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 $n$th 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) perturbations 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 $m_i r_j^{1/2} A_{ij} = m_j r_i^{1/2} A_{ji}$, where $m_i$ = mass of $i$th 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

\[
AX = \lambda BX \quad \text{with related differential equations} \quad B\bar{Y} = AY,
\]

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

¹The 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 \( X_{rs} \) is a system of coefficients such that \( X_{rr} = 1 \) and \( X_{sr} = -X_{rs} \), then the system of coefficients \( a_{rs} \) defined by

\[
(2) \quad a_{rs} = 2D_{rs}/D - \delta_{rs}, \quad D = |\lambda_{rs}|, \quad D_{rs} = \partial D/\partial \lambda_{rs},
\]

has the Lagrangian orthogonality properties. Thus the coefficients \( a_{rs} \) of the orthogonal substitution are expressible as rational functions of \( n(n-1)/2 \) parameters—the \( \lambda_{rs} \).

Expressed in modern symbolism, Cayley’s solution can be written as

\[
(3) \quad U = 2(I + S)^{-1} - I = (I + S)^{-1}(I - S), \quad U := (a_{rs}), \quad I + S = (\lambda_{rs}).
\]

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:

\[
(X, Y, Z, \ldots) = \begin{bmatrix}
    \begin{vmatrix}
        a, h, g, \ldots \\
        b, f, \ldots \\
        c, \ldots \\
    \end{vmatrix}
    & a, h - \nu, g + \mu, \ldots \\
    h + \nu, b, f - \lambda, \ldots \\
    g - \mu, f + \lambda, c, \ldots \\
\end{bmatrix}
\]

(4)

In modern notation

\[
X = A^{-1}(A - S)(A + S)^{-1}Ax.
\]

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 characteristic 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 $\leq n−1$. 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 $\lambda$ 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

7. A. L. Cauchy, Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opérées entre les variables qu’elles renferment, J. École Polytech. 10 (1815) = Oeuvres (2) 1, 91–169.