation of a cell, and a the production by the cell of some hormone for use outside the cell, whose production-rate needs to be funneled precisely.

Level 5. Noise. We may superimpose on $\{f, D, F\}$ stochastic noise in the form of random small displacements of control and behaviour. For most noise the dynamic D carries the state rapidly back onto G, and the slow flow F proceeds as before, and so the noise can be ignored. However in two cases noise can cause catastrophes, firstly if control-noise crosses the bifurcation set, and secondly more interestingly if the behaviour-noise crosses a separatrix.

EXAMPLE 7. In the stock exchange example noise represents external events and consequent jumpiness in the market, and may cause recessions to occur before the bifurcation set is reached.

EXAMPLE 9. RIOTS [4]. This model reports joint work in progress with prison psychologists P. Shapland, C. Hall and H. Marriage, and statistician J. Harrison. We start with a truism: The more tension in an institution the more disorder. This applies not only to institutions such as prisons, universities, firms, or countries but also to individuals. In the case of prisons, an analysis of data suggests that the tension (or distress or frustration) can be measured by the numbers reporting sick, suitably smoothed, and the disorder can be measured by correlating independent assessments of the seriousness of incidents. Alienation (or lack of communication) seems to be a splitting factor, producing the two modes that we have labelled quiet and disturbed in Figure 7, and the data suggest that this may be measured by the numbers of disciplinary reports. The feedback flow represents the increase in tension during quiet (over months) and the release during disturbance (over days). Noise describes incidents, and if the noise level crosses the separatrix AA' at B then the incidents will escalate and spark a riot causing a catastrophe. Some types of prison population (e.g., young long-term) have a higher noise level, and are therefore more susceptible to riots. When the tension has subsided after a few days an incident may cause the reverse catastrophe at B'. The same incident might not have done so earlier, which explains the advantage of playing it cool.

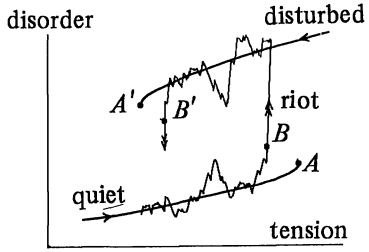


FIGURE 7

EXAMPLE 10. PHASE TRANSITION [3], [11], [15]. If the noise is frequent, and the noise-level high, the state will, averaged over time, seek the absolute maximum (or minimum) of f. This explains why Van der Waals' equation for liquid/gas phase transition has to be supplemented with Maxwell's rule [3], [5], [14], instead of obeying the delay rule. On the other hand, if the noise level is kept low then partial delays can be induced, such as in the supersaturated and superevaporated states of the cloud and bubble chambers. The usual proof of Maxwell's rule in statistical mechanics involves integration by steepest descent, but since this method breaks down near the critical point, it would be interesting if a new abstract proof could be devised, parallel to the proof of Thom's theorem, in order to enhance critical point analysis.

Level 6. Diffusion. The following arises out of joint work in progress with Sharon

Hintze, stimulated by papers of Winfree [20], [21] and Kopell and Howard [6] on the Zhabotinsky reaction. First the mathematics.

Let Y be a manifold and g be a C^∞-vector field on Y. The associated ordinary differential equation is

$$\dot{y} = g(y). \tag{1}$$

In particular we shall be concerned with the type of differential equation given by Level 4, namely $Y = C \times X$ and $g = \{D, F\}$. Suppose now that Y represents the space of local states of some medium in space-time $S \times T$, and that g represents the reaction of that medium. Suppose further that the medium not only reacts but also diffuses. Then, following [6], the global state $y: S \times T \to Y$ of the medium satisfies the reaction-diffusion partial differential equation:

$$\partial y/\partial t = g(y) + k\nabla^2 y \tag{2}$$

where k is a constant (more precisely a vector bundle map $k: TY \to TY$) representing the different rates of diffusion of the various components of Y. We are particularly interested in whether or not the medium can sustain stable periodic wave trains, or stable pulses (isolated waves). If it can, and θ is the speed, then the global state y can be factored $S \times T \to R \to Y$ such that $\partial y/\partial t = \theta \dot{y}$ and $\nabla^2 y = \ddot{y}$, where the dot denotes differentiation with respect to R. Therefore the partial differential equation (2) reduces to the ordinary differential equation

$$\theta \dot{y} = g(y) + k\ddot{y}. \tag{3}$$

This equation (3) is the central interest of Level 6; compare it with equation (1) above. If k is small then (3) can be regarded as a singular perturbation of (1), but in important applications k is large, and so new methods are needed.

For instance (3) can be regarded as a flow on TY, with the same fixed points as (1), on the zero section Y. An attractor of (1) may be a saddle point of (3), and a homoclinic orbit of this saddle will represent a pulse solution of (2). Meanwhile a closed orbit of (3) represents a wave train solution of (2). Therefore we seek homoclinic and closed orbits of (3), that are stable with respect to (2). As yet relatively little is known, even when g represents a canonical elementary catastrophe with the simplest form of feedback.

EXAMPLE 6. In the heartbeat and nerve impulse dynamics [23] Conley and Carpenter [2] have shown the existence of homoclinic and closed orbits, and the next problems are to prove stability and fit data.

EXAMPLE 11. ZHABOTINSKY REAGENTS. Belousov discovered a mixture of chemicals that oscillates in colour at about twice a minute, and later Zhabotinsky and Zaiken observed that circular wave trains would propagate through this reagent, entraining the oscillation. Winfree [21] then modified Belousov's reagent by adding a little more bromide and a little less acid, so as to stop the oscillation. He called his mixture the Z-reagent, after Zhabotinsky and Zaiken, and showed that it could sustain both pulses and rotating scroll-shaped wave trains. In [21] Winfree offers equations which beautifully explain the geometry of the patterns, but which can be mildly criticised on four counts. Firstly his dynamic is discontinuous, and the

obvious way to make the model differentiable is to approximate it by a catastrophe model. Indeed as Kopell and Howard [6] point out there are both fast (fractions of a second) and slow (minutes) reactions, as well as a very slow (hours) loss of energy. Therefore one would normally expect the reaction dynamic to belong to Level 4. Secondly Winfree's equations do not illustrate the modification Belousov $\rightarrow Z$. However this can be illustrated naturally in catastrophe theory, by modifying one constant, causing a Hopf bifurcation, as we show below. Thirdly his equations exhibit a jump return, like the heartbeat, whereas his photographs illustrate a smooth return, blue $\rightarrow$ red, as opposed to the catastrophic hard edge, red $\rightarrow$ blue, more like the repolarisation of the nerve impulse. This feature can be accommodated by using the cusp-catastrophe [20], [23]. Fourthly he does not offer mathematical proof of existence and stability.

As Winfree has pointed out [20], the first two criticisms are answered by a 2-dimensional fold-catastrophe model as follows (cf. [23, Figures 7, 9]). Let $Y = R^2$, and let g be given by

$$D, \text{ fast dynamic:} \quad \dot{x} = -(x^3 - 3x + a),$$
$$F, \text{ slow feedback:} \quad \dot{a} = \varepsilon(x - \lambda),$$

where ε, λ are constants and ε is small. For the Belousov reagent choose $\lambda < 1$, and for the Z-reagent $\lambda > 1$. Then by [23] the decrease of the parameter λ past the value $\lambda = 1$ gives the Hopf bifurcation. The resulting flows are illustrated in Figure 8, with the catastrophe slow manifold shown dotted. For the Belousov reagent a theorem of Kopell and Howard [6] ensures the existence and stability of closed orbits for equation (3) near the Van der Pol attractor, but only provided diffusion is sufficiently small. For the Z-reagent Conley and Carpenter [2] have proved the existence, but not yet the stability, of homoclinic and closed orbits, provided ε is sufficiently small, for the case of a large diffusion of x. What is needed is to handle both cases together and prove stability for large diffusion, giving an estimate on ε, the ratio of slow/fast. Then extend the results to the cusp-catastrophe [23, Example 8]. Finally identity the equations with the explicit chemical reactions, make a quantitative model, and predict the speeds of the various waves.

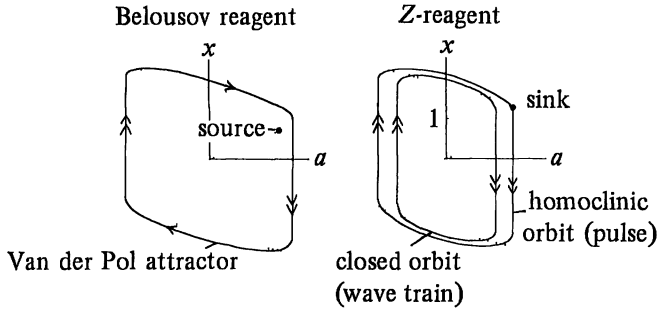


FIGURE 8

References

1. V. I. Arnold, *Singularities of differentiable functions*, these PROCEEDINGS.

2. G. A. Carpenter, *Travelling wave solutions of nerve impulse equations*, Thesis, University of Wisconsin, Madison, Wis., 1974.

3. D. H. Fowler, *The Riemann-Hugoniot catastrophe and Van der Waal's equation*, Towards a Theoretical Biology 4, edited by C. H. Waddington, Edinburgh Univ. Press, 1972, pp. 1–7.

4. C. Hall, P. J. Harrison, H. Marriage, P. Shapland and E. C. Zeeman, *A model for prison disturbances*, Jour. Math. and Stat. Psychology (to appear).

5. C. A. Isnard and E. C. Zeeman, *Some models for catastrophe theory in the social sciences*, Use of Models in the Social Sciences, edited by L. Collins, Tavistock, London, 1974.

6. N. Kopell and L. N. Howard, *Pattern formation in the Belousov reaction* (A.A.A.S., 1974, Some Math. Questions in Biology, VIII), Lectures on Math. in the Life Sci., vol. 7, Amer. Math. Soc., Providence, R. I., 1974, pp. 201–216.

7. K. Lorenz, *On aggression*, 1963; English transl., Methuen, London, 1967.

8. J. N. Mather, *Right equivalence*, Warwick University, 1969 (preprint).

9. T. Poston and A. E. R. Woodcock, *Zeeman's catastrophe machine*, Proc. Cambridge Philos. Soc. **74** (1973), 217–226.

10. ———, *A geometrical study of the elementary catastrophes*, Lecture Notes in Math., vol. 373, Springer-Verlag, Berlin, 1974.

11. L. S. Shulman and M. Revzen, *Phase transitions as catastrophes*, Collective Phenomena **1** (1972), 43–47.

12. F. Takens, *Constrained differential equations*, Dynamical Systems, Warwick, 1974, edited by A. K. Manning, Lecture Notes in Math., vol. 468, Springer-Verlag, Berlin, pp. 80–82.

13. R. Thom, *Topological models in biology*, Topology **8** (1969), 313–335. MR **39** #6629.

14. ———, *Stabilité structurelle et morphogénèse*, Benjamin, New York, 1972.

15. ———, *Phase-transitions as catastrophes*, Conf. on Stat. Mechanics, Chicago, Ill., 1971.

16. ———, *A global dynamical scheme for vertebrate embryology* (A.A.A.S., 1971, Some Math. Questions in Biology, IV), Lectures on Math. in the Life Sci., vol. 5, Amer. Math. Soc., Providence, R. I., 1973, pp. 1–45.

17. D. J. A. Trotman and E. C. Zeeman, *The classification of elementary catastrophes of codimension* $\leqq 5$, Lecture Notes, Warwick University, 1974.

18. G. Wassermann, *Stability of unfoldings*, Lecture Notes in Math., vol. 393, Springer-Verlag, Berlin, 1974.

19. ———, *(r, s)-stability of unfoldings*, Regensburg Universität, 1974 (preprint).

20. A. T. Winfree, *Spacial and temporal organisation in the Zhabotinsky reaction*, Aakron Katchalsky Memorial Sympos., Berkeley, Calif., 1973.

21. ———, *Rotating chemical reactions*, Scientific American **230** (1974), 82–95.

22. E. C. Zeeman, *Geometry of catastrophes*, Times Literary Supplement, 1971, pp. 1556–1557.

23. ———, *Differential equations for heartbeat and nerve impulse*, Dynamical Systems, edited by M. M. Peixoto, Academic Press, New York, 1973, pp. 683–741.

24. ———, *A catastrophe machine*, Towards a Theoretical Biology 4, edited by C. H. Waddington, Edinburgh Univ. Press, 1972, pp. 276–282.

25. ———, *Applications of catastrophe theory*, Manifolds, Tokyo, 1973, Univ. Tokyo Press, 1975, pp. 11–23.

26. ———, *On the unstable behaviour of stock exchanges*, J. Math. Economics **1** (1974), 39–49.

27. ———, *Primary and secondary waves in developmental biology* (A.A.A.S., 1974, Some Math. Questions in Biology, VIII), Lectures on Math. in the Life Sci., vol. 7, Amer. Math. Soc., Providence, R. I., 1974, pp. 69–161.

UNIVERSITY OF WARWICK
COVENTRY, ENGLAND CV4 7AL

Section 20

History and Education

Труды Международного Конгресса Математиков
Ванкувер, 1974

Об Исследованиях по Истории Математики, Проводящихся в Советском Союзе

Б. В. Гнеденко

Господин президент!

Дамы и господа!

Я счастлив представившейся возможностью подготовить и произнести в столь высоком собрании обзор исследований советских ученых в области истории математики. Нет нужды говорить о том, что эта тема необъятна по своим масштабам, разнообразию возникающих в ней вопросов и возможных подходов к их решению. Естественно поэтому, что мне придется оставить вне моего внимания большое число интересных результатов, гипотез и даже целых направлений исследований. К тому же я и сам не стремлюсь превратить мой доклад в перечень имен и частных достижений.

Прежде всего хотелось бы обратить внимание на то, что последнее время интерес к вопросам истории математики усилился и ими занимаются не только специалисты в области истории науки, но и собственно математики. В связи с этим расширился диапазон исследований и их глубина. Если за период с 1917 по 1947 г. в Советском Союзе было опубликовано около 250 работ по истории математики, то за следующее десятилетие появилось свыше 400 монографий и статей, а в десятилетии с 1957 по 1967 уже более 500. Причин для такого систематического роста внимания к истории математики (как и вообще к истории науки) много и они далеко выходят за рамки только количественного роста той части общества, которая занимается научными исследованиями. Пожалуй, еще большее значение имеют иные причины, в том числе и следующие: отчётливое понимание того, что история науки является (а) непременной составной частью всеобщей истории, (б) одним из важнейших условий современного развития науки, (в) методом

совершенствования обучения, (г) базой научной методологии, (д) одним из важнейших источников анализа процессов мышления.

Всеобщая история до сих пор уделяет основное внимание борьбе группировок за государственную власть, сражениям, смене правителей. Однако прогресс человеческого общества далеко не исчерпывается только этими аспектами истории. Для человечества несравненно важнее чем перечисление имен завоевателей, ознакомление с тем изумительным путем побед знания над незнанием, которым сопровождалась вся его история. Для всего развития человеческого общества ни с чем несравнимую роль сыграло то, что в Древней Греции математика превратилась из сборника практических рецептов в дедуктивную науку. На судьбы всей истории открытие законов Ньютона и создание дифференциального и интегрального исчислений оказали куда большее влияние чем битвы при Ватерлоо или же при Каннах. Однако, до сих пор в учебниках истории, которые изучают школьники всего мира, открытиям Ньютона и Лейбница уделяют в лучшем случае несколько слов, тогда как даже в самых кратких курсах истории указанным битвам посвящены многие страницы текста. Нужно думать, что здесь сказывается не столько неведение и отрицание роли физико-математических наук, сколько давняя традиция, которая владеет умами историков и отодвигает на задний план то обстоятельство, что развитие точных наук не только обусловлено общественными отношениями, но и само общественное развитие в значительной мере зависит от состояния точных знаний. Мы должны признать, что история науки является важнейшей частью истории человеческого общества и его культуры и от нее нельзя отрывать изложение исторического процесса.

Относительно значения истории науки для современного развития самой науки установившегося мнения до сих пор нет. И мы можем в научной среде встретить многочисленных представителей, которые придерживаются того мнения, что занятия историей науки не полезны, а вредны для прогресса науки, поскольку они отнимают знания, силы и время на изучение того, что бесконечно устарело и уже не имеет реального значения. Конечно, история науки необходима для изучения истории общественного развития, для философии и для общего образования, но сама то наука непрерывно движется вперед, обогащается новыми идеями и фактами, новыми концепциями и областями исследования. Как же в таких условиях знание прошлого может оказаться полезным для научных исследований наших дней? Нельзя ли скорее спросить: не оказывает ли груз прошлого тормозящего влияния на современные исследования, мешая появлению новых идей?

Я убежден, что значение истории науки для развития самой науки со временем будет возрастать, а не убывать. Дело в том, что задача истории математики не сводится только к описанию пройденного математикой пути, но и осмысливание его. История математики, как любая живая наука со временем изменяет свое содержание и по-новому подходит к своим задачам. Если на первых порах её развития основное её содержание сводилось к описательной части, то теперь собирание фактов и их описание является

лишь начальным моментом. Основное же содержание истории математики мы видим в выяснении причин появления тех или иных руководящих идей, направлений исследования, формулировке закономерностей развития математики, выявлении тех тормозящих факторов, которые приводили к краху прогрессивных начинаний. Естественно, что в таком плане историей математики могут успешно заниматься лишь те лица, которые сами близки к самостоятельному творчеству. В какой то мере это подтверждается тем, что каждый ученый математик в какой то мере вынужден заниматься историей науки, знакомясь минимум с историей того вопроса, которым он занимается. Пусть при этом он и не уходит в глубину веков, а лишь тщательно изучает работы и идеи последних десяти-пятнадцати лет. Но при этом он узнает ту научную атмосферу, в которой рождались проблемы и в которой работали его предшественники и стремится понять, что следует изменить, чтобы найти нужный путь. Постоянно происходит то, о чем в свое время красочно писал И. Ньютон: "Если я увидел больше других, то только потому, что стоял на плечах гигантов."

Собственно тоже самое происходит постоянно с каждым исследователем: новое удается открыть в значительной мере потому, что используется опыт и результаты предшественников. А ведь история науки как раз и имеет своей целью собрать этот опыт и, обобщив его, найти те закономерности прогресса науки, которые толкают её вперед. Недаром сейчас каждой большой специальной работе предпосылают исторический обзор. Это делается не потому, что такова традиция, а потому, что такой обзор позволяет охватить предмет исследования глубже, полнее и многостороннее.

Знание прошлого науки позволяет в концентрированном виде получить сведения об истоках идей и фактов, о формировании научных понятий о полетах мысли, опередивших свое время, а потому забытых и теперь уже неизвестных. Немало того, что было найдено в свое время, потом было забыто, вышло из моды, но затем в новых условиях заиграло всеми красками и оказалось нужным для современности. Таких случаев математика знает много, также как и все естествознание. Во многих случаях подлинный смысл открытия выяснялся на дальнейших этапах развития науки. Так в свое время случилось с открытием теории относительности в работах А. Пуанкаре. На глазах нашего поколения возродилась и развилась математическая теория полетов в космос, созданная К.Э. Циолковским. Он значительно опередил свое время и при его жизни его работы не получили признания. Но то, что история науки не забыла о них, оказало существенное влияние на развитие теории и практики космических полетов наших дней.

Для воспитания новых поколений история науки важна сама по себе, безотносительно к её ценности для развития современной науки. На примерах прошлого следует учить молодежь страсти к поиску нового, к творчеству, научному упорству и научной честности. Примеры великих представителей науки прошлого, их упорство в преодолении трудностей способны вызвать молодых людей на подобные же стремления. Талантливых представителей

юношества такое воспитание может натолкнуть на выбор научной карьеры, остальных же приучит к упорной работе, непрестанному поиску пути, который максимально соответствовал бы стоящей перед ними проблеме. На примерах творческой жизни ученых, на примерах истории их открытий можно вдохновить учащихся на поиски неизведанного, привить им веру в их творческие силы и стремление испытать их на решении задач, возникающих перед современной наукой.

История науки может и должна вводить учащихся в творческую лабораторию ученого и показать как постепенно возникали мысли об открытии, формулировки результатов, идеи доказательств. Для этой цели необходимо увидеть не только окончательный продукт работы ученого—оформленную статью, монографию или созданную им теорию, но сам процесс творчества: возникновение вопроса, прикидки доказательства, поиски метода и т.д.

В этом плане мне хотелось бы обратить внимание на поучительное исследование Г.П. Матвиевской [**1**], посвященное неопубликованным записным книжкам Л. Эйлера, хранящимся в архивах Академии наук СССР.

Я позволю себе привести здесь заключительную часть статьи Г. П. Матвиевской: "... записные книжки показывают, что весьма часто Эйлер приходил к своим результатам чисто экспериментальным путем, производя вычисления, которые поражают своей сложностью. Подтвердив свой окончательный вывод рядом таких числовых примеров, он констатирует полученный математический факт и зачастую этим ограничивается на первой стадии работы. Иногда он намечает пути дальнейшего исследования, а затем возвращается к заинтересовавшему его вопросу много раз на протяжении ряда лет. В соответствующих заметках из записных книжек Эйлер делает попытки доказательства высказанного утверждения, пробуя для этого различные пути и способы. Некоторые из полученных доказательств не являются строгими, и Эйлер обычно сам указывает на эту нестрогость. Но после нескольких попыток он приходит к окончательному строгому доказательству и только тогда считает вопрос исчерпанным."

До сих пор история науки была исключительно историей успехов человеческого разума. Хотелось бы думать, что наступила пора когда история науки может стать и историей мышления. Математика в силу своей специфики, может дать в этом отношении много полезного. В известных работах А. Пуанкаре и Ж. Адамара увлекательно изложены факты, в которых авторы сами были действующими лицами. Каждому, кто испытал радость научного творчества, хорошо известно какой путь для этого следует пройти. Прежде всего должна быть та проблема, которая увлекает исследователя и заставляет думать о ней всё время и во время работы и во время отдыха. Зачастую пути, которые приводили к успеху в других, казалось бы близких, вопросах, оказываются недостаточными. Мысль ищет решения, сопоставляет. Рассматриваются частные случаи, на которых делаются попытки уловить общий метод доказательства и общую формулировку результата. Проходят дни, месяцы, а порой и годы, пока простая мысль, простой геометрический образ

или физическое сопоставление не подскажут то, что так долго и безуспешно разыскивалось. Внезапное озарение, в котором содержится искомое решение, прекрасно описано А. Пуанкаре. Далее начинается период жатвы, когда мысль из основной идеи выводит многочисленные следствия.

То, что только что было сказано, является лишь внешним описанием творческого процесса. Это еще не модель мышления, которая может быть исследована. Однако для создания модели мышления нужны наблюдения и история математики может дать для этого буквально неограниченный материал.

История науки важна не только потому, что она необходима для решения ряда научных и педагогических проблем. Она важна и сама по себе, как памятник человеческому гению, который позволил человечеству пройти великий путь от полного незнания и полного подчинения силам природы до великих замыслов о завоевании космоса и до первых шагов в их осуществлении.

История науки является тем факелом, который освещает новым поколениям путь дальнейшего прогресса и передает им священный огонь Прометея, толкающий их на новые открытия.

О значении истории науки прекрасно сказал еще Лейбниц в одном из сочинений, оставшихся неопубликованными при его жизни [2] "Весьма полезно познать истинное происхождение замечательных открытий, особенно таких, которые были сделаны не случайно, а силою мысли. Это приносит пользу не столько тем, что история воздает каждому свое и побудит других добиваться таких же похвал, сколько тем, что познание метода на выдающихся примерах ведет к развитию искусства открытия."

Исследования по истории математики, которые проводились в Советском Союзе, можно разбить на следующие группы:

1. Разработка вопросов истории Древнего Востока (Вавилон, Древние Египет и Китай);
2. Математика Древней Греции;
3. Математика средневековья, особенно Средней Азии и мусульманского Востока;
4. Математика нового времени;
5. Развитие математики в России, Грузии, Армении и др. частях Советского Союза;
6. Издание трудов классиков математической науки с комментариями и творческими биографиями;
7. Разработка проблем истории отдельных математических дисциплин.

Естественно, что работа в области истории математики проводилась во многих планах. Во-первых большое число исследований посвящено разработке отдельных конкретных вопросов истории математики; во-вторых в ряде работ изложены общие концепции истории математики; в третьих выполнено большое число работ чисто биографического характера, в четвертых опубликованы работы по истории математики специально для целей педагогики.

Работам первого из указанных направлений будет посвящена основная часть настоящего доклада. О работах остальных направлений мы дадим сейчас же некоторые сведения.

Из оригинальных работ, трактующих историю математики в целом, мне хотелось бы отметить в первую очередь большую статью А.Н. Колмогорова, написанную им для Большой Советской Энциклопедии [3]. Этот сжатый очерк истории математики от её возникновения до наших дней представляет значительный интерес. А.Н. Колмогоров предложил периодизацию развития математики. Согласно этой концепции можно выделить четыре периода: (1) зарождения математики как отдельной теоретической науки. Этот период заканчивается в VI-V веке до нашего летоисчисления; (2) период элементарной математики или иначе—период математики как науки о постоянных величинах. Этот период продолжался до XVII века. (3) Период математики как науки о переменных величинах. Продолжается до начала XIX века. (4) Современная математика характеризуется значительным расширением её предмета и более высокой степенью абстракции её понятий.

Высокой оценки заслуживает коллективный трёхтомный труд "История математики" под редакцией профессора А.П. Юшкевича [4]. Изложение ведётся от глубокой древности до начала XIX века. Авторы отдельных глав являются глубокими специалистами в соответствующих областях математики или же её истории и сами внесли много нового как в трактовки известного материала, так и в привлечение новых данных. Это относится ко всем трем томам. Специалисты заметят много новых мыслей почти во всех главах книги. Это относится в особенности к изложению материала касающегося развития математики в Древней Греции, средневековом Китае, в странах ислама. Полезные трактовки и факты содержатся и в главах, посвященных развитию отдельных математических дисциплин. Это трехтомное сочинение трудно рассматривать в качестве учебника, поскольку оно задумано и выполнено как монографическое исследование, написанное с единых методологических позиций.

Мне хотелось бы указать здесь на две следующие монографии А.П. Юшкевича "История математики в средние века" [5] и "История математики в России" [6], в которых дана широкая картина развития математики. В первой из них—в средние веке и при том значительное внимание уделено странам мусульманского мира, изучению которых в Советском Союзе занимается весьма большой коллектив ученых. Во второй—дано последовательное изложение математической жизни в России со времен образования Древней Руси. Как первая, так и вторая монографии основаны на многочисленных новых материалах, изученных главным образом за годы, прошедшие после окончания второй мировой войны.

Здесь мне хотелось бы несколько большее внимание уделить работам, посвященным математике средних веков, поскольку, пожалуй, в этой области исследований советским математикам принадлежит наиболее широкий фронт работ. В результате многочисленных критических публикаций текстов,

принадлежащих авторам средних веков, в значительной мере математикам мусульманского востока, Закавказья, Китая, Индии и Западной Европы прояснилась впечатляющая картина серьезного математического прогресса именно в эпоху Средневековья. За короткий срок появилось большое число переводов на русский язык сочинений математиков Средней и Ближней Азии, писавших на арабском языке. Ряд из этих сочинений до последнего времени вообще не был опубликован ни на одном из европейских языков. Конечно, математики Ближнего и Среднего Востока—арабы, иранцы, сирийцы, народы средней Азии (узбеки, таджики) и др.—были учениками в первую очередь греков и в значительно меньшей степени индусов и китайцев. Но ученики не только продолжали традиции своих учителей. Их внимание привлекало не только направление классических изысканий в духе Евклида, значительного развития достигло вычислительно-алгоритмическое направление, которое только намечалось в Древней Греции и не получило развития в эпоху распада Римской империи. Интерес к этого рода задачам тесно связан с исследованиями по астрономии и географии.

Я позволю себе кратко коснуться некоторых из проведенных работ. Большое внимание было уделено математическим работам Омара Хайяма, русское издание которых появилось в 1962 г. Изучению его вклада в арифметику, теорию параллельных линий были посвящены работы А.П. Юшкевича, Б.А. Розенфельда, С.А. Красновой. Тщательному изучению были подвергнуты трактаты Насир ад-Дина ат-Туси в ряде работ Б.А. Розенфельда, Г.Д. Мамедбейли, С.А. Ахмедова. Последний из названных лиц заметил, что в арифметическом трактате ат-Туси имеется правило возведения в степень бинома, а также таблица биномиальных коэффициентов до 12 степени включительно. Там же изложен алгоритм извлечения корней с любым натуральным показателем по способу "Горнера-Руффини". Этим самым отодвинуто с XV века на XIII открытие как формулы бинома Ньютона, так и способа извлечения корней Горнера-Руффини. Ранее это относилось к 1427 г., когда появился "Ключ арифметики" ал-Коши, теперь же выяснилось, что они были известны уже в 1265 г. Интересны результаты исследования "Ключа арифметики" Гияс ад-Дина ал-Каши и его же "Трактата об окружности", в которых были изложены исключительно тонкие и изящные методы вычисления числа π и вычисления синуса одного градуса посредством приближенного решения уравнения третьей степени итерационным методом. Содержательны работы М.И. Медового по изучению арифметических методов в трактате Абу-л-Вафы. Он нашел, что этот ученый уже прибегал к использованию отрицательных чисел, восстановил во всех деталях технику операций в деловой практике с дробями и подверг критическому анализу ряд ранее утвердившихся в истории положений. В частности он обоснованно отвергает наличие в странах ислама двух школ—прогреческой и проиндийской.

Конечно, сказанным не ограничивается работа советских историков математики по изучению вклада ученых средних веков в прогресс математической науки. Большое число исследований было посвящено и анализу произведений,

выполненных авторами средневековой Западной Европы, Грузии, Армении, а также России. Уже 90 лет прошло после первых публикаций В.В. Бобынина о русских математических рукописях XIV-XVII столетий, но до сих пор не было дано ни исчерпывающего сравнительного анализа их содержания, ни полной их публикации с комментариями. Подробное изучение имеющихся в хранилищах русских математических рукописей только в 1955 г. было завершено К.И. Швецовым. Позднее он же вместе с Ю.А. Белым опубликовал и прокомментировал русскую геометрическую рукопись первой четверти XVII века. Интересные работы были выполнены В.П. Зубовым, Т.И. Коншиной, Л.Е. Майстровым по анализу первой математической русской рукописи Кирика Новгородского, а также новгородских берестяных грамот.

Всё, что было только что нами упомянуто существенно продвинуло наши знания прошлого математики в разных странах и у разных народов, позволило выявить преемственность этих знаний и вклад различных народов в человеческую культуру.

Несомненно, что советские историки математики не могли пройти мимо развития математики в новое время. Особенно интенсивно и разнообразно развивались наши знания о периоде, связанном с появлением и становлением математического анализа и аналитической геометрии. Значительное внимание привлекали такие титаны математической мысли как Ньютон, Лейбниц, Декарт, Л. Эйлер, Коши, Лежандр, Лаплас. История и предистория дифференциального и интегрального исчислений были предметом многочисленных исследований, в том числе таких крупных математиков как А.Н. Колмогоров. Большой труд был вложен в издание на русском языке и комментирование работ классиков—Кеплера, Кавальери, Декарта, Ньютона, Монжа, Карно, Л. Эйлера и др. Особенно много было в этом отношении сделано по наследию Л. Эйлера. Так были изданы его знаменитые "Введение в анализ", "Дифференциальное и интегральное исчисления", "Теория Луны" и т.д. Монографическому исследованию подверглись работы Л. Эйлера по дифференциальным уравнениям (Н.И. Симонов, Ф.И. Франкль), вариационному исчислению (К.А. Рыбников, Н.С. Кошляков), теории функций (А.И. Маркушевич, С.Е. Белозеров), теории чисел (Б.А. Венков, И.Г. Башмакова), механике (Ю.А. Крутков, Г.К. Михайлов) и т.д. Интересный по содержанию коллективный труд, посвященный анализу творчества Л. Эйлера был подготовлен и опубликован к 250-летнему юбилею Л. Эйлера. Сделаны первые шаги по изучению рукописного наследия, хранящегося в архивах Академии наук СССР этого титана математической мысли. Глубокие по содержанию монографии, основанные на изучении архивных материалов и подлинников появились в печати относительно И. Кеплера, Б. Паскаля, Г. Лейбница. Эти монографии содержат не только новые трактовки, но и значительный свежий фактический материал.

Не остались в стороне и труды крупнейших ученых XIX и XX столетий. Естественно, что максимальное внимание при этом было уделено творцам неевклидовой геометрии Н.И. Лобачевскому и Я. Бойяи. Изданы на русском

языке собрания их сочинений в сопровождении подробных и квалифицированных комментариев. Лобачевскому посвящены несколько специальных монографий и биографических очерков. В настоящее время заканчивается подготовка к изданию рукописных материалов, посвященных педагогическим и философским взглядам этого выдающегося ученого. Изданы труды П.Л. Чебышева, А.М. Ляпунова, М.В. Остроградского, Н.Е. Жуковского, Г.Ф. Вороного, а также А. Пуанкаре, Риманна, С.Н. Бернштейна и ряда других выдающихся ученых.

Издание собраний сочинений классиков математической науки является не только важным моментом в развитии историко-математических исследований. Не меньшее значение имеют и другие аспекты их использования. Прежде всего я хотел бы указать на стимулирование научных исследований и на возрождение ценных идей, которые не были своевременно замечены и развиты. Я позволю себе сказать об этом несколько позднее об одном таком случае. Но, пожалуй, особенное значение такого рода издания играют в воспитательных целях. Действительно широте взглядов и смелости мысли следует учиться не у тех, кто эти мысли пересказывает, а у тех, кто их породил и развил. Сейчас же, в период массового увлечения научными исследованиями, нередко внимание молодежи обращается не на действительно фундаментальные проблемы науки а на доделку мелочей, оставшихся незавершенными после исследователей далеко не первого класса.

В 1950–1952 г.г мне и моему коллеге Е.Я. Ремез Академия наук Украины поручила в связи со 150-летним юбилеем М.В. Остроградского изучить рукописный фонд, хранящийся в Киевской публичной библиотеке. Рукописи Остроградского были изучены очень плохо и состояли из листов и клочков бумаги разного формата. Тщательное их изучение принесло нам много неожиданных находок. Об одной из них я хотел бы рассказать сейчас. На глаза нам попались несколько обрывков, на которых были рассмотрены примеры вычисления приближенных значений иррациональных чисел. Оказалось, что Остроградский на числовых примерах разработал два алгоритма приближения иррациональных чисел рациональными числами с исключительно быстрой сходимостью (значительно превышающей быстроту сходимости свойственную непрерывным дробям). Повидимому, эти подсчёты производились Остроградским в последние дни его жизни, а потому не были доведены им до публикации. Е.Я. Ремез [7] удалось расшифровать эти заметки Остроградского и подарить математике хорошие математические результаты, оставшиеся неизвестными в течение ста лет. Я позволю себе вкратце дать описание этих алгоритмов.

Пусть число ω таково, что $0 < \omega < 1$. Представим это число в виде отношения двух отрезков A и B ($A < B$). Для некоторого целого p_0 имеет место равенство $B = p_0 A + A_1$, где $A_1 < A$. Далее последовательно $B = p_1 A_1 + A_2$ ($A_2 < A_1$), $B = p_2 A_2 + A_3$ ($A_3 < A_2$) и т.д. Легко показать, что имеет место представление

$$\omega = \frac{1}{p_0} - \frac{1}{p_0 p_1} + \frac{1}{p_0 p_1 p_2} \mp \cdots \quad (p_0 < p_1 < p_2 < \cdots).$$

Это разложение единственно для всех иррациональных чисел. При аппроксимации числа ω суммой первых n членов погрешность будет менее чем $1/(n+2)!$. Из единственности бесконечных разложений указанного типа можно без всяких дополнительных рассуждений сделать выводы относительно иррациональности ряда чисел. Так числе

$$1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots = 1 - \frac{1}{e}$$

можно сразу сказать, что оно иррационально, поскольку для него разложение ранее указанного вида бесконечно.

Второй алгоритм состоит в следующем: пусть $q_0, q_1, q_2, \cdots$ натуральные числа, определяемые посредством равенств:

$$B = q_0 A + A' \quad (A' < A), \qquad q_0 B = q_1 A' + A'' \quad (A'' < A'),$$
$$q_0 q_1 B = q_2 A'' + A''' \quad (A''' < A''), \qquad q_0 q_1 q_2 B = q_3 A''' + A^{\mathrm{iv}} \quad (A'' < A'''),$$
$$\cdots$$

Можно доказать, что

$$\omega = \frac{1}{q_0} - \frac{1}{q_1} + \frac{1}{q_2} - \frac{1}{q_3} + \cdots,$$

причем выполняются неравенства $q_1 > q_0 (q_0 + 1)$, $q_2 > q_1 (q_1 + 1), \cdots$.

Погрешность, которую дает приближение числа ω посредством использования первых n членов не превосходит величины $\varepsilon_n = 2^{-2^n}$. Мы видим, что точность приближения не только превосходит аппарат непрерывных дробей, но и первый из указанных нами алгоритмов Остроградского.

Естественно, что для развития интереса к истории науки, необходимы содержательные, хорошо написанные учебники и монографии. За последние годы появилось большое число монографий, как советских авторов, так и переводных. Некоторые из монографий широкого плана мной были названы ранее. Еще большее число мной даже не были упомянуты. Мне хотелось бы упомянуть сейчас лишь два таких издания. Во-первых в высшей степени интересную книгу В.А. Медведева "Развитие понятия интеграла" [8], в которой с глубоким знанием дела дан исторический очерк развития понятия интеграла с глубокой древности. Во-вторых считаю нужным упомянуть фактически пятитомное сочинение большого коллектива авторов "История отечественной математики" [9]. Как всегда, труд, в котором принимали участие десятки авторов, не может быть однородным по литературным достоинствам и объективно выявлять место каждой из ветвей математической науки в современном её развитии. Однако, для общего ознакомления с тем, что было сделано для развития математики в России, а позднее в Советском Союзе эта монография вполне пригодна.

К сожалению, до сих пор еще не создан учебник истории математики,

который удовлетворял бы современным к нему требованиям. За последние годы был сделан ряд таких попыток. Недавно появилось 2-ое издание оригинального труда этого жанра, принадлежащего К.А. Рыбникову [**10.**] Несомненно, что работа по созданию учебников истории математики должна продолжаться и должны быть представлены различные научные подходы к их построению. О том, что могут быть существенно различные точки зрения на характер построения учебника истории математики говорит хотя бы опыт непериодического советского сборника "Историко-математические исследования", в котором лет пятнадцать-двадцать назад были опубликованы вводные лекции к курсу истории математики нескольких авторов. Все изложенные подходы были существенно различны и каждый из них заслуживает внимательного рассмотрения [**11**]—[**13**].

Мне хотелось бы еще остановиться на вопросе, связанном с местом истории математики в школьном курсе математики. Конечно, никакого систематического изложения истории математики в школе не может быть, но беседы о развитии математики, о великом вкладе математиков в культуру человечества и его техническую мощь абсолютно необходимы. На примерах великих ученых прошлого молодые люди должны учиться настойчивости в достижении цели и умению выбирать достойную цель. Для этого как учебники математики для школьников, так и методические пособия для учителей должны содержать краткие сведения из истории науки. Должны быть написаны книги для чтения по истории науки, о выдающихся ученых прошлого. Всё это должно быть написано увлекательным языком и содержать творческие биографии ученых, а также показывать как тесно связано развитие математики с требованиями жизни, научной и житейской практики.

По моему предложению в журнале для учителей математики "Математика в школе" теперь систематически печатается календарь исторических дат. Своевременно за месяц-два каждый преподаватель узнает какие события происходили в прошлом в математической науке. Этим в руки педагога даётся серьезная возможность своевременно сообщать школьникам интересные исторические сведения.

Первоначально исследователи по истории математики концентрировались в Московском университете около семинара по истории математики. Позднее в 1945 году был создан институт истории естествознания в Академии наук СССР, где стал формироваться костяк исследователей в области истории науки, в том числе и математики. После того, как Союз стал оправляться от разрухи, причиненной войной 1941–1945 г.г., интерес к истории математики стал возникать и в других научных центрах страны—Киеве, Ереване, Тбилиси, Ташкенте, Ленинграде, Баку, Казани и др. Я счастлив, что мне довелось быть организатором семинара по истории математики в Киеве. Позднее я передал руководство этим семинаром И.З. Штокало, а в настоящее время фактическим соруководителем этого семинара является увлеченный историк науки А.Н. Боголюбов.

Работы советских историков математики публикуются в различных изда-

ниях, но в первую очередь в "Трудах института истории естествознания и техники", "Историко-математических исследованиях", а также на страницах журналов "Успехи математических наук", "Математика в школе", а также в трудах университетов, педагогических и технических институтов.

Советские историки математики уверены, что они выполняют исключительно полезную работу, поскольку человечество должно знать о взлетах мысли лучших своих представителей и, опираясь на знание прошлого, уверено строить великое здание науки. В этом труде необходимо международное сотрудничество и мы готовы участвовать во всех международных акциях, которые направлены на развитие науки, культуры и мирного сотрудничества народов.

Литература

1. Г.П., Матвиевская, *О неопубликованных рукописях Леонарда Эйлера по диофантову анализу*, Историко-математические исследования, вып. XII, 1960, 107–186.

2. G. Leibnitz, *Matematische Schriften*, B. V. Halle, 1858, S. 392.

3. А. Н. Колмогоров, *Математика*, Большая советская энциклопедия, т. 26, стр. 464–483.

4. *История математики*, т.т. 1–3, 1970–1972; под редакцией А.П. Юшкевича.

5. А.П. Юшкевич, *История математики в средние века*, Гос. издат. физ.-мат. лит., Москва, 1961, стр. 1–448.

6. ———, *История математики в России*, "Наука", Москва, 1968, стр. 1–591.

7. Е.Я. Ремез, *О знакопеременных рядах, которые могут быть связаны с двумя алгорифмами М.В. Остроградского для приближения иррациональных чисел*, Успехи матем. наук **6** (1951), 33–42.

8. Ф.А. Медведев, *Развитие понятия интеграла*, "Наука", Москва, 1974, стр. 1–423.

9. *История отечественной математики*, т.т. 1–4, Киев, 1970.

10. К.А. Рыбников, *История математики*, изд. 2-е, Издат. Москов. ун-та, 1974, 1–455.

11. С.А. Яновская, *Вводная лекция к курсу "История математики"*, Историко-математические исследования, т. XI, 1958, стр. 185–208.

12. К.А. Рыбников, *О предмете истории математики*, Историко-математические исследования, т. XI, 1958, стр. 209–224.

13. Б.В. Гнеденко, *Вводная лекция к курсу истории математики*, Математика в школе, No. 1, 1963, стр. 3–13.

Московский Государственный Университет
Москва, СССР

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

The Theory of Matrices in the 19th Century

Thomas Hawkins

Although the origins of the theory of matrices can be traced back to the 18th century and although it was not until the 20th century that it had become sufficiently absorbed into the mathematical mainstream to warrant extensive treatment in textbooks and monographs, it was truly a creation of the 19th century.

When one contemplates the history of matrix theory, the name that immediately comes to mind is that of Arthur Cayley. In 1858 Cayley published *A memoir on the theory of matrices* in which he introduced the term "matrix" for a square array of numbers and observed that they could be added and multiplied so as to form what we now call a linear associative algebra. Because of this memoir, historians and mathematicians alike have regarded Cayley as the founder of the theory of matrices; he laid the foundations in his 1858 memoir, so the story goes, upon which other mathematicians were then able to erect the edifice we now call the theory of matrices.

For convenience I shall refer to this interpretation of the history of matrix theory as the Cayley-as-Founder view. It is a very simplistic interpretation which, as I will indicate, does not make much historical sense. The history of the theory of matrices is much more complex than the Cayley-as-Founder view would imply. Indeed its history is truly international in scope and hence seems an especially appropriate subject for a Congress such as this. I will begin by indicating several reasons why Cayley's memoir of 1858 does not have the historical significance that the Cayley-as-Founder view suggests.

In the first place, Cayley's celebrated memoir went generally unnoticed, especially outside of England, until the 1880's. This was before the days of comprehensive abstracting journals—the first began ten years later in 1868—and Cayley had published his memoir in the Transactions of the Royal Society of London, some-

thing that does not seem to have been widely read for its mathematical content.

Secondly, the ideas Cayley expressed in 1858 were not particularly original. The idea of representing a linear substitution (i.e., a linear transformation) by the square array of its defining coefficients is already found in Gauss's treatment of the arithmetical theory of quadratic forms as presented in his *Disquisitiones arithmeticae* of 1801. There we also find the idea of composing two linear substitutions to form a third and the idea of representing substitutions by single letters for convenience.

Furthermore Gauss's notational practices were carried one step further by Eisenstein [1], [2] in his efforts to develop further the general theory of forms envisioned by Gauss. Eisenstein observed that if linear substitutions (in any number of variables) are considered as entities and denoted by letters, then they can be added and multiplied much as ordinary numbers, except, as he stressed, the order of multiplication does matter: ST need not be the same as TS. Eisenstein also introduced the common algebraic notation for products, inverses and powers of linear substitutions and used it to good advantage in his papers on the arithmetical theory of forms during the period 1844–1852 (the year of his untimely death). In the early 1850's, Eisenstein's contemporary, Charles Hermite, who was in contact with Eisenstein, continued the latter's use of the symbolical algebra of linear substitutions, both in his work on the theory of forms and on the transformation of abelian functions.

Thus by the mid-1850's the idea of treating linear substitutions as objects which can be treated algebraically much like ordinary numbers was not very novel. Hence it is not entirely surprising to find that during the period when Cayley's 1858 memoir lay unread, two other mathematicians, Laguerre in France and Frobenius in Switzerland, further developed the consequences of the symbolical algebra of linear substitutions in a fashion similar to that taken by Cayley but without a knowledge of Cayley's memoir. Laguerre's work, which was published in 1867 in the journal of the École Polytechnique, suffered the same fate as Cayley's 1858 memoir. Frobenius' work was published in 1877 in Crelle's journal—one of the leading journals of the time—and was more widely known. Frobenius' paper was also much more substantial than those by Laguerre and Cayley, and I shall have more to say about it further on.

There is another reason why the Cayley-as-Founder view of the history of matrix theory is misleading. By focusing, as it does, upon the *form* of the theory, i.e., the matrix symbolism, it tends to ignore its *content*: the concepts and theorems that make it a bona fide theory. For example: the notion of an eigenvalue, the classification of matrices into types, such as symmetric, orthogonal, Hermitian, unitary, etc., and the theorems on the nature of the eigenvalues of the various types and, above all, the theory of canonical matrix forms—in short, what I shall refer to as the spectral theory of matrices.

Spectral theory did not originate with, or depend upon, the work of Cayley. It originated in the 18th century when various physical investigations led to the consideration of eigenvalue problems. During the 19th century these problems were extricated from their physical contexts and transformed into a purely mathematical

theory at the hands of mathematicians such as Cauchy, Jacobi, Kronecker, Weierstrass and Camille Jordan. This activity spanned, roughly, the 50-year period from 1826–1876, and Cayley played no role in these developments. It was a spectral theory of linear substitutions and quadratic and bilinear forms.

Because spectral theory is an important part of the theory of matrices, I would like to make a few remarks about its history at this point, particularly about the contributions of Cauchy and Weierstrass. The spectral theory of the 19th century was initiated by Cauchy in a paper of 1829 [3]. There he gave the first valid proof that the eigenvalues of an n-by-n symmetric matrix must be real. (Cauchy did use matrices but called them "systems".) The significance of Cauchy's paper and its relation to the work of the 18th century geometers, however, have not been correctly portrayed by historians, and I would therefore like to say a few things along these lines.

One of the great achievements of the 18th century geometers was the successful application of the new analysis of the 17th century to problems in terrestrial and celestial mechanics. In the course of making such applications they were led to consider eigenvalue problems. Most eigenvalue problems arose in connection with the integration of systems of linear differential equations with constant coefficients. It was the physical contexts of these equations, i.e., stability considerations, that focused interest on the *reality* of the eigenvalues. The 18th century geometers had not correctly worked out the solution to these differential equations when multiple roots exist, but they could handle the case of distinct roots quite well—thanks especially to the work of Lagrange—and they could see that stability required that the eigenvalues be real.

As long as the system of differential equations was sufficiently simple—i.e., as long as most of the coefficients were zero—reality could be established by special means, for example by actually solving for the eigenvalues. But in the second half of the 18th century the work of Lagrange and Laplace led to the consideration of differential equations and eigenvalue problems of a much more general nature—i.e., the coefficients were not specific numbers. The problem of establishing the reality of the eigenvalues seemed a terribly difficult problem in these cases, for it meant demonstrating the reality of the roots of a very general nth degree polynomial —the characteristic polynomial. For some time they (understandably) had no idea that the symmetry of the coefficients was relevant; and, in fact, in one such problem they considered—the secular (= long-term) pertubations of the parameters determining the planetary orbits—the symmetry property of the coefficients was not immediately in evidence. Because of the apparent mathematical difficulty of the problem and because they were primarily concerned with the analysis of mechanical problems, Lagrange and Laplace introduced various physical arguments, together with some questionable mathematical reasoning (by modern standards), to establish reality.

Eventually, in a somewhat fortuitous manner (which perhaps proves that two wrongs make a right) Laplace discovered [4] that the symmetry properties of the coefficients in the secular perturbation problem could indeed be used to demon-

strate the reality of the eigenvalues. (The symmetry property in this case is $m_i r_i^{1/2} A_{ij} = m_j r_j^{1/2} A_{ji}$, where m_i = mass of ith planet and r_i = its mean distance to sun.) By modern standards Laplace's reality proof was not valid because what he did was to use symmetry together with the differential equations to derive an equality that implied the solutions had to be bounded as functions of time. Then from the form of the solutions to the differential equations, which, as noted, were not correctly formulated for multiple roots, he inferred the reality of the eigenvalues.

Despite these flaws, Laplace's discovery of the relationship between reality and symmetry was a real breakthrough; Lagrange, for example, never realized the relationship. It is therefore somewhat ironical to find that it was Lagrange, and not Laplace, who had more influence upon Cauchy. When Cauchy wrote his paper in 1829 he was not mindful of the eigenvalue problems that arise in integrating systems of differential equations. He was writing some lectures that dealt with a favorite subject at the École Polytechnique in the early 19th century: the classification of quadric surfaces. And in this connection he was interested in the transformation of a quadratic form in three variables into a sum of square terms only. This problem also arose in the mathematical analysis of the rotational motion of a rigid body as studied by Lagrange in the 18th century [5], [6].

Cauchy was especially influenced by Lagrange's treatment of the transformation of a quadratic form, which was unlike anyone else's in terms of its essentially abstract formulation. Most mathematicians, both before and after Lagrange, regarded the problem as follows: Given a quadratic form in 3 variables, write down the equations for the change to another rectangular system of coordinates. These equations involved sines and cosines of 3 angles, and by using some trigonometry one could show how to eliminate successively the nonsquare coefficients. The proofs involved using the fact that a cubic equation—not the characteristic equation—has a real root.

Lagrange's approach, on the other hand, was this: Consider an arbitrary linear substitution in three variables x, y, z which has the property that it leaves $x^2 + y^2 + z^2$ unchanged. Lagrange showed that the coefficients of such a substitution must satisfy the now-familiar orthogonality conditions that characterize an orthogonal substitution. His problem was therefore to prove the existence of such an orthogonal transformation which takes the given quadratic form into a sum of square terms. Furthermore, unlike the other treatments of the problem, Lagrange's made the consideration of an eigenvalue problem central. That is, he showed that the eigenvalue problem determined by the coefficients of the quadratic form yields, as the eigenvectors, the coefficients of the desired orthogonal transformation; and the eigenvalues are the coefficients of the square terms in the transformed form.

Naturally the existence of the orthogonal transformation depended upon the reality of the eigenvalues—the roots of the characteristic equation. This problem did not seem as overwhelming as the others because it was simply a case of a cubic, and Lagrange succeeded in demonstrating that this cubic has the "remarkable" property of having all its roots real, no matter what values are assigned to the coefficients of the associated quadratic form. What is especially significant about

Lagrange's formulation of the principal axis theorem is that it is immediately generalizable to n variables: It is clear what is meant by a quadratic form in n variables, and, thanks to Lagrange, it is clear what an orthogonal transformation in n variables can be taken to mean.

Lagrange was, of course, exclusively interested in the 3-variable case because that was the physically relevant case. Cauchy, however, was in a position—and of a frame of mind—to see that not only was Lagrange's formulation of the principal axis problem generalizable, so was his proof; it was only necessary to translate Lagrange's proof into the language of determinants to see that it was valid for any number of variables. In this connection I must point out that in the 18th century the notion of a determinant was only vaguely formulated and no significant properties were established. Many mathematicians, including Lagrange, made no use of determinants. It was Cauchy who, in 1812, wrote a brilliant memoir [7] which essentially created the theory of determinants as we know it. With this background, it was only natural that he should look at Lagrange's proof in terms of determinants.

In this manner Cauchy established the reality of the eigenvalues of a symmetric matrix, as a part of his generalization of the principal axis theorem. (It is worth noting that Cauchy's 1829 paper thus also represents the beginning of n-dimensional analytic geometry.) Cauchy's generalization was, moreover, simply that—a generalization simply for the sake of an interesting generalization; he did not see the relation of his theorem to the eigenvalue problems stemming from differential equations until it was pointed out to him by Charles Sturm, one of Fourier's students, while he was in the process of writing up his results for publication.

Historians have attempted to capture the significance of Cauchy's work by attributing to him the discovery of the underlying similarity of many of the mechanical problems of the 18th century—i.e., that they involved eigenvalue problems with symmetric coefficients. If that honor can be said to belong to anyone, it belongs to Sturm.[1] The actual historical significance of Cauchy's paper is to be found in its methodology: the theory of determinants. His contemporaries, led by Jacobi, saw in Cauchy's theory of determinants a new and powerful algebraic tool for dealing with n-variable algebra and analysis. It was they, especially Weierstrass, who used determinant-theoretic methods to give a purely mathematical treatment of many of the other eigenvalue problems of the 18th century, viz. those of the form

$$(1) \qquad AX = \lambda BX \quad \text{with related differential equations } B\ddot{Y} = AY,$$

where A and B are symmetric and B is definite.

I cannot go into the work of Weierstrass in any detail, but I would like to point out that Weierstrass' treatment of (1) involved, implicitly, the notion of elementary divisors. In effect, Weierstrass proved [8] that it is possible to transform simultaneously (by an orthogonal substitution) B into the identity I and A into diagonal form *because* the elementary divisors are linear. (Recall that an elementary divisor $(\lambda - a)^k$ of the characteristic polynomial of a matrix corresponds to a k-by-k Jordan

[1] The title of Cauchy's paper [3] has misled historians. Evidence exists which suggests he added it at the last minute, after his encounter with Sturm.

block with eigenvalue $\lambda = a$.) These results were then generalized to bilinear forms and became Weierstrass' theory of elementary divisors [9]. I should point out that Weierstrass introduced the so-called Jordan canonical form in developing his theory, and an immediate consequence of it is the theorem that two matrices (or bilinear forms) are similar iff they have the same Jordan canonical form.

The above-mentioned work of Weierstrass spanned the years 1858–1868. A word is in order concerning the motivation behind it. We are all familiar with Weierstrass' work in analysis, which was characterized by his concern for rigor and hence for the foundations of analysis. Certain generalities seem to have been drawn from this, namely that a concern for rigor comes at the end of a mathematical development, after the "creative ferment" has subsided, that rigor in fact means rigor mortis. Weierstrass himself provides a good counterexample to this generality, for all his work on the spectral theory of forms was motivated by a concern for rigor, a concern that was vital to his accomplishments.

Weierstrass was dissatisfied with the kind of algebraic proofs that were commonplace in his time. These proofs proceeded by formal manipulation of the symbols involved, and no attention was given to the singular cases that could arise when the symbols were given actual values. One operated with symbols that were regarded as having "general" values, and hence such proofs were sometimes referred to as treating the "general case", although it would be more appropriate to speak of the generic case. Generic reasoning had led Lagrange and Laplace to the incorrect conclusion that, in their problems, stability of the solutions to the system of linear differential equations required not only reality but the nonexistence of multiple roots. (Hence their problem had seemed all the more formidable!) In fact, Sturm who was the first to study the eigenvalue problem (1) proved among other things the "theorem" that the eigenvalues are not only real but distinct as well. His proof was of course generic, and he himself appears to have been uneasy about it; for at the end of his paper he confessed that some of his theorems might be subject to exceptions if the matrix coefficients are given specific values. Cauchy was much more careful to avoid what he called disparagingly "the generalities of algebra," but multiple roots also proved problematic for him. As he realized, his proof of the existence of an orthogonal substitution which diagonalizes the given quadratic form depended upon the nonexistence of multiple roots. He tried to brush away the cases not covered by his proof with a vague reference to an infinitesimal argument that was anything but satisfactory.

It was to clear up the muddle surrounding multiple roots by replacing generic arguments with truly general ones that Weierstrass was led to create his theory of elementary divisors. Here is a good example in which a concern for rigor proved productive rather than sterile. Another good example is to be found in the work of Frobenius, Weierstrass' student, as I shall shortly indicate.

So far I have indicated some features of the history of the theory of matrices which show that the significance of Cayley's memoir on matrices of 1858 has been grossly exaggerated. But I do not intend to imply that Cayley played no role whatsoever. Indeed he did, but neither the nature of that role nor the motivation that

led Cayley to write his 1858 memoir has been correctly partaged by historians. Cayley's role and the motivation for his development of the symbolical algebra of matrices are linked with a problem of considerable historical importance for the theory of matrices. The problem is this: Given a nonsingular quadratic form in n variables, determine all the linear substitutions of the variables that leave the form invariant. For reasons that will be clear momentarily, I shall refer to this as the Cayley-Hermite problem.

The problem originated in a paper that Cayley wrote in 1846—and fortunately published in Crelle's Journal. He had read a paper in Liouville's journal by Olinde Rodrigues in which the latter showed among many other things that the 9 coefficients of a linear transformation of rectilinear axes could be expressed rationally in terms of three parameters. This was the period—the early 1840's—when Cayley was preoccupied with learning and applying the theory of determinants, and he showed that Rodrigues' result could easily be established using determinants and, in fact, extended to n variables. Cayley's general result was that if λ_{rs} is a system of coefficients such that $\lambda_{rr} = 1$ and $\lambda_{sr} = -\lambda_{rs}$, then the system of coefficients α_{rs} defined by

$$\alpha_{rs} = 2D_{rs}/D - \delta_{rs}, \quad D = |\lambda_{rs}|, \qquad D_{rs} = \partial D/\partial\lambda_{rs}, \tag{2}$$

has the Lagrangian orthogonality properties. Thus the coefficients α_{rs} of the orthogonal substitution are expressible as rational functions of $n(n-1)/2$ parameters—the λ_{rs}.

Expressed in modern symbolism, Cayley's solution can be written as

$$U = 2(I+S)^{-1} - I = (I+S)^{-1}(I-S), \quad U = (\alpha_{rs}), \qquad I+S = (\lambda_{rs}). \tag{3}$$

The significance of Cayley's solution was that a succinct symbolical representation can be given if *both* the operations of addition and multiplication are employed. I already pointed out that in the late 1840's and early 1850's Eisenstein and Hermite recognized the possibility of a symbolical algebra of linear substitutions under addition and multiplication. But in their work they had—or saw—no occasion to make any symbolical use of the addition of substitutions.

In his 1846 paper Cayley himself introduced no new symbolism. The notation of determinants provided a very succinct form for the solution, as (2) indicates. (Cayley did not, however, use Kronecker deltas.) The matter would probably have ended there were it not for Hermite and his interest in the theory of numbers. In the course of pursuing his research on the arithmetical theory of ternary quadratic forms Hermite had occasion to pose and solve an algebraic problem: Determine all the substitutions of a nonsingular ternary form which leave it invariant [**10**]. Hermite seems to have been aware of Cayley's paper of 1846 which can be interpreted as a solution to the problem when the form is $x^2 + y^2 + z^2$. Although Hermite gave another proof for the more general problem, he used Cayley's idea of generating solutions from skew symmetric systems and he also generalized the result to n variables.

Hermite, however, left his solution in a somewhat incomplete form in the sense

that he did not explicitly write down the coefficients of the solutions to the problem except in the case of 2 variables. He could have written the solution down in succinct explicit form had he thought to employ his symbolical algebra of linear substitutions, but this thought did not occur to him.

It did occur, however, to Cayley, who responded with a paper in 1855 in which he showed that Hermite's solution could be written using the composition of matrices:

$$(X, Y, Z, \cdots) = \left(\begin{vmatrix} a, & h, & g & \cdots \\ h, & b, & f & \cdots \\ g, & f, & c & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}^{-1} \begin{vmatrix} a, & h - \nu, & g + \mu, & \cdots \\ h + \nu, & b, & f - \lambda, & \cdots \\ g - \mu, & f + \lambda, & c, & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} \right.$$

$$\left. \cdot \begin{vmatrix} a, & h + \nu, & g - \mu, & \cdots \\ h - \nu, & b, & f + \lambda, & \cdots \\ g + \mu, & f - \lambda, & c, & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}^{-1} \begin{vmatrix} a, & h, & g & \cdots \\ h, & b, & f & \cdots \\ g, & f, & c & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} \right) (x, y, z, \cdots). \tag{4}$$

In modern notation

$$X = A^{-1}(A - S)(A + S)^{-1}Ax. \tag{5}$$

There was nothing new about expressing results using the composition of matrices; Eisenstein and Hermite had been doing it regularly. Notice also that Cayley, like Eisenstein and Hermite, used only composition, not addition, of matrices. Cayley's originality consisted in seeing a new application for such symbolism, although in 1855, he himself did not develop the symbolism and its application any further than (4). As in the case of Eisenstein and Hermite, the symbolism was primarily used to express succinctly results obtained by other means.

The Cayley-Hermite problem was, however, conducive to the fuller development of the algebra of matrices, and Cayley undertook this in his 1858 memoir on matrices. The importance of the Cayley-Hermite problem in motivating the 1858 memoir is confirmed by the fact that Cayley actually wrote two companion memoirs in 1858; the second dealt with the Cayley-Hermite problem and treated it in terms of the more fully developed notation. (Cayley is thus able to express the solution in a form similar to (5).) Although the 1858 memoirs remained unknown for over 20 years, Cayley's papers of 1846 and 1855 were published in Crelle's journal, and through them Cayley did exert some influence.

Before leaving Cayley, I must point out a characteristic of his mathematics. I have already stressed the importance in the history of the theory of matrices of distinguishing between form—the symbolical algebra of matrices—and substance. This distinction is especially meaningful in connection with Cayley's papers on matrices, for they are primarily on the formal level and lacking in much substance. The sole theorem contained in the 1858 memoir is that a matrix satisfies its characterstic equation—a theorem immediately suggested by the symbolism. Cayley, however, did not prove it. He gave a computational verification for two-by-two matrices, assured his readers that he had also verified the computations for 3-by-3 matrices and added: "I have not thought it necessary to undertake a formal proof of the theorem in the general case of a matrix of any degree." This reflects not only

Cayley's lack of interest in proofs, where inductive evidence seemed convincing, but also his failure to realize that his symbolical algebra of matrices makes it possible to give a simple general proof. Cayley never seemed to realize fully the power of the symbolical algebra of matrices as a method of reasoning.

The only other general theorem in Cayley's 1858 memoir is the assertion that "the general expression" for the matrices which commute with a given matrix M are polynomials in M of degree $\leqq n-1$. This theorem is, of course, literally false; it is true when the minimal polynomial is the same as the characteristic polynomial, as Frobenius was to prove. But Cayley failed to introduce such distinctions in his extremely vague proof. It was typical of the generic level of reasoning in algebra that Weierstrass had just begun to criticize. Incidentally, this very problem of determining the linear substitutions that commute with a given one was to motivate Camille Jordan, 10 years later, to introduce the Jordan canonical form [**11**], [**12**]. Cayley, however, failed to do what Jordan did.

The solutions that Cayley and Hermite had given to the Cayley-Hermite problem were also generic; they had not obtained all possible solutions to the problem. Generic proofs were the order of the day in the 1850's, and no one raised any objections. The Cayley-Hermite problem thus sank into oblivion where it perhaps would have remained had it not once again been for the theory of numbers. Just as Hermite's interest in the arithmetical theory of ternary forms had reawakened Cayley's interest in the Cayley-Hermite problem, so now in the early 1870's it was Paul Bachmann's interest in ternary forms that led him to re-examine Hermite's solution to the Cayley-Hermite problem and to discover its completeness.

Bachmann's observations brought forth a reply from Hermite in which he patched up his earlier solution to cover the singular cases—but only for the ternary case. Also Bachmann's colleague at Breslau, Jacob Rosanes, attempted to deal with the n-variable case by making use of the fact, observed by Hermite and Cayley concerning their solutions, that if λ is a characteristic root so is $1/\lambda$. Rosanes' results were, however, incomplete, especially because he could not handle the case of multiple roots.

Here then was a good "Weierstrassian" problem, a challenge similar to that faced earlier by Weierstrass. This time the challenge was taken up by Frobenius, who had received his doctorate under Weierstrass in 1870. The result was a 63-page paper which was published in Crelle's journal in 1877. Let me explain why it was so long. Frobenius saw that to provide a rigorous and elegant solution to the Cayley-Hermite problem and the related questions raised by Rosanes' paper, it was desirable to fuse together the spectral theory he had learned from Weierstrass and Kronecker with the symbolical algebra of linear substitutions. To treat the Cayley-Hermite problem Frobenius in effect composed a masterful and substantial monograph on the theory of matrices—or forms as he called his symbols. In its pages he convincingly demonstrated the power of the new symbolical method of reasoning when used in conjunction with spectral theory.

Frobenius' paper thus represents an important landmark in the history of the theory of matrices, for it brought together for the first time the work on spectral

theory of Cauchy, Jacobi, Weierstrass and Kronecker with the symbolical tradition of Eisenstein, Hermite and Cayley. I should add that Frobenius' role in the history of the theory of matrices goes beyond what I have indicated. He also established the importance of matrix theory in important new areas of mathematics, such as hypercomplex numbers and group representations.

Although I have stressed Frobenius' contributions, I trust that my necessarily incomplete historical sketch has been sufficient to indicate that the history of matrix theory involved the efforts of many mathematicians, that it was indeed an international undertaking.

References

1. G. Eisenstein, *Allgemeine Untersuchungen über die Formen dritten Grades mit drei Variabeln, welahe der Kreistheilung ihre Entstehung verdanken*, J. Reine Angew. Math. **28** (1844), 289–374.

2. ———, *Ueber die Vergleichung von solchen ternaren quadratischen Formen, welche verschiedene Determinaten haben*, Verhandlungen Akad. Wiss. Berlin, **1852**, 350–389. See esp. p. 354.

3. A. L. Cauchy, *Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes*, Exer. Math. **4** (1829) = Oeuvres (2) **9**, 174–195.

4. P. S. Laplace, *Mémoire sur les variations séculaires des orbites des planètes*, Mém. Acad. Sci. **1787** (1879) = Oeuvres **11**, 295–306.

5. J. L. Lagrange, *Nouvelle solution du problème du rotation d'un corps de figure quelconque qui n'est animé par aucune force accélératrice*, Nouv. Mem. Acad. Berlin **1773** (1775) = Oeuvres **3**, 577–616.

6. ———, *Mechanique analitique*, Paris, 1788, Pt. II, Sect. 6.

7. A. L. Cauchy, *Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des tranpositions opérées entre les variables qu'elles renferment*, J. École Polytech. **10** (1815) = Oeuvres (2) **1**, 91–169.

8. K. Weierstrass, *Ueber ein die homogenen Functionen zweiten Grades betreffendes Theorem*, M'ber. Akad. Wiss. Berlin **1858** = Werke **1**, 233–246.

9. ———, *Zur Theorie der quadratischen und bilinearen Formen*, M'ber. Akad. Wiss. Berlin **1868** = Werke **2**, 19–44.

10. C. Hermite, *Sur la théorie des formes quadratiques ternaires indéfinies*, J. Reine Angew. Math. **47** (1854) = Oeuvres **1**, 193–199.

11. C. Jordan, *Sur la résolution algébrique des equations primitives de degré p^2 (p etant premier impair)*, J. Reine Angew. Math. **13** (1868), 111–135.

12. ———, *Traité des substitutions et des équations algébriques*, Paris, 1870, p. 114ff.

BOSTON UNIVERSITY
BOSTON, MASSACHUSETTS 02215, U.S.A.

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Science as Handmaiden of Mathematics

Geoffrey Matthews

This is a survey report, and a limited one at that. It refers to teaching at school level, the whole range from nursery to college, but mainly to children aged about 11 to 16. These children go from class to class during the school day, and if the mathematics teacher and the science teacher tell different stories, or even similar stories at different times and in different languages, the children can become confused. There is a need, therefore, at least for the mathematics and science staffs to speak to each other, and this dialogue will be more peaceful, and even constructive, if interdisciplinary material is available for them to work on together. Some work has been done elsewhere in this field, notably by Max Bell in Chicago, but the present paper describes only some activities in my own country.

There are four main traditional areas of disagreement between science and mathematics teachers:

(1) *Notation.* Mathematicians brought up on Newton's second law of motion as $\boldsymbol{P} = m\boldsymbol{f}$ can be disconcerted to find the physics man propounding $\boldsymbol{F} = m\boldsymbol{a}$. This is not a serious problem; after all, $\boldsymbol{F}$ is for Force and $\boldsymbol{a}$ for acceleration, and anyway mathematicians are used to a bizarre range of notations (otherwise they would not be able to teach calculus).

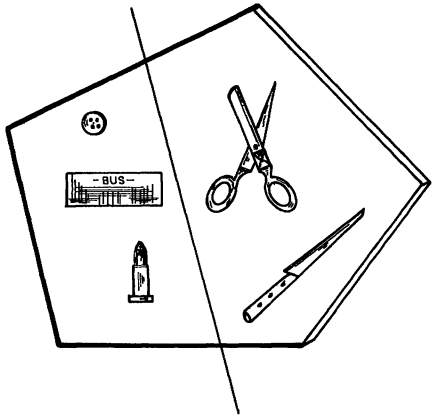
(2) *Timing.* Each year when I was a teacher of 17-years-olds, the class expressed surprise when I set up the differential equation for simple harmonic motion. 'But, Sir, we've just been doing this in physics, with a little man going round a circle dropping ping-pong balls onto a diameter.' I would express suitable surprise. But in less happy schools, the science teacher is forever attacking his mathematical colleague; 'Why haven't they done logarithms?', 'Why haven't they done π?' The answer may be (i) 'They're not ready for it', (ii) 'Sorry, I'll do it tomorrow', or even

(iii) 'We've been doing it for the past fortnight', but the question should not have been necessary.

(3) *Respectability.* When I once asked a science teacher why his colleagues wrote 'Sin $\theta = \theta$' as if it were an identity, he replied 'Well it is generally, isn't it?' (Alright, he was pulling my leg.)

(4) *Style.* Quote from a chemistry teacher: Weight of a beaker + X = 38.55 g. This sort of hybrid sentence of course makes us all wince, particularly the use of ' = ' for something which in the nature of things must be approximate. There is something to be said for being a bit fussy over this; see for example [8] where it is 'proved' in 16 moves that a fly weighs the same as an elephant using ' = ' to mean 'weighs the same as' in the sense of balancing on a pair of scales as near as the eye can judge.

A child of 2 understands the true nature of equivalence.



After turning out mother's handbag, he partitions the set of objects into *things he likes* (button, bus-ticket, lipstick) and *things he does not like* (scissors, knife). If he picks up the button and replaces it, it will still go into the same subset, namely *things he likes* (reflexive property) and the symmetric and transitive properties are equally readily verified; the 2-year-old has had no trouble in establishing the well-known result that

every partition of a set induces an equivalence relation

though he is unlikely to use those words.

When he goes to school some 3 years later, he will find no conflict at first between mathematics and science; they are both being taught by the same teacher (who may not have a special taste for either). Science here is quite happy as handmaiden to

mathematics; for example, an investigation into 'things that float and things that sink' is giving further experience towards the abstraction quoted above ('every partition ···').

At this level, it is indeed difficult to distinguish between mathematics and science. The Schools Council Project 'Science 5—13' for example has produced a publication *Ourselves* [12], largely concerned with pictorial representation of a type which could well have been marketed as mathematics. Again, my wife has devised a simple form of recording progress for young children [10] in which some of the maths and science concepts are difficult to distinguish.

Before turning to the 11 to 16 year olds, I will deal briefly with 'advanced level', i.e., for 16 to 18 year olds. Here the Royal Society has taken remarkable initiative and groups of physicists, chemists and biologists have been working to determine mathematical needs. The physicists have produced a report [11] and the others are on the way. The work has been undertaken in a friendly spirit, but mathematics is considered firmly in its traditional role as handmaiden of science. Is it just possible that the roles may sometimes be reversed even at this level? Two examples must suffice. First, the Nuffield advanced level physics course includes an obviously very simplified account of the Schrödinger equation applied to the hydrogen atom. It may well be debated how much this means in terms of physics (it seems to me a quite admirable preview but not everyone agrees with this), but certainly the students are intrigued and motivated to explore solutions of the differential equation concerned both with and without a computer. (See Matthews and Lewis [9].) This is a good lead-in to numerical methods in general. My second example is contained in a small book *Growth and Decay Models* by my colleague Bob Lewis [4]. This pioneer work starts by looking at a variety of growth and decay situations (population, bacteria, compound interest, radioactive decay, etc.) and gradually building up the mathematics required, i.e., a study of the exponential function. The second part investigates the efficiency of the mathematical model of various situations (rumour-spreading, biological and economic growth, chemical decomposition, capacitor discharge in electrical circuits). The value of this exercise lies not so much in the handling of one particular topic in science as the realisation of the power of mathematics and the capacity of the student to transfer his mathematical expertise from one field to another.

Capacity to think for oneself and transfer knowledge and skills to new situations must be rated very high among the aims of mathematical education; no other discipline affords such opportunities. Our central problem is: 'How, then, can science help with this aim?'

I now turn to our 11 to 16 year olds and start with Project SCISP (Schools Council Integrated Science Project). This project is for abler children of about 14 to 16 years and is remarkable for being conscious not only of the underlying mathematics but also the social implications of science teaching. The following is an extract from one of their first publications [13]:

Penguin	mass/kg.	approximate temperature range of habitat(s) / °C
Emperor	29.15	−10 to 0
King	16.0	0 to +7
Adelie	5.0	−10 to +5
Magellanic	5.0	0 to +12
Chinstrap	4.5	−5 to +5
Peruvian	4.3	+6 to +16
Rockhopper	2.8	+5 to +10
Blackfooted	2.8	+10 to +17
Northern blue	1.8	+12 to +19
Southern blue	1.8	+8 to +15

What pattern is clearly shown by the information? Suggest an explanation for this pattern. Devise a way to test your explanation using ordinary laboratory apparatus as models.

Clearly the penguins get larger the nearer they get to the South Pole and the problem has to do with the ratio of surface area to volume. It is of pedagogical interest to examine other ways of presenting the problem to children:

1. Investigate penguins.
2. Investigate the size of penguins.
3. Do the sizes of penguins vary with distance from the South Pole?
4. Do penguins get bigger the nearer you get to the South Pole?
5. Why do penguins get bigger the nearer you get to the South Pole?
6. (i) *Cubes*

Length of Side	Volume	Surface Area	SA / V
1	1	6	6
2	8	24	3
3	27	54	2
4			
5			

Complete this table.

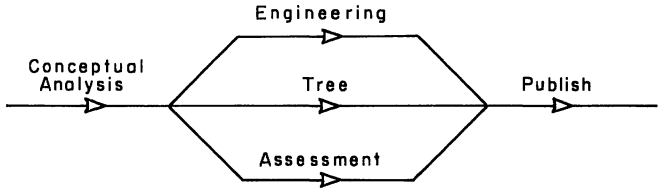
(ii) The ratio 'surface area/volume' affects your capacity to keep warm (your teacher will explain why). As you approach the South Pole, it gets colder; research has shown that penguins get bigger the nearer you get to the South Pole. Can you use the above table to help you explain why?

Here we have the whole range from the completely open-ended (1) to the highly-structured, giving the game away completely (6). But whichever approach is used, it is an improvement on the mathematics textbook which starts with a rectangular lawn with a dull path all round it (and neither access nor flowerbeds) and ends with a mad sculptor who has made his model the wrong size and of the wrong

material and wishes to enlarge it by some outlandish factor. Here is surely a case where science can help mathematics, through motivation.

Such thoughts gave rise to Project SUM (Science Uses Mathematics), now in its trial stages. At the Centre for Science Education, Chelsea College, University of London, where I work, we had made several attempts to get mathematics and science teachers together, but the result was invariably argument about notation, timing, etc., as described earlier. It therefore appeared necessary to produce some interdisciplinary materials so that they had something they could work on together. (See Dudley [1], [2], [3] and Malpas [5], [6], [7].) As a result, a number of 'modules of work' are in course of preparation, the main themes being *Rates, Ratio* and *Measurement.* I will illustrate the flavour of these by reference to such a module called 'Indices and Molecules'. This has roughly the dimensions of an airline ticket. The first part consists of bound-in notes for the teacher and the second a set of loose cards for issue to a group of say 3 or 4 pupils. One sequence consists of dissolving a gram of crystals of potassium permanganate in a litre of water, removing 1 cm^3 of the coloured liquid, mixing this with more water and so on. This gives a genuine reason for introducing negative indices, so that the two things —experiment and mathematics—are going on together, and the science man need no longer growl 'Why haven't they done indices?'. Later in the same module, there is a sequence on the so-called oil-drop experiment (a drop of oil forms a patch on a surface of water, so that by estimating its volume and the area it then covers, some idea can be grasped of a molecule and its possible length). Here again the subjects are in harmony instead of 'Why haven't they done π?'.

So much for Project SUM. But underlying the conflicts between the teachers there is the fundamental fact that little research has been done on the hierarchy of concepts in either mathematics or science at this secondary level. We have recently received a grant from our Social Science Research Council for a five-year project to investigate this problem. There is only time to give a diagram of our plans:



After preliminary analysis of conceptual difficulties in various contemporary courses, the work splits into three branches:

(i) *Tree.* This will be an attempt to make a partially-ordered 'tree of knowledge' to suggest preferred orders of treatment of topics.

(ii) *Engineering.* By simply looking at the topics, ways of reducing difficulties (e.g., transferring from 'formal' to 'concrete' level) will suggest themselves.

(iii) *Assessment.* Methods will be devised for production of reliable tests for the use of teachers to check on children's levels of understanding.

Work in all these branches has begun this year.

Of course there is a lot more work to be done in forging links at this level between mathematics and other subjects—geography, economics, etc. Pupils will be really educated when they can appreciate mathematical applications to subjects other than their own chosen speciality—but we are a long way at present from that goal.

References

1. B.A.C. Dudley, *Teaching size and form in biology*, J. Biol. Ed. **5** (1971), 187–191.
2. ———, *The mathematical basis of Mendelian genetics*, Int. J. Math. Educ. Sci. Technol. **4** (1973), 193–204.
3. ———, *The mathematical basis of population genetics*, Int. J. Math. Educ. Sci. Technol. **4** (1973), 329–333.
4. R. Lewis, *Growth and decay models*, Penguin, Baltimore, Md., 1972.
5. A. J. Malpas, *Mathematical interpretation of experimental data in schools*, Int. J. Math. Educ. Sci. Technol. **4** (1973), 263–270.
6. ———, *Mathematics and science in the secondary school*, Education in Science **52** (1973), 27–32.
7. ———, *Combining masses, ratios and chemical formulae*, School Science Review **54** (1973), 188, 542–545.
8. Geoffrey Matthews and John Bausor, *Models and measurement*, Int. J. Math. Educ. Sci. Technol. **3** (1972), 275–279.
9. Geoffrey Matthews and Bob Lewis, *Quanta for the quick*, Math. Gaz. **58** (1974), 8–14.
10. J. Matthews, *Progress records for young children*, John Murray and W. and R. Chambers, London, 1974.
11. Royal Society/Institute of Physics Joint Committee for Physics Education, *The mathematical needs of A-level physics students*, Physics Bull. **24** (1973), 226–228.
12. Science 5–13 Project, *Ourselves*, Macdonald, for Schools Council, London, 1973.
13. Schools Council Integrated Science Project (SCISP), *Patterns. I. Building blocks*, Penguin, Baltimore, Md., 1973.

CHELSEA COLLEGE, UNIVERSITY OF LONDON
LONDON, ENGLAND SW6 4HR

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

How to Understand and Teach the Logical Structure and the History of Classical Thermodynamics*

C. Truesdell

Prologue. I do not think it is possible to write the history of a science until that science itself shall have been understood, thanks to a clear, explicit, and decent logical structure. The exuberance of dim, involute, and undisciplined historical essays upon classical thermodynamics reflects the confusion of the theory itself. Thermodynamics, despite its long history, has never had the benefit of a magisterial synthesis such as that which Euler gave to hydrodynamics in 1757 or that which Maxwell gave to electromagnetism in 1873; the expositions in the works of discovery in thermodynamics stand a pole apart from the pellucid directness of the notes in which Cauchy presented his creation and development of the theory of elasticity from 1822 to 1854. Thermodynamics was born in disorder if not confusion, and there the common presentations of it have remained.

Here I provide a simple logical structure for the classical thermodynamics of homogeneous fluid bodies. I take as primitive *just those quantities which every early author took as primitive*, namely:

t time,

and, associated at the time t with a fluid body:

$V(t)$ volume ($V(t) > 0$),

$\theta(t)$ temperature (by convention restricted to the range $0 < \theta(t) < \infty$),

$p(t)$ pressure ($p(t) > 0$),

$Q(t)$ heating

—these and no other. I deal only with classical thermodynamics, to which the term

*The research reported here was supported in part by grants of the United States National Science Foundation to the Johns Hopkins University. A full account, with proofs, many corollaries, applications to cases, and historical analyses, will appear in the Archive for Rational Mechanics and Analysis

"reversible" was attached in later years, and only with two independent variables, which I choose to interpret as being the volume and temperature of a homogeneous fluid body. Finally, I employ no mathematical tools that were not already available in the 1820's, and every thermodynamic concept I use was introduced by 1854 at the latest. The entire theory is based on the systematic, elementary, and rigorous use of Carnot cycles. The central axiom is Carnot's own: *The motive power of a Carnot cycle is a function of its two temperatures θ^+ and θ^- and of the heat absorbed at the higher temperature.*

My scope is threefold:

1. *Conceptual.* For those already expert in thermodynamics, to show how all the concepts can be derived from simple and natural assumptions about heat engines, developed by simple and rigorous mathematics, with no "physical" arguments and no appeal to metaphysics.

2. *Pro-Historical.* For those who would study the pioneer researches, by logical analysis to reveal the features of principle *common* to Carnot's thermodynamics and Clausius', and to discern the irreconcilable *differences* of principle between them, never before carded of metaphysics and rhetoric.

3. *Pedagogical.* For those who wish to learn a clean elementary thermodynamics so as to teach it to beginners.

Thus the past and the present illuminate each other.

My work owes much to a long memoir of F. Reech, published in 1853. You will recall that Clausius introduced and developed the concept of internal energy of an ideal gas in 1850, that he did not publish his treatment of more general equations of state until 1854, and that only in that year did he introduce the concept of entropy or distinguish between reversible and irreversible processes. Reech's great memoir appeared after Clausius' first note and before the second one.

Frédéric Reech was born in 1805 at Lampertsloch, Alsace. He studied at the École Polytéchnique and became director of testing and development of marine steam engines at the French naval port of Lorient. Later he became director of the École d'Applications du Génie Maritime. He published a book on engines in 1844 and a general course of mechanics in 1852. When Regnault, perhaps influenced by Joule, began to doubt that heat was an indestructible substance, Reech undertook a logical examination of the whole matter, starting from the most general form of Carnot's axioms, axioms Carnot himself had applied only subject to further restrictions. Reech was ready to reject the axiom that heat is conserved but not ready to adopt the axiom that heat and work are uniformly interconvertible.

Reech's work on thermodynamics passed unread and has been scarcely noticed to this day. Reech's axioms are exactly those Carnot had stated a quarter of a century earlier, apart from the assumption that heat is conserved. Carnot's own treatment is inextricably entwined with that assumption, which Reech cast aside. Reech attempted to develop all the consequences of Carnot's axioms, but he was not skilled enough in elementary analysis to do so. Indeed, he did construct a theory that subsumes both Carnot's and Clausius'; he did introduce and exploit generalizations of the entropy and the thermodynamic potentials before Clausius and

others introduced the special cases now familiar; and he did derive, in generalized form, every identity, definition, or other formal relation in the classical part of the subject. Unfortunately, the conditions he obtained, although necessary, are not sufficient. They are too general in that they leave unrestricted certain functions which the axioms in fact restrict severely.

The theory I present now is based on Reech's axioms. It completes his work by exhausting the consequences of those axioms. It is a prolegomenon to the history of thermodynamics: an outline of the theory such as could have been written in 1854, had thermodynamics been blessed with an Euler, a Cauchy, or a Maxwell. That the discoverers did not find proofs so simple as mine and that they failed to draw some central conclusions which in fact do follow easily from their assumptions have nothing to do with the subject itself or the date at which it was developed, the organization of science, the social or economic circumstances of the scientists involved or not involved in its early development, or any of the other extranea now popularly invoked so as to smother thought and truth and to exalt what might well be called "group science".

A list of the definitions, axioms, and main theorems follows. The proofs, which are omitted here, use a mathematical tool which the authors of textbooks, following the works of discovery, seem unwilling if not unable to apply. That tool is integral calculus.

I. Calorimetry.

DEFINITION 1 (CARNOT, 1824). A *process* is a pair of continuous and piecewise continuously differentiable functions V and θ, defined over a closed interval $[t_1, t_2]$ of times t, $t_1 < t_2$, and having as values $V(t)$ and $\theta(t)$ a volume and a temperature, respectively, of some given body.

DEFINITION 2. A process is *isothermal* if $\theta(t) = \text{const.}$ for all $t \in [t_1, t_2]$.

DEFINITION 3. Let a process have been given, and let p be an integrable function in $[t_1, t_2]$. Set $L \equiv \int_{t_1}^{t_2} p\dot{V}\,dt$. If $p(t)$ is the pressure exerted upon the given body at the time t, then the number L is the *work done* by the body subjected to the pressure p in the given process.

DEFINITION 4. Let a process have been given, and let Q be an integrable function in $[t_1, t_2]$. Set

$$C \equiv \int_{t_1}^{t_2} Q\,dt, \quad C^+ \equiv \frac{1}{2}\int_{t_1}^{t_2} (|Q| + Q)\,dt, \quad C^- \equiv \frac{1}{2}\int_{t_1}^{t_2} (|Q| - Q)\,dt,$$

so that $C = C^+ - C^-$. If $Q(t)$ is the heating of the given body at the time t, then the numbers C, C^+, and C^- are, respectively, the *net gain of heat*, the *heat absorbed*, and the *heat emitted* by the body in the process.

Definitions 5—8 define the terms *path*, simple *path*, *reverse* $-\mathscr{P}$ of a path $\mathscr{P}$, and *cycle*.

AXIOM I (EXISTENCE OF A THERMAL EQUATION OF STATE, EULER, 1757). *In a certain nonempty connected open set $\mathscr{D}$ of the positive V-θ quadrant, the pressure p acting upon a given body is determined by the volume and temperature of that body:*

$$p = \varpi(V, \theta).$$

The function ϖ is continuous and has continuous partial derivatives $\partial\varpi/\partial V$ and $\partial\varpi/\partial\theta$; also $\partial\varpi/\partial V < 0$.

EXAMPLE. A body is an *ideal gas* if $\varpi = R\theta/V$, R being a positive constant.

THEOREM 1 (FIRST REVERSAL THEOREM). *The work L done by a body in a process depends only upon the path $\mathscr{P}$ of that process and is given by the line integral*

$$L = L(\mathscr{P}) = \int_{\mathscr{P}} \varpi(V, \theta)\, dV.$$

Also $L(-\mathscr{P}) = -L(\mathscr{P})$.

COROLLARY 1.1. *If $\mathscr{C}$ lies in a simply connected part of $\mathscr{D}$ and is a simple cycle, so oriented that the region $\mathscr{A}$ it encloses lies on its right-hand side, then*

$$L(\mathscr{C}) = \iint_{\mathscr{A}} \frac{\partial\varpi}{\partial\theta}\, dV\, d\theta.$$

AXIOM II (DOCTRINE OF LATENT AND SPECIFIC HEATS). *There are functions Λ_V and K_V, defined over $\mathscr{D}$, continuous and having continuous partial derivatives there, such that in any process (V, θ) the heating Q of the body has the form*

$$Q = \Lambda_V(V, \theta)\dot{V} + K_V(V, \theta)\dot{\theta}$$

at all times when $\dot{V}$ and $\dot{\theta}$ exist. Moreover, $K_V > 0$.

THEOREM 2 (SECOND REVERSAL THEOREM). *The net gain of heat in a process depends only on the corresponding path*:

$$C = C(\mathscr{P}) = \int_{\mathscr{P}} (\Lambda_V\, dV + K_V\, d\theta),$$

and $C(-\mathscr{P}) = -C(\mathscr{P})$.

COROLLARY 2.2 (CLAUSIUS, 1850). *If $\mathscr{C}$ is a simple cycle in $\mathscr{D}$, so oriented that the region $\mathscr{A}$ it encloses lies on its right-hand side, then*

$$C(\mathscr{C}) = \iint_{\mathscr{A}} (\partial\Lambda_V/\partial\theta - \partial K_V/\partial V)\, dV\, d\theta.$$

THEOREM 3 (FUNDAMENTAL THEOREM OF CALORIMETRY). $Q = \Lambda_p\dot{p} + K_p\dot{\theta}$; *$\Lambda_p$, the latent heat with respect to pressure, and K_p, the specific heat at constant pressure, are related as follows to Λ_V and K_V*:

$$\Lambda_p = \frac{\Lambda_V}{\partial\varpi/\partial V}, \qquad K_p - K_V = -\Lambda_V \frac{\partial\varpi/\partial\theta}{\partial\varpi/\partial V}.$$

Also

$$(K_p - K_V)Q = \Lambda_V\left[\frac{K_V}{\partial\varpi/\partial V}\,\dot{p} - K_p\dot{V}\right].$$

DEFINITION 10. A process such that $Q = 0$ is *adiabatic*. The path of an adiabatic process is called an *adiabat*.

THEOREM 4 (LAPLACE, 1816–1823, IN PRINCIPLE). *In an adiabatic process, p is a differentiable function of V, and*

$$dp/dV = \gamma\, \partial w/\partial V, \qquad \gamma \equiv K_p/K_V.$$

HISTORICAL REMARK. The position of this theorem in the development, before any relation between heat and work has been assumed, shows that Laplace's calculation of the speed of sound is independent of his assumption that heat is an indestructible substance consisting of molecules that act upon each other through central forces.

COROLLARY 4.1 (LAPLACE AND POISSON, SUFFICIENCY ONLY, IN PRINCIPLE). *Let a body of ideal gas undergo an adiabatic process. Then in order that* $pV^\gamma = \text{const.}$, *or, equivalently,* $\theta V^{\gamma-1} = \text{const.}$; *it is necessary and sufficient that* $\gamma = \text{const.}$ *on the adiabats.*

THEOREM 5 (EMISSION-ABSORPTION ESTIMATE). *Let the cycle* $\mathscr{C}$ *lie in a part of* $\mathscr{D}$ *over which exist a positive continuous function* f *and a differentiable function* H_f *such that* $Q = f\dot{H}_f$. *Then*

$$C^-(\mathscr{C}) \geqq \frac{\min_{\mathscr{C}^-} f}{\max_{\mathscr{C}^+} f}\, C^+(\mathscr{C}).$$

Apart from cases in which $C^+(\mathscr{C}) = 0$ *or* $C^-(\mathscr{C}) = 0$, *equality holds if and only if* f *has a constant value on* $\mathscr{C}^-$ *and a constant value on* $\mathscr{C}^+$, *the two constants not necessarily being the same.*

II. Carnot's general axiom.

DEFINITION 11. A cycle is a *Carnot cycle* if at all of its points where Q exists, one of the following three alternatives holds:

$$\begin{aligned} &\theta = \theta^+, && Q > 0, \\ &Q = 0, && \\ &\theta = \theta^- < \theta^+, && Q < 0. \end{aligned}$$

That is, a Carnot cycle is formed by adiabats connecting two isotherms, and heat is absorbed on the isotherm at higher temperature but emitted on that at lower temperature.

REMARK. The assumptions of smoothness and the inequalities included in the statements of Axioms I and II ensure also that if the adiabats and isotherms of a given Carnot cycle be extended slightly, another Carnot cycle is formed, and that a sufficiently small Carnot cycle is subdivided into smaller ones by any interior adiabats or isotherms.

AXIOM III (CARNOT'S GENERAL AXIOM, 1824). *If* $\mathscr{C}$ *is a Carnot cycle, then*

$$L(\mathscr{C}) = G(\theta^+, \theta^-, C^+(\mathscr{C})).$$

Moreover, the function G *is of such a kind that if the isotherms* $\theta = x$ *and* $\theta = y$ *cut* $\mathscr{D}$, *then*

$$\begin{aligned} &G(x, y, z) > 0 \quad \text{if } x > y \text{ and } z > 0, \\ &G(x, x, y) = 0 \quad \text{if } x > 0 \text{ and } y \geqq 0, \\ &G(x, y, 0) = 0 \quad \text{if } x > y. \end{aligned}$$

If x *and* y *are fixed,* $\lim_{z\to 0+} G(x, y, z) = 0$.

THEOREM 6 (REECH, 1851). *$L(\mathscr{C}) = F(\theta^+, \theta^-)C^+(\mathscr{C})$, and $F(x, y) > 0$ if $x > y > 0$. $F(x, x) = 0$.*

THEOREM 7 (EFFICIENCY THEOREM). *There exist functions g and h such that*

$$F(\theta^+, \theta^-) = \frac{g(\theta^+) - g(\theta^-)}{h(\theta^+)}$$

and

$$h(x) > 0, \qquad g(x) - g(y) > 0 \text{ if } x > y.$$

The function h is unique to within a positive constant factor; when h is determined, g is unique to within an additive constant.

COROLLARY 7.1. *In a Carnot cycle*

$$C^-(\mathscr{C}) = \frac{h(\theta^-)}{h(\theta^+)} C^+(\mathscr{C}).$$

SUBCOROLLARY 7.2 (REECH). *In order that $C^-(\mathscr{C}) = C^+(\mathscr{C})$ in all Carnot cycles, it is necessary and sufficient that $h(\theta) \equiv$ const. Equivalently,*

$$L(\mathscr{C}) = [g(\theta^+) - g(\theta^-)]C^+(\mathscr{C}),$$

and $g(x) - g(y) > 0$ if $x > y$.

HISTORICAL REMARK. Though Carnot stated his general axiom in words, he never used it except in the special case when $G(\theta^+, \theta^-, C^+(\mathscr{C})) = [g(\theta^+) - g(\theta^-)] \cdot C^+(\mathscr{C})$. We may call this case *Carnot's special axiom.* Reech's subcorollary shows that Carnot need not have assumed, as he did, that heat was conserved. He could have proved that as a theorem. Also, the common claims that Carnot "knew" the "Second Law" without knowing the "First Law" are ridiculous. He knew neither. His special axiom contradicts the "First Law", so it cannot be consistent with what is called today the "Second Law".

COROLLARY 7.3. *In order that for all Carnot cycles the work done by a given body shall be one and the same constant multiple of the net gain of heat, that is, $L = J(C^+ - C^-)$, J being a constant, it is necessary and sufficient that for that body $g' = Jh'$.*

COROLLARY 7.4 (CHARACTERIZATION OF CLASSICAL EFFICIENCY). *In order that, for a given body,*

$$F(\theta^+, \theta^-)/J = 1 - \theta^-/\theta^+,$$

it is necessary and sufficient that $g \propto h +$ const., $h \propto \theta$.

THEOREM 8. *The heat-loss function h is continuously differentiable at every θ that corresponds to some point in $\mathscr{D}$. Moreover, h is an integrating factor for $\Lambda_V dV + K_V d\theta$. That is,*

$$\frac{\partial}{\partial\theta}\left(\frac{\Lambda_V}{h}\right) - \frac{\partial}{\partial V}\left(\frac{K_V}{h}\right) = 0.$$

THEOREM 9 (BASIC CONSTITUTIVE RESTRICTIONS OF THERMODYNAMICS). *The function g is continuously differentiable at every θ that corresponds to a point of 𝒟, and*

$$(g'/h)\Lambda_V = \partial\varpi/\partial\theta.$$

The partial derivative $\partial^2\varpi/\partial V\partial\theta$ exists and is continuous throughout 𝒟; $\partial^2\varpi/\partial\theta^2$ exists at all points in 𝒟 on the isotherm $\theta = \theta_0$ if and only if $g''(\theta_0)$ exists. At such points

$$\frac{g'^2}{h}K_V = \int\left[g'\frac{\partial^2\varpi}{\partial\theta^2} - g''\frac{\partial\varpi}{\partial\theta}\right]dV + k,$$

k being a function of θ alone.

III. Energy, efficiency, and potentials.

THEOREM 10. *Let Axiom* II *and the formula for $\partial\varpi/\partial\theta$ given by Theorem* 9 *be assumed, and let g and h be given. Then, whether or not Axiom* III *holds, in any simply connected part of 𝒟 there is a function $E_{h,g}(V, \theta)$, called the internal pro-energy of the body, such that*

$$(g/h)\Lambda_V = \varpi + \partial E_{h,g}/\partial V, \qquad (g/h)K_V = \partial E_{h,g}/\partial\theta.$$

Moreover, if $\dot{E}$ denotes the derivative of $E_{h,g}(V, \theta)$ in the process (V, θ), then $\dot{E} = (g/h)Q - p\dot{V}$.

COROLLARY 10.1 (CONDITION IMPOSED BY THE PRESSURE FUNCTION ON THE PRO-ENERGY). *With a choice of g such that $g > 0$ and $g' > 0$,*

$$\frac{g^2}{g'}\cdot\frac{\partial}{\partial\theta}\left(\frac{\varpi}{g}\right) = \frac{\partial E_{h,g}}{\partial V}.$$

THEOREM 11. *If 𝒞 is a cycle in a simply connected part of 𝒟,*

$$L(\mathscr{C}) = \int_{t_1}^{t_2}\frac{g}{h}Q\,dt = \int_{\mathscr{C}}\frac{g}{h}(\Lambda_V\,dV + K_V\,d\theta).$$

THEOREM 12 (CONVERSE OF THEOREMS 9 AND 7). *Let Axioms* I *and* II *be assumed, and suppose that there are functions g and h such as to relate the constitutive functions ϖ, Λ_V, and K_V in the way asserted by Theorem* 9. *Then in any simply connected part of 𝒟 where $g' \neq 0$ and where $\partial^2\varpi/\partial\theta^2$ exists, Carnot's general axiom holds, and in it G has the form asserted by Theorems* 6 *and* 7.

THEOREM 13.

$$L(\mathscr{C}) \leqq \left[\max_{\mathscr{C}}\frac{g}{h} - \left(\min_{\mathscr{C}}\frac{g}{h}\right)\frac{\min_{\mathscr{C}} h}{\max_{\mathscr{C}} h}\right]C^+(\mathscr{C}).$$

If both g/h and h are monotone nondecreasing and if at least one of them is strictly increasing, the upper bound is achieved by Carnot cycles and by them only.

COROLLARY 13.1 (APPLICATION TO THE CALORIC THEORY). *If $h = 1$, as Carnot's special axiom asserts, then for any cycle 𝒞 with greatest and least temperatures θ^+ and θ^-, respectively,*

$$[g(\theta^-) - g(\theta^+)]C^+(\mathscr{C}) \leqq L(\mathscr{C}) \leqq [g(\theta^+) - g(\theta^-)]C^+(\mathscr{C}).$$

The maximum is achieved if and only if $\mathscr{C}$ is a Carnot cycle having θ^+ and θ^- as the temperatures of its isotherms. The minimum is achieved if and only if $\mathscr{C}$ is the reverse of such a cycle.

COROLLARY 13.2 (APPLICATION TO THE CASE WHEN HEAT IS UNIFORMLY INTERCONVERTIBLE WITH WORK). *If $L(\mathscr{C}) = J[C^+(\mathscr{C}) - C^-(\mathscr{C})]$ for every cycle $\mathscr{C}$, then*

$$L(\mathscr{C}) \leqq J[1 - h(\theta^-)/h(\theta^+)]C^+(\mathscr{C}),$$

and equality is achieved if and only if $\mathscr{C}$ is a Carnot cycle.

DEFINITION 12 (REECH). For a given process of a given body,

$$\Phi \equiv E - gH, \qquad X \equiv E + pV, \qquad Z \equiv X - gH.$$

The functions Φ, X, and Z are the *free energy*, the *enthalpy*, and the *free enthalpy*.

DEFINITION 13. Let g and h be known. Then a function is a *thermodynamic potential* of a given body if it determines the three constitutive functions of that body: ϖ, Λ_V, and K_V.

THEOREM 14 (REECH).

$$-\dot{\Phi} = \varpi(V, \theta)\dot{V} + g'(\theta)H(V, \theta)\dot{\theta},$$
$$\varpi = -\partial\Phi/\partial V, \qquad g'H = -\partial\Phi/\partial\theta,$$
$$\Lambda_V = -\frac{h}{g'}\frac{\partial^2\Phi}{\partial V\partial\theta}, \qquad K_V = -h\frac{\partial}{\partial\theta}\left(\frac{1}{g'}\frac{\partial\Phi}{\partial\theta}\right).$$

Thus Φ is a thermodynamic potential. If θ is a function of V and H, then E may be expressed as a function of V and H and is a thermodynamic potential. If V and θ are functions of p and H, then X may be expressed as a function of p and H and is a thermodynamic potential. If V is a function of θ and p, then Z may be expressed as a function of θ and p and is a thermodynamic potential.

IV. Universal efficiency.

AXIOM IV (UNIVERSALITY). *The function G in Axiom* III *is universal, not constitutive.*

That is, all bodies have the same efficiency in all the Carnot cycles possible for them with given temperatures θ^+ and θ^-.

AXIOM V (KEY EXAMPLE). *Some ideal gas has constant specific heats.*

HISTORICAL REMARK. Although this axiom may seem trivial to the modern student, who is accustomed to taking it for granted, it is the key to classical thermodynamics. The caloric theory contradicts it except in the trivial case when $\gamma = 1$. The reader may see this for himself by comparing Corollary 2.2, which implies for the caloric theory that $\partial\Lambda_V/\partial\theta = \partial K_V/\partial V$, with Theorem 3 after specialization to ideal gases. The pioneers' cry for new experiments so as to determine Carnot's function g was misdirected. Theorem 4 shows that the squared speed of sound, conceived as an adiabatic undulation, cannot be proportional to but greater than the isothermal

squared speed unless $\gamma = \text{const.} > 1$. In Carnot's thermodynamics this condition forbids K_p and K_V to be constants and makes both be decreasing functions of temperature. Carnot's thermodynamics is inconsistent with simple facts of experiment already well known when it was published. That the inconsistency remained unnoticed until 1973 results from Carnot's reluctance to apply clean mathematical reasoning, a reluctance abundantly reflected by every textbook and every historical study of thermodynamics that I have seen.

In forming classical thermodynamics it was not experiment that wanted. It was mathematics.

THEOREM 15 (UNIVERSAL EFFICIENCY). *The functions g and h have the following universal forms:*

$$g \propto h + \text{const.}, \qquad h \propto \theta.$$

COROLLARY 15.1 (THE "FIRST LAW OF THERMODYNAMICS"; EFFICIENCY OF CARNOT CYCLES). *Every body converts heat uniformly into work in any cycle: $L = JC$; each body has an internal energy $E(V, \theta)$ such that $\dot{E} = Q - p\dot{V}$; also*

$$J\Lambda_V - \varpi = \theta^2 \frac{\partial}{\partial \theta}\left(\frac{\varpi}{\theta}\right) = \frac{\partial E}{\partial V}, \qquad JK_V = \frac{\partial E}{\partial \theta};$$

for any body undergoing a Carnot cycle $C^-/C^+ = \theta^-/\theta^+$; the efficiency of a body undergoing a Carnot cycle is given by

$$F(\theta^+, \theta^-)/J = 1 - \theta^-/\theta^+;$$

if θ^- and θ^+ are the least and greatest temperatures of a cycle that is not a Carnot cycle, then the efficiency of that cycle is less than $1 - \theta^-/\theta^+$.

COROLLARY 15.2 (THE "SECOND LAW OF THERMODYNAMICS"). *There exists an entropy $H_\theta(V, \theta)$ such that*

$$\Lambda_V = \theta\, \partial H_\theta/\partial V, \qquad K_V = \theta\, \partial H_\theta/\partial \theta$$

Moreover,

$$\Lambda_V = \theta\left(\frac{\partial \Lambda_V}{\partial \theta} - \frac{\partial K_V}{\partial V}\right), \qquad J\Lambda_V = \theta \frac{\partial \varpi}{\partial \theta}$$

$$J(K_p - K_V) = -\theta\left(\frac{\partial \varpi}{\partial \theta}\right)^2 \Big/ \frac{\partial \varpi}{\partial V}.$$

Epilogue. Reech's axioms refer to Carnot cycles. Thus in the thermodynamics based upon them, the theory of heat engines is not an application of general principles but the source of them. The traditional "First Law" and "Second Law" of thermodynamics for "reversible" processes in fluid bodies appear here as corollaries of theorems on heat engines. Of course, also my proof that a Carnot cycle is more efficient than any other with the same extremes of temperature is mathematical. The common circular and declaratory appeal to the absurdity of creating "something" from "nothing", though it has indeed discouraged many mathematicians and engineers so much as to bring them to contemn classical thermodynamics, by no means implies any defect inherent in that theory.

Thermodynamics, like mechanics and electromagnetism, is a rational science. As such, it may be taught to beginners, provided they master the elements of differential and integral calculus. They will like it better than their teachers, who may be loath to give up the profound metaphysical taboos of the tribes which now possess the field. Such beginners, after they have learnt without starting from something to unlearn, may see lynx-eyed straight on into practical applications, into research upon the modern rational thermodynamics of deformation and reaction, into the history of the mathematical sciences.

JOHNS HOPKINS UNIVERSITY
BALTIMORE, MARYLAND 21218, U.S.A.

Short Communications*

Johan F. Aarnes, *Differentiably induced representations and Frobenius reciprocity*
Johannes Aarts, *Dimension modulo a class of spaces for nonmetrizable spaces*
N. N. Abdelmalek, *An efficient method for the discrete linear L_1 approximation*
Wazir Abdi, *Stability of linear difference-differential equations via bibasic functional equations*
Shingo Abe, *Finite generalized motions in X_4*
J. Abrham, *A note on nonconcave continuous programming*
M. H. Afghahi, *On the Berry-Esseen theorem*
B. Agrawal, *A general orthogonal polynomial*
Michael Aissen, *Combinatorial aspects of subassociative operations*
Charles Akemann, *Computing operator norms in the group algebra of a discrete group*
H. Al-Amiri, *On a linear combination of some expressions in the theory of univalent functions*
M. V. Al-Dhahir, *On the ternary ring of a projective plane*
W. R. Allaway, *The representation of orthogonal polynomials in terms of a differential operator containing their generating function*
Seth Alpert, *Nonorientable graph embeddings*
Charles Amelin, *Sectorial pairs of linear operators*
J. W. Andrushkiw, *A probabilistic method for determining the number of zero terms in the expansion of a determinant*
Roman Andrushkiw, *Quadratic eigenvalue problem*
R. Ansorge, *Kombination zweier Monotoniesatze von Redheffer und J. Schroder zur numerischen Konstruktion von Schrankenfunctionen*
D. U. Anyanwu, *A turning point theory with applications for a first-order ordinary differential equation in Hilbert space*
R. N. Aranake, *Effects of mass transfer and free convection currents on Stokes problem for an infinite vertical plate*
Joseph Arkin, *Magic Latin k-cubes of order n*
F. M. Arscott, *Transform theorems for multiparameter eigenvalue problems*
P. Artzner, *Decomposition of some random positive quadratic forms*
S. P. Arya, *On the locally countable sum theorem*
C. E. Aull, *Metrization of topological spaces with a delta-theta base*
B. Auslander, *A variation on the self-paced method of teaching elementary college mathematics*
Nils A. Baas, *Structural stability of composed mappings*

---

*In the case of more than one author only the name of the person presenting the paper is listed.

P. C. Baayen, *Comprehensive topological transformation groups*
John Baker, *Extension of homeomorphisms over 0-dimensional separable metric spaces*
Catherine Bandle, *On a nonlinear Dirichlet problem*
D. O. Banks, *A Prüfer transformation for the equation of an axially loaded vibrating beam*
Erwin M. Bareiss, *Diffusion and transport theory solutions near interior corners and edges*
R. Barnhart, *A class of simply connected double-toroidal manifolds which are 3-balls*
Harm Bart, *Poles of the resolvent of an operator function*
K. F. Barth, *An annular function whose derivative is not annular*
Lucilla Bassotti, *Sous éspaces invariants pour les opérateurs differentiels à coefficients constants*
B. Beauzamy, *Convexifying operators between Banach spaces*
Ilse Becker, *A geometric interpretation of group extensions*
John Beekman, *Compound Poisson processes as modified by Ornstein-Uhlenbeck processes*
J. Behboodian, *On symmetric distributions*
S. J. Benkoski, *The probability that k integers are relatively r prime*
H. L. Bentley, *Alexander cohomology for nearness spaces*
Christian Berg, *Fourier transformation of positive definite measures*
F. Berquier, *Differential calculus over modules*
William A. Beyer, *The asymptotic form of partition functions of linear chain molecules: An application of renewal theory*
P. Bhatnagar, *Transport properties of a collisionless plasma from finite-Larmor-radius equations in a general magnetic field*
M. C. Bhattacharjee, *Iterates of the initial distribution of a stationary renewal process*
K. D. Bierstedt, *Inductive limits of weighted spaces of continuous (resp. holomorphic) functions*
Denis Blackmore, *On genericity in dynamical systems*
D. Bobrowski, *On the nonoscillatory solutions of a certain differential inequality*
Klaus Boehmer, *A generalized Fuchsian theory and applications*
V. Bohun-Chudyniv, *Applications of quasigroups and quasiloops for constructing sets of mutually orthogonal Latin squares of order bigger than 2 and not equal to 6*
Juergen Bokowski, *Lattice polyhedra and Wills' conjecture*
Ethan Bolker, *Combinatorial topology solves some combinatorial problems*
D. D. Bonar, *On four questions raised by Erdös*
E. Boorman, *Beta operations and a unigeneration theorem for the Burnside ring of the symmetric group*
John F. Bowers, *Soluble groups of finite rank*
Ivan Bozovic, *Undecidability and new demands imposed on physical theories*
Willem Brandenburg, *Mathematical models in education*
Alfred Brauer, *A lower bound for the greatest root of an irreducible nonnegative matrix*
George Brauer, *Interpolation of a class of analytic functions*
Robert Braun, *On a theorem of W. Koppelman about Cauchy-Fantappié forms*
Izak Broere, *An abstract theory of ideals*
Arne Broman, *Pythagorean twins*
Lawrence Brown, *Convexity in measure algebras and semigroups of partial isometries*
Margaret Brown, *Cognitive demand levels in secondary mathematics*
Robert F. Brown, *Fixed points of endomorphisms of compact groups*
F. Brownell, *Time expectation value asymptotics and pseudo-eigenvalues*
J. Bruning, *A global estimate of the spectral function of an elliptic operator*
H. D. Brunk, *Uniform inequalities for conditional p-means given σ-lattices*
Lothar Budach, *Solution of the labyrinth problem for finite automata*
Gerhard Burde, *Representations of knot groups and link groups*
Stefan Burr, *Extremal Ramsey theory for graphs*
G. Buyukyenerel, *Projective group representations and fiber bundles*
Colin M. Campbell, *Some results on Fibonacci type groups*
V. Cantoni, *Lorentz-invariant operators on continuously decomposable fields*

Renate Carlsson, *The Wedderburn principal theorem for associative triple systems*
D. S. Carter, *Collinearity-preserving transformations of Desarguesian planes*
John Casti, *Matrix Riccati equations, optimal control theory, and radiative transfer*
Marshall Cates, *Schur's theorem and a related problem*
Jean-P. Caubet, *Quantization and space-time Brownian motion*
Jack G. Ceder, *On the sum of Darboux functions*
Arsalan Chademan, *Sur la limit radiale des fonctions sous harmoniques*
O. P. Chandna, *Hamel's theorems in MHD*
Gerald Chapman, *Mixed extensions*
K. H. Chen, *Reduced symmetric hyperbolic system with discrete wave fronts*
Daniel E. Cohen, *Groups with free subgroups of finite index*
Harvey Cohn, *Ternary forms as invariants for Markoff forms and* $SL_2(Z)$ *bundles*
J. H. E. Cohn, *Length of period of continued fraction for the square root of d*
W. Comfort, *Uniform continuity in topological groups*
S. B. Conlon, *Finite linear p-groups of degree p*
W. Connett, *A better multiplier theorem for ultraspherical expansions*
J. L. B. Cooper, *Group properties of linear transformations*
M. A. Coppini, *Nouvelles considérations sur la méthode pour mesurer l'intensité de la redistribution du revenu opérée par la sécurité sociale*
Bruce D. Craven, *Lagrangean conditions for constrained optimization in Banach spaces*
R. M. Crownover, *Isometries that commute with shift operators*
Akos Csaszar, *Real compactifications, function classes, uniformities*
Thomas Cusick, *Gaps in the Markoff spectrum*
Taen-yu Dai, *On a class of dsc-pola*
Brian Davey, *Dualities for varieties of relative Stone algebras*
Anthony Davis, *Surface waves incident on a submerged horizontal cylinder*
Gary Davis, *Strict interpolation orders on ordered-permutation groups*
Donald Dawson, *Stochastic measure-valued processes*
M. De Guzman, *On the halo problem in differentiation theory*
Michael Deakin, *Curriculum priorities in developing countries*
S. Deb, *On group preference function*
L. Debnath, *Inertial oscillations and hydromagnetic multiple boundary layers in a rotating fluid*
Paul Dedecker, *Structure and invariants of multidimensional claculus of variations*
Ralph E. Demarr, *Thin elements in a dsc-pola*
M. Dempster, *A necessary and sufficient condition for optimal control*
John Dettman, *Saturation theorems connected with the abstract wave equation*
G. Di Antonio, *Solution of Poincaré's phase plane problem*
J. B. Diaz, *A generalization to higher dimensions of a theorem of Lucas concerning the zeroes of the derivative of a polynomial of one complex variable*
George Dinolt, *Truncations of principal geometries*
Jean-Pierre Dion, *Estimation of the variance of a branching process*
John D. Dixon, *Algorithmic problems in matrix groups*
D. Djokovic, *Matrices representable as* $A\bar{A}$
Claude Do, *Problèmes de valeurs propre d'inéquations variationelles: Application aux plaques minces*
Albrecht Dold, *The fixed point transfer of fiber preserving maps*
Y. Dold-Samplonius, *Two Arabic manuscripts on elementary geometry*
James Donaldson, *The Cauchy problem for systems of operator differential equations*
Joerg Donig, *The Cauchy problem for a parabolic equation with singular coefficients*
M. Doob, *Algebraic properties of magic graphs*
Alberto Dou, *On Korn's inequality*
Allan Douglas, *Computational aspects of diakoptics and the method of modified matrices*
Ed Dubinsky, *Applications of basis theory to expansions of analytic functions*

Patrick Eberlein, *The behaviour of geodesics on certain finitely connected surfaces*
W. F. Eberlein, *Primitive sequences*
Jurgen Eckhoff, *On combinatorial properties of convex sets*
R. B. Eggleton, *Cannibals: A game with captures*
Byron Ehle, *On certain order constrained Chebyshev rational approximations*
W. Eichhorn, *Axiomatic theory of the price index*
Calvin C. Elgot, *Matricial theories*
Erich Ellers, *The length problem*
George Elliott, *Inductive limits of sequences of matrix algebras*
John Emerson, *Simple points on affine algebraic varieties: An undergraduate course proposal*
M. Engber, *Fiber spaces over Teichmüller spaces*
Folke Eriksson, *The law of sines for tetrahedra and n-simplices*
John Ernest, *Decomposition and spectral multiplicity theory for bounded operators on a separable Hilbert space*
W. N. Everitt, *On an integral inequality associated with an ordinary differential equation*
J. O. C. Ezeilo, *Periodic solutions of a certain differential equation*
Edward Fadell, *Equivariant maps and Haar systems*
M. Farkas, *On isolated periodic solutions of differential systems*
Istvan Fary, *On the Mayer-Vietoris addition theorem*
A. Fasano, *Free boundary problems of general type for linear parabolic equations*
L. Ferguson, *Muntz-Szasz theorem with integral coefficients.* II
Herbert Fischer, *Tolerance arithmetic*
Bernard Fleishman, *Comparison and approximation of nonlinear boundary value problems*
Patrick J. Fleury, *Morita contexts and* 2×2 *matrix rings*
R. M. Fraker, *Optimization of bounds for boundary sets*
E. Frank, *Computation of certain special functions and applications*
Leonid Frank, *Spectral function of elliptic difference operator*
Jan F. Frankena, *Optimal control problems with delay*
Michael Fredman, *The complexity of some sorting problems*
J. O. Friel, *Adaptive prediction of stationary time series by modified conjugate direction methods*
Rudolf Fritsch, *Relative semisimplicial approximation*
Federico Gaeta, *A geometric description of period relations between holomorphic abelian differentials*
E. A. Gal'perin, *Equivalence classes of functionals and solvability of optimal control problems*
J. Galambos, *The distribution of values of arithmetical functions*
J. M. Gandhi, *On Fermat's last theorem*
Robert Gardner, *Vector relative invariants of* GL(n, F)
A. C. Garg, *Penny shaped crack at the interface of two dissimilar elastic layers*
Dean Gerber, *Jordan algebras and the explicit solution of quadratic differential equations*
L. Gerhards, *Projektive Darstellungen in der Galoistheorie und zugeordnete Einbettungsprobleme*
A. Ghaffari, *Motion of a satellite in an axially symmetric gravitational field*
M. A. Ghalib, *Fatigue fracture analysis*
John R. Giles, *The numerical range of unbounded operators*
J. Gillis, *Integrals of products of Laguerre polynomials*
Evarist Gine, *On quadratic variation of processes with Gaussian increments*
G. Glauberman, *A signalizer functor theorem*
L. S. Goddard, *Bases for certain prime polynomial rings*
D. Goelman, *The critical set of the reduced norm as an algebraic set*
Michael Goldberg, *Well-spaced bundles of lines*
A. W. Goodman, *A note on the zeros of Faber polynomials*
Gerald Goodman, *A new definition of Lebesgue measure*
S. H. Gould, *Mathematical analysis of genealogical spaces*
Douglass Grant, *B(𝒞) topological groups*

Curtis Greene, *Partitions associated with a partially ordered set*
T. N. E. Greville, *Smoothest moving-average interpolation formulae for equally spaced data*
Bronius Grigelionis, *The representation of square integrable martingales by stochastic integrals*
Alberto F. Grünbaum, *Identifying an input linear process from the output of a nonlinear device*
Branko Grunbaum, *How many hexagons are there?*
Claus Güenther, *Cubature formulas of polynomial precision with real nodes*
Robert Gulliver, *False branch points of minimal surfaces of higher topological type*
Hansraj Gupta, *Magic partitions*
Karl Gustafson, *Absorption for an arbitrary selfadjoint operator*
Bela Gyires, *On a limit distribution theorem of linear order statistics*
Alexander Haan, *The isomorphisms of the integral classical groups*
Peter Hagedorn, *A differential game with two pursuers and one evader*
William S. Hall, *Discontinuous periodic solutions of an autonomous nonlinear wave equation*
Noboru Hamada, *Construction of group divisible type PBIB designs*
V. L. Hansen, *Spaces of maps into a space of type* $(II, 1)$
Frank Harary, *On clique-extremal* (p, q)*-graphs*
H. Harborth, *On drawings of the complete graphs in the plane*
K. A. Hardie, *Approximating the EHQ sequence by the exact sequence of a quasifibration*
B. Hargrave, *Asymptotic representations for ellipsoidal wave functions of the third kind*
Hiroshi Haruki, *On the addition theorem* $f(x+y) = F(f(x), f(y))$
Egbert Harzheim, *On an elementary construction of the mapping degree*
Emami-Rad Hassan, *Semi-groupe distribution engendre par* $-A^{\alpha}$
Morisuke Hasumi, *Invariant subspaces on open Riemann surfaces*
W. Hausmann, *Recent developments on the existence of Haar subspaces*
Hiroshi Hayashi, *On the theory of hysteresis functions*
Stephen Hechler, *A new combinatorial structure and the Baire number of* $\beta N - N$
L. I. Hedberg, *Spectral synthesis in Sobolev spaces*
Erhard Heil, *Line families and evolutes*
Irvin Hentzel, *Generalization of right alternative rings*
John Herriot, *Procedures for interpolating natural quintic splines*
Gary Hewer, *Periodicity detectability and the matrix Riccati equation*
Neil Hindman, *Applications of the finite sum theorem*
F. Hirzebruch, *The Hilbert modular group and class number relations*
Xuan Hoang, *Le problème d'inverser des objets*
S. Hocquenghem, *Tabloides*
L. Hodgkin, *The K-theory of many well-known spaces*
William Hoffman, *The role of Lie transformation groups in the life sciences*
P. C. Hogarth, *Decomposition of graphs into stars*
H. Hogbe-Nlend, *Sur les rapports entre espaces de Banach et espaces nucleaires*
Charles Holmes, *Two groups of order* 189 *with identical subgroup structures*
A. Howe, *Perturbed linear Hamiltonian systems with periodic coefficients*
Po-fang Hsieh, *On nonlinear differential equations with an irregular type singularity*
D. Huet, *Singular perturbations*
Kenneth Hughes, *Descent of homotopy structures*
Paul Humke, *Oppositely ambiguous points of sets*
J. H. V. Hunt, *The Phragmen-Brouwer theorem for separated sets*
L. R. Hunt, *Extension of C-R functions*
James Hurley, *Normal closures in elementary subgroups of Chevalley groups over rings*
A. Iarrobino, *Vector spaces of forms in* $k(x, y)$*, and Wronskians*
C. Infantozzi, *Sur la* (S)*-derivation (derivation de Stieltjes): Une généralisation de la formule de Taylor*
S. Ishihara, *On harmonic and relatively affine mappings*
Kiro Isobe, *On representation of a certain partially ordered linear space*

V. F. Ivanoff, *On the parabola and pentagon inscribed into it*
S. O. Iyahen, *On the closed graph theorem*
Thomas Jech, *Construction of trees and applications*
W. E. Jenner, *Some remarks on the construction of simple nonassociative algebras*
D. L. Johnson, *The Fibonacci groups*
Roy Johnson, *The product of a nonregular Borel measure and a regular Borel measure*
John Jones, *On the zeros of functions with values in Hilbert space*
Gerard Joseph, *Numerical range of operators on locally convex spaces*
E. Jucovic, *On face and vertex vectors of cell decompositions of orientable 2-manifolds*
M. A. Kaashoek, *On the reduced algebraic multiplicity of a meromorphic operator function at a certain point*
John Kalman, *Axiomatizations of logics with values in groups*
P. N. Kaloni, *A simplified continuum theory of structure fluids*
H. Kamowitz, *The spectra of composition operators on* H^p
P. Kannappan, *On a generalization of some measures in information theory*
Hans G. Kaper, *Continuity and differentiability properties of the solution of the integral transport equation*
D. A. Kappos, *Probabilistic characterization of the observables in the quantum mechanics*
Takahiro Kawai, *Finite dimensionality of cohomology groups*
A. Kearsley, *Applications of isotropic Cartesian tensors to field theories of deformable media*
David Kelly, *Planar partially ordered sets*
Frank Keogh, *The Schwarzian derivative and estimates of functions analytic on the unit disc*
William Kerby, *Sharply 3-transitive groups of characteristic congruent to* 1 (mod 3)
G. Khosroshahi, *Growth properties of solutions of Schrodinger type equations*
Hong Kim, *Hyponormal operators with infinite essential spectrum*
Ronald Kirk, *The Haar integral via nonstandard analysis*
Abel Klein, *Singular Hamiltonians in quantum field theory via the Feynman-Kac-Nelson formula*
Larisse Klein, *A perturbation-Galerkin method for transverse vibrations*
R. J. Knill, *On the Hopf flow in projective* $2n-1$ *space*
R. A. Knoebel, *Regular expressions for universal algebras*
John Knopfmacher, *Maximum orders of magnitude of arithmetical functions*
Helmut Koch, *On some questions in the Galois theory of algebraic number fields*
M. Komhoff, *Some problems on stack polytopes*
J. Komlos, *An approximation of partial sums of independent RV's, and the sample DF*
B. Korenblum, *Closed ideals in certain Banach algebras of analytic functions*
Helen Kosachevskaya, *The approximate solution of Oldroyd equations of nonstationary viscoplastic boundary layer under the wall slip condition*
U. Koschorke, *Frame fields and other bundle monomorphisms*
Arthur Krener, *High order maximal principle for singular extremals*
V. Krishnamurthy, *On a problem of tree counting related to finite maximal connected* T_0 *topologies*
Frantisek Krnan, *Preliminary report on an algorithm for formation of a sequence of twin primes*
Eric Kronstadt, *Interpolating sequences for analytic functions satisfying a Lipschitz condition*
Djuro Kurepa, *On a new function*
Thomas Kurtz, *Semigroups and singular perturbation of ordinary differential equations*
W. Kuyk, *Modular function theory of the groups* $G(\sqrt{Q})$, Q *a prime*
R. C. Lacher, *Embeddings with mapping cylinder neighborhoods*
G. S. Ladde, *Variational comparison theorem and perturbations of nonlinear systems*
Hang Chin Lai, *Retrictions or liftings of Fourier transforms on AP*
H. F. Lai, *Pseudo conformal geometry in codimension* 2
Ilpo Laine, *Results on harmonic B1-mappings*
V. Lakshmikantham, *Stability theory through abstract cones*
B. S. Lalli, *Asymptotic equivalence and functional differential systems*
C. Lamoureux, *Qualitative properties of codimension one foliations*

D. Laugwitz, *A characterization of normed Minkowski spaces*

Richard B. Lavine, *Time estimates in scattering theory*

Terry Lawson, *Low dimensional failures of the S-cobordism conjecture*

David Lay, *Existence of holomorphic projection functions and generalized inverses of operator functions*

W. G. Leavitt, *A note on upper radicals*

S. J. Lee, *Generalized resolvents for definitely symmetric systems*

C. R. Leedham-Green, *On p-groups of maximal class*

S. Leela, *A new technique in perturbation theory*

Eugene Lehman, *Méthode d'approximation de la somme d'une série pratique pour étudients du deuxième semestre*

R. Sherman Lehman, *Evaluation of all polynomials*

D. H. Lehmer, *Characters of linear permutations*

Roy Leipnik, *Quadratic differential equations*

F. Wm. Lemire, *A new family of irreducible representations of* A_n

J. Lepowsky, *Conical vectors in induced modules*

H. Levinson, *An upper bound for the surface genus of a knot*

Chong-Keang Lim, *A result on free pro-nilpotent group*

T. Y. Lin, *Adams type spectral sequence for stable homotopy and its applications*

W. Lindenberg, *Bizyklische P-gruppen, ihre Klassifizierung und ihre Automorphismen*

William Ling, *On varisolvency and generalized rational approximation*

M. Lipschutz-Yevick, *Elements of a language of patterns based on holographic logic*

I. S. Luthar, *Geometry of numbers in local fields of characteristic p not equal to* 0

S. M. Makky, *Optimum design of domes*

Gary Makowski, *A convergence rate for an estimator of a doubly nondecreasing repression surface*

A. R. Manwell, *Weak shock waves in plane transonic flow*

K. Marathe, *Geometric theory of gravitation*

Sibe Mardesic, *The Whitehead theorem in shape theory*

Charles Marshall, *Differentiable spaces and vector pseudobundles*

O. Martio, *Period quasiregular mappings into n-space*

P. Masani, *Ergodic theorem for locally integrable semigroups of continuous linear operators on a Banach space*

John Masley, *Solution of class number one problem for cyclotomic fields*

Klaus Matthes, *Limit theorems for branching schemes, homogeneous in space, with a finite number of types*

Meinard Mayer, *Connections in principal fibrations and gauge theories in physics*

James R. McConnell, *Multiplication of Schur functions*

Sean McElwain, *A relationship between the eigenvalues of two relaxation matrices*

J. Meldrum, *Group d. g. near-rings*

Abolghassen Miamme, *Positive operators on Banach space and the Fejer-Riesz theorem*

E. Michael, *A selection theorem with countable domain*

Philippe Michel, *Problème des inégalites. Application aux problèmes d'optimisation*

Anna Maria Micheletti, *On periodic solutions of nonlinear variational ordinary differential equations*

M. Mikolas, *On the recent progress of the theory of fractional differential and integral operators*

B. Arthur Miller, *Varieties of radical rings*

Edward Miller, *The homotopy Euler characteristic of manifolds with vector fields*

Sanford S. Miller, *Generating functions for some classes of univarient functions*

Kenneth Millett, *Fibered geometric topology: Concordances*

R. Millman, *Analytic vector harmonic expansions on* SU(2) and S^2

Pier Milman, *Total index theory of weakly continuous mappings*

Paul Milnes, *Extension of continuous functions on topological semigroups*

Edward Miranda, *An analysis of the impact of computers in the calculus sequence*

R. S. Mishra, *Certain specified structures in a differentiable manifold*

Shaw Mong, *The signature of homogeneous spaces*

J. J. Moreau, *Systèms elastoplastiques*

Christopher Morgan, *On embedding metric spaces in Euclidean space*

G. R. Morris, *A fixed-point theorem for multifunctions of finite type*

John A. Morrison, *Charge singularity at the tip of a slender cone of general cross section*

Louise Moser, *On the impossibility of obtaining the product* $S^2 \times S^1$ *by elementary surgery along a knot*

Jack Mosevich, *A numerical method for appoximating solutions to functional equations arising in a model of an epidemic*

Some Mukherjee, *Some theorems on the convergence and Cesàro summability of multiple orthogonal series*

Helmut Müller, *Asymptotic distribution of Beurling's generalized integers*

C. J. Mulvey, *Compact representations of rings*

Hyo Myung, *Prime and primary ideal theories in a nonassociative ring*

Kazumi Nakano, *Connector theory*

Mysore Narasimahn, *Thermomicropolar theory of heat-conducting nematic liquid crystals*

C. Nash-Williams, *Marriage in denumerable societies*

Maher Nassif, *Gontcharoff polynomials generating functions of two complex variables*

Melvyn Nathanson, *Minimal bases and maximal nonbases in additive number theory*

M. K. Nestell, *Branching of solutions of Chandrasekhar's H-equation*

John Neuberger, *Itera for partial differential equations*

Edward Newberger, *The Cauchy problem with data in non-quasi-analytic classes*

M. F. Newman, *Groups of exponent four*

Jean-Louis Nicolas, *Nombres hautement composés et optimisation en nombres entiers*

José I. Nieto, *Gâteaux differentials in Banach algebras*

F. Niiro, *On a set including the spectra of positive operators*

Albert Nijenhuis, *Smith canonical forms of matrices*

John Niman, *The teaching of graph theory*

N. Ninomiya, *On the potential taken with respect to complex valued and symmetric kernels*

Karl Norton, *On the number of restricted prime factors of an integer*

Chike Obi, *Periodic oscillations of a certain differential equation*

Shin-Ichi Ohwaki, *General boundary problems for linear differential equations*

J. Okee, *A lattice-theoretic study of the algebra of species*

T. Okubo, *Geometry of tangent bundles viewed by submersion*

P. Olin, *Elementary equivalence of free products*

Andrew Olson, *On symbolic-numeric series methods for nonlinear operator equations*

Tamio Ono, *In a finite* AW^* *algebra* $x^*x \leqq xx^*$ *implies* $x^*x = xx^*$

Gerhard Opfer, *Numerical methods for calculating conformal maps with Tchebycheff approximations*

P. Orlik, *Resolution of singularities of projective surfaces with* C^* *action*

C. P. Ormell, *The consciousness gap in mathematical education*

Charles Osgood, *Finding good rational approximations to power series*

Tominosuke Otsuki, *Geodesics of* O_n^2 *and an analysis on a related Riemann surface*

Gunther Pabst, *Quantitative axioms for affine planes*

R. Padmanabhan, *On lattices with unique complements*

Carlo Pagani, *On an ultraparabolic equation*

Laszlo Pal, *A new modification of Hermite-Fejér-interpolation*

J. N. Pandey, *Hilbert transforms of generalized functions*

M. Papadopoulos, *A new result in orthogonal function theory relating to functions which satisfy linear linkage conditions at the ends of segments of a broken line*

Eli Passow, *Copositive polynomial approximation*

Francis Pastijn, *A generalization of Green's equivalence relations for half groupoids*

J. Patera, *Maximal solvable subgroups of $SU(p, q)$ and $O(p, q)$ groups*
Daniel Pedoe, *Ausz disem wirdt vil dings gemacht*
Donald Pelles, *Iteration of endomorphisms of $P^n(C)$*
J. P. Penot, *Sous-differentiels de fonctions non-convexes*
Jan Persson, *Nonuniqueness of Tacklind type in the characteristic Cauchy problem*
Gordon Petersen, *Metric density and Lusin's theorem*
N. C. Petridis, *Mappings of bounded dilatation*
John Phythian, *Open University mathematics*
S. Piccard, *Les groupes des transformations periodiques de l'ensemble des entiers rationnels*
Livio Piccinini, *Stability results for systems of ordinary differential equations with discontinuous coefficients*
A. Pietsch, *Minimal and maximal operator ideals*
D. Pigozzi, *Universal varieties of algebras*
W. Plesken, *The Bravais group and the normalizer of a reducible finite subgroup of* $GL(n,Z)$
B. S. Popov, *Generating functions for powers of certain sequences of numbers*
Giovanni Porru, *Pseudo conformal mappings in space*
Henry Potoczny, *Closure preserving families and metacompactness*
A. Potter, *A degree for a class of noncompact operators*
N. Prakash, *Differential geometry of framed manifolds via Hodge theory*
David Price, *M-triples and symplectic action*
R. S. Printis, *On controlling shocks in solutions of nonlinear partial differential equations*
Pratap Puri, *Plane thermo elastic waves in a rotating medium*
T. K. Puttaswamy, *Stokes phenomenon for a certain third-order differential equation*
G. F. Raggett, *On the solution of hyperbolic partial differential equations using cubic splines*
M. Rajagopalan, *On rigidity and groups of homeomorphisms.* II
Arlan Ramsay, *The Mackey machine without regular imbedding*
Kishan Rao, *Thermal stresses in a solid elastic cone*
R. K. Rao, *Asymptotic theory for a class of perturbed functional equations*
M. Razzaghi, *On the matrix Riccati equation in optimal control*
M. Reichert, *Losungsmengen abstrakter Volterra-Operatoren*
Ivan L. Reilly, *On the quasimetrization problem*
P. Rejto, *On a theorem of Titchmarsh-Neumark-Walter concerning absolutely continuous operators*
John A. Rickard, *A mathematical model of the basilar membrane in mammals*
Giovanni Rizza, *Monogenic functions on algebras and conformal mappings*
Raphael Robinson, *The structure of certain triple systems*
D. Roeder, *Necessary conditions for p-exceptionality of integral quadratic forms*
A. Ronveaux, *États Lies à energie nulle en théorie de Schroedinger*
Ivo G. Rosenberg, *A Galois correspondence between universal algebras and relations on the carrier*
Millu Rosenblatt-Roth, *Compound communication channels*
G. Roskes, *Some unstable multimode wavetrains*
Peter Rowlinson, *Involutions fixing at most four points*
Jose Rubio, *Nets of subgroups in locally compact groups*
W. Ruckle, *Absolutely divergent series and isomorphism of subspaces*
John Rutter, *A homotopy equivalence extending the Bott-Samelson theorem and giving a dual to Sugawara's fiber sequence*
G. S. Sacerdote, *Infinite coforcing in model theory*
G. Saitoti, *K-theory* mod 2 *of second iterated loop space*
Victor Saks, *Countably compact and pseudocompact powers*
Pablo Salzberg, *Liapunov stability in pre-ordered topological spaces*
N. G. F. Sancho, *A two stage algorithm for the solution of optimization problems in dynamic programming*
A. D. Sands, *Radical properties in associative rings*
Bill Sands, *Finite weakly embeddable lattices which are embeddable*

I. Sawashima, *Reduction of a sub-Markov operator to its irreducible components*
H. Sayeki, *Scale of functions and the continuum*
Joel Schiff, *Harmonic null sets and the Painlevé theorem*
Asmus Schmidt, *Diophantine approximation of complex numbers*
Rolf Schneider, *Transformations of convex bodies*
Irwin Schochetman, *The dual topology of certain group extensions*
Barbara Schumacher, *Audiovisuelle Medien im Mathematikunterricht*
Stefan Schwarz, *Ideal structure of C-semigroups*
D. Schweigert, *On a concept of completeness for finite lattices*
Ridgway Scott, *Optimal Galerkin approximation of singular functions*
Jack Segal, *Locally well-behaved paracompacta in shape theory*
P. Seibert, *A unified theory of Liapunov stability*
B. R. Seth, *Transitions: Basic concepts*
S. M. Shah, *Relations between the radius of convergence of a power series and the radii of univalence of its derivatives*
Gerald Shanholt, *Some stability results for neutral functional differential equations*
A. Sharma, *Cardinal lacunary interpolation by G-splines*
R. N. Shekoury, *On quasi-trivializable spaces*
Max Shiffman, *Measurable and nonmeasurable sets*
J. Shiue, *Uniform distribution in the ring of Gaussian integers*
Richard A. Shore, *Some counterexamples in α-recursion theory*
H. Shuard, *Mathematical work of colleges of education in England*
Frederic Shultz, *Noncommutative spectral theory on convex sets*
J. A. Siddiqi, *Approximation by polynomials on arcs*
Udo Simon, *Some differential operators of second order on homogeneous manifolds*
S. Simons, *Convergence theorems and minimax theorems*
M. K. Singal, *On the closure preserving sum theorem*
Donald Singley, *The rank of the difference tensor on parallel hypersurfaces*
P. Sjolin, *Fourier multipliers and estimates of the Fourier transform of measures carried by smooth curves in* R^2
Helge Skovgaard, *A statistical investigation of the scores and background of* 7000 *engineering students in* 1957–1970
A. D. Sloan, *Analytic domination by fractional powers with linear estimates*
Edwin Smith, *The role of the mathematician in the preparation of elementary school teachers*
Harry Smith, *Generalization* (I) *of alternative algebras*
J. C. Smith, *Theta-type characterizations for generalized paracompact spaces*
N. Smythe, *Grusko's theorem via groupoids*
J. Sniatycki, *Bohr-Sommerfeld conditions in geometric quantization*
Donald R. Snow, *Functional equations and sums of powers of arithmetic sequences*
R. Soares, *Récurrences en polynômes orthogonaux*
D. Socolescu, *Existence of a viscous fluid flow over a bottom with declivity*
Edward Soetens, *On nonmetric geodesics*
M. Sonis, *On ideals in Banach algebras generated by linear functionals*
G. C. Sornberger, *Random response queue*
Andrade Soto, *The representations of* GSP(4, k) *and* SP(4, k) *for a finite field* k
V. M. Soundalgekar, *Free convection effects on MHD oscillatory flow past vertical porous plate*
G. Spinelli, *Possible nonlinear Lagrangian tensor field theories of gravitation*
F. W. Stallmann, *Interpolating splines in higher dimensions*
Richard Stanley, *Shelling, homological dimension, and the upper bound conjecture*
M. Starbird, *Normality in products and closed maps*
Sherman Stein, *Algebraic packing and tiling*
Ralph Steinlage, *Equicontinuity of transformation groups*
Jacques Stern, *Caractérisation par des formules d'espaces de Banach et d'operateurs*

Ian Stewart, *Infinite-dimensional soluble Lie algebras*

A. Stieglitz, *Invariant cohomology of G-spaces*

Arthur Stone, *Tower adjoint functors and elementary properties*

William Stone, *On a generalization of the distribution of the significant digits under computation*

M. J. Strauss, *Uniqueness of analytic continuation*

S. Strelitz, *The asymptotic behaviour of solutions of differential equation on a ray*

Li Pi Su, *Round subsets of Wallman-type compactifications*

W. H. Summers, *The weighted approximation theorem*

Dan Sunday, *Pulse to wave transformations in neuron populations*

J. G. Sunday, *Hypergraphs and abstract polygons*

C. Sutti, *On a new unconstrained minimization algorithm without derivatives*

Tatsuo Suwa, *Compact quotient spaces of C^3 by affine transformation groups*

Komei Suzuki, *On the notion of probability*

Halina Swiatak, *Regularity problems for nonlinear functional equations*

J. Szabados, *Saturation problems of discrete linear operators*

Nanfred Egon Szabo, *Classically, simple rationality = rational simplicity, categorically*

Earl J. Taft, *The antipode of a finite-dimensional Hopf algebra*

Masayuki Takahashi, *A generalization of Kuhn's theorem for an infinite game*

Shuichi Takahashi, *On Boolean-valued automata*

Franklin Tall, *Between Martin's axiom and Souslin's hypothesis*

Takayuki Tamura, *On Z-$\mathcal{N}$ semigroups*

T. T. Tanimoto, *Determination of distributions*

O. Taussky-Todd, *Additive commutators between 2×2 integral representations of quadratic orders*

D. E. Taylor, *On the normal structure of the unit group of a group ring*

Peter Taylor, *Embedding a compact convex set into a LCTVS*

Armin Thedy, *Structure of simple right alternative rings*

H. Thiele, *Note on the connection between search theory and the theory of experiments on finite automata*

Charles B. Thomas, *Contact structures on 1-connected 5-manifolds*

P. Tondeur, *Nontriviality of secondary characteristic classes for foliated bundles*

L. V. Toralballa, *A geometric solution of the problem of Plateau in 3-space*

J. Totten, *Restricted linear spaces*

Neil Trudinger, *A family of maximum principles*

Aubrey Truman, *Spectrality, cluster decomposition and small distance properties in Wightman field theory*

G. Tsagas, *On the manifolds with strictly negative sectional curvature*

P. M. Tuchinsky, *Linear asymptotic stress equations for thin cylindrical shells*

Edward Tully, *Inverses relative to inner identity elements in a semigroup*

Atwell Turquette, *Minimal axioms for Pierce's triadic logic*

Joseph L. Ullman, *Extremal polynomials associated with a compact set in the plane*

G. Ungar, *ANR's in the mapping category*

R. G. Van Meter, *Some special equations over a finite field*

F. M. J. Van Oystaeyen, *Noncommutative localization and sheaves*

W. Vasilaky, *A probabilistic approach to the approximation of solutions to the parabolic differential equation*

Jerry Vaughn, *Nonnormal products of ω_{M_μ} metrizable spaces*

Theresa Vaughn, *Factorization of $Q(h(T)(x))$ over a finite field, where $Q(x)$ is irreducible and $h(T)(x)$ is linear.* II

R. Venkataraman, *A characterization of topological groups that satisfy Pontryagin duality*

B. G. Verma, *Determination of flow field behind magnetogasdynamic blast wave under the influence of viscosity*

Jaime Vinuesa, *Orthogonal polynomials on harmonic algebraic curves*

H. Von Lienen, *On a certain Pell's equation*

M. Von Renteln, *Finitely generated ideals in subalgebras of the disc algebra*
Eldon J. Vought, *On decompositions of hereditarily unicoherent continua*
William Wade, *Sets of uniqueness and α-capacity on the group 2^{ω}*
Stanley Wagon, *Saturated ideals in Boolean extensions of L^{U}*
Martin Ward, *Quasivarieties of partial algebras*
Richard Warren, *Neighborhoods, bases and continuity in fuzzy topological spaces*
Simon Wasserman, *Commutation in tensor products of C*-tensor products*
H. Wefelscheid, *Sharply 2-transitive groups and finiteness conditions*
Gerd Wegner, *D-collapsing and families of convex sets*
Wolfgang Weil, *The characterization of balls by brightness and width*
Hanns J. Weinert, *Rings with noncommutative addition*
Joseph Weinstein, *Reconstructing colored graphs: Some reductions*
W. Wesselius, *Lipschitz semialgebras*
Alvin White, *Beyond behavioral objectives*
James Whittaker, *Covering the exterior of the horned sphere*
H. H. Wicke, *On topological structures defined by primitive sequences*
Fletcher D. Wicker, *Bloch and normal mappings on Hermitian manifolds*
J. B. Wilker, *Equidistant sets and their connectivity properties*
Scott Williams, *Paracompact products*
Patricia Willis, *Computer symbolic integration of nonlinear differential equations: Preliminary report*
Paul Willis, *Distortion mappings.* II: *Preliminary report*
Jorg Wills, *Convex bodies and lattice points*
Robert Lee Wilson, *A structural characterization of the simple Lie algebras of Cartan type*
Wolfgang Winkler, *Sequential statistical problems for processes with independent increments*
M. B. Wischnewsky, *Sheaves in topological categories*
Claudio Withalm, *Hyperpseudo-holomorphic Funktionen*
Joseph Wolf, *Geometric realizations of unitary representations*
John Wolfe, *Expectations of the anticommutation relations*
Karl Wolff, *Fast-Blockpläne*
Chi Wong, *Close-to-normal structure*
Yung-Chow Wong, *M-Tensors and M-connections on tangent bundle TM*
A. Wood, *Eigenfunction expansions by δ-function and contour integral*
D. H. Wortman, *Some results concerning the maximal ideal space of H^{∞}*
Aubrey Wulfsohn, *Tensor products of Jordan algebras*
Chung-Chun Yang, *On entire functions taking the same values at the same points*
A. Yaqub, *Finiteness of semisimple rings*
D. Younger, *A minimax theorem for directed graphs*
Don Zagier, *A Kronecker limit formula for real quadratic fields*
A. J. Zajta, *Semi-solutions of a Diophantine equation*
W. R. Zame, *Ideal structure of rings of holomorphic germs*
Phillip Zenor, *A metrization theorem for normal Moore spaces*
F. Zirilli, *Ginsburg-Landau equations and bifurcation theory*

Index

Roman numbers refer to pages in Volume 1.
Italic numbers refer to pages in Volume 2.