Weierstrass and the Theory of Matrices

THOMAS HAWKINS

Communicated by M. KLINE

Table of Contents

1. Introduction 119
2. Multiple Roots and Generic Reasoning 121
3. Weierstrass' Paper of 1858 124
4. The Theory of Elementary Divisors 132
5. Related Developments in England and France 139
6. Frobenius 147
7. Weierstrass and the Rise of the Theory of Matrices 155
8. Bibliography 159

1. Introduction

The following essay is the third in a series devoted to the history of the theory of matrices.[1] In [1975a] I related CAUCHY's important memoir on the characteristic roots of a real quadratic form to the mechanical investigations of LAGRANGE and LAPLACE. The origins and development of the symbolical algebra of matrices are traced in [1977]. Now I present the keystone of the history: the theory of canonical matrix forms and, in particular, the work of KARL WEIERSTRASS. The emphasis upon WEIERSTRASS reflects my conclusion that through his theory of elementary divisors [1868] and its influence upon other mathematicians, he, more than anyone, was responsible for the emergence of the theory of matrices as a coherent, substantial branch of twentieth-century mathematics.

In view of the importance of WEIERSTRASS' memoir on elementary divisors, special attention has been given to the background against which it was created and to the mathematicians who influenced him (Sections 2–4). In [1868] and a preliminary memoir [1858] WEIERSTRASS demonstrated more than theorems. He also demonstrated the possibility and desirability of a more rigorous approach to algebraic analysis that did not rest content with the prevailing tendency to reason vaguely in terms of the "general" case. WEIERSTRASS' memoir thus served as a paradigm for further research by his colleagues and students such as KRONECKER and FROBENIUS. It was by virtue of the influence of his work on others that WEIERSTRASS played such a key role in the evolution of the theory of matrices. Section 6 is entirely devoted to the fundamental contributions of FROBENIUS while in Section 7 I establish the dominant role played by the Berlin school (WEIERSTRASS, KRONECKER, FROBENIUS) in the development, application, syste-

[1] I gratefully acknowledge the support of the U.S. National Science Foundation.

matization and popularization of the theory. Finally, in order to clarify WEIERSTRASS's role, it is also necessary to consider the related work done in England by SYLVESTER, CAYLEY and SMITH and in France by JORDAN (Section 5).

To appreciate fully the material discussed in Sections 2 through 4 it is necessary to be mindful of the essentials of WEIERSTRASS' theory of elementary divisors as formulated by him. The main ideas and results are as follows. Consider a family or system of bilinear forms $s\Phi-\Psi$, where s is a complex parameter and $\Phi=\sum_{i,j=1}^{n} A_{ij}x_iy_j$, $\Psi=\sum_{i,j=1}^{n} B_{ij}x_iy_j$. Let $D(s)$, or $D_0(s)$, denote the determinant $|sA_{ij}-B_{ij}|$ $=|s\Phi-\Psi|$. The greatest common divisor of all the first minors of $D_0(s)$, *i.e.* the $(n-1)\times(n-1)$ minors, considered as polynomials in s, is denoted by $D_1(s)$. In similar fashion $D_2(s)$ is defined as the greatest common divisor of all the second, or $(n-2)\times(n-2)$, minors of $D_0(s)$, $D_3(s)$ as the greatest common divisors of the third minors, and so on. It can then be shown that $D_i(s)$ divides $D_{i-1}(s)$ and that if $E_i(s)$ denotes the polynomial $D_{i-1}(s)/D_i(s)$, then $E_i(s)$ divides $E_{i-1}(s)$. It follows from the definition of the $E_i(s)$ and the divisibility properties of the $D_i(s)$ that $D(s)=|s\Phi-\Psi|$ differs from the product of the $E_i(s)$ by a constant. Thus of $s_1, s_2, \ldots, s_k$ are the distinct roots of $D(s)$,

$$E_i(s)=\gamma_i\prod_{j=1}^{k}(s-s_j)^{m_{ij}}$$

where γ_i is a constant and the m_{ij} are positive integers or zero. Each factor $e_{ij}=(s-s_j)^{m_{ij}}$ with $m_{ij}>0$ is called an *elementary divisor* of $D(s)$ or of the family $s\Phi-\Psi$. (Following FROBENIUS, many authors refer to the $E_i(s)$ as the elementary divisors while the $e_{ij}(s)$ are called the simple elementary divisors.)

Two families of bilinear forms, $s\Phi-\Psi$ and $s\Phi'-\Psi'$, are *equivalent* if one can be transformed into the other by (possibly different) nonsingular linear transformations of the x and y variables. The main theorem of WEIERSTRASS' theory concerns families $s\Phi-\Psi$ which are non-singular in the sense that $|s\Phi-\Psi|\not\equiv 0$. It asserts that two such families are equivalent if and only if they possess the same elementary divisors. To prove the theorem, WEIERSTRASS introduced the equivalent of the canonical form discovered independently by CAMILLE JORDAN.

The theory of elementary divisors does not explicitly involve matrices, and the question therefore arises as to the role of matrix algebra in the history of what is now referred to as the theory of matrices. I have discussed the matter in [1977]. Here I simply wish to point out that CAYLEY's oft-cited memoir containing the elements of matrix algebra [1858] was unknown outside of England until the 1880's. By that time the content of the theory of matrices had been developed extensively on the Continent within the framework of quadratic and bilinear forms and linear transformations. Furthermore, in [1878] FROBENIUS introduced the equivalent of the symbolical algebra of matrices without knowledge of CAYLEY's memoir [1858], although he was familiar with some earlier suggestive observations by CAYLEY [1855a, 1855b] that eventually led to his memoir [1858]. FROBENIUS' treatment of matrix algebra was vastly superior to CAYLEY's because he was able to apply the results of his teachers, WEIERSTRASS and KRONECKER, on bilinear forms. Matrix algebra acquired substance and became of general interest to mathematicians only after it was fused with the theory of canonical matrix

forms, and FROBENIUS' efforts in this direction, in [1878] and subsequent publications, were especially important. (See Sections 6 and 7 and my [1977].)

Although the mathematicians whose work is presented in Sections 2–5 made no use of matrix symbolism (unless otherwise noted), for the sake of brevity I use some modern matrix notation. No confusion or misunderstanding should arise as long as it is kept in mind that any equations I express with matrices should be interpreted as a system of equations. In Sections 6 and 7 the notation is that of FROBENIUS with the exception that I have omitted the boldface type he favored.

2. Multiple Roots and Generic Reasoning

The geometers of the 18th century had inherited two magnificent achievements from the previous century. One was an entirely new way of doing mathematics, the method of symbolical analysis, which had been introduced by VIÈTE and DESCARTES and further developed by NEWTON and LEIBNIZ. The second great achievement was NEWTON's *Philosophiae naturalis principia mathematica* (1687). Like most great works the *Principia* left many problems to clarify and resolve. Furthermore, although NEWTON had earlier been an enthusiastic exponent of DESCARTES' method of symbolical analysis, by the time he composed the *Principia* he preferred the less controversial geometrical mode of presentation by means of synthetic infinitesimal geometry.

It was the Continental mathematicians of the 18th century, men such as D'ALEMBERT, EULER, LAGRANGE and LAPLACE, who perfected the analytical methods of their predecessors and demonstrated their power by systematically and successfully applying them to the problems in terrestrial and celestial mechanics suggested by NEWTON's *Principia*. Two works which summed up and symbolized these accomplishments were LAGRANGE'S *Méchanique analitique* [1788] and LAPLACE's *Traité de méchanique céleste* [1799]. These works represented unequivocal triumphs for the method of analysis. LAGRANGE and LAPLACE were well aware of this fact which seems to have induced them to place unlimited confidence in the method of analysis.

Thus in the preface to *Méchanique analitique* [1788] LAGRANGE boasted that his book contained neither geometrical figures nor geometrical constructions, that the methods expounded by him involved only "algebraic" operations subjected to a systematic and uniform progression" so that mechanics becomes a new branch of analysis. And LAPLACE, in the course of contrasting the geometrical methods of NEWTON with his own, wrote that although algebraic analysis necessitates consideration of "abstract combinations" far removed from the particulars of the problem at hand, "by abandoning oneself to the operations of Analysis ... one is led, by the generality of this method and by the inestimable advantage of transforming the reasoning into mechanical procedures, to results often inaccessible to synthesis. Such is the fruitfulness of Analysis that it suffices to translate particular truths into this language in order to see emerge from their very expression a multitude of new and unexpected truths." [1835, 465].

Being understandably preoccupied with the power of the method of symbolical analysis, LAGRANGE and LAPLACE tended to overlook certain concomitant difficulties. As the above quotations indicate, the great virtue of analysis was its

generality and the resulting uniform, even mindlessly mechanical, nature of its procedures. The generality of the method of analysis had been viewed as its great virtue ever since its inception. Thus VIÈTE stressed that the new method of analysis "does not employ its logic on numbers—which was the tediousness of the ancient analysts—but uses its logic through a logistic which in a new way has to do with species. This logistic is much more successful and powerful than the numerical one for comparing magnitudes with one another in equations.... " [1591, 321-322]. In the new analysis one reckons with species by means of "signs" (symbols) which do not represent specific magnitudes, but an entire species of magnitudes. Analysis became a method for reasoning with, manipulating, expressions involving symbols with "general" values and a tendency developed to think almost exclusively in terms of the "general" case with little, if any, attention given to potential difficulties or inaccuracies that might be caused by assigning certain specific values to the symbols. Such reasoning with "general" expressions I shall refer to for the sake of brevity as *generic reasoning*.[2]

LAGRANGE's treatment of the principal axis theorem of mechanics [1788, Pt. II, Sect. VI] affords a simple example of generic reasoning. In the course of it algebraic expressions arise which for certain values of the symbols involved possess zeros in denominators. Since these expressions are however meaningful on the generic level the proof proceeds. In this fairly simple situation the singular cases could have been treated separately and individually and the "general" conclusions maintained without exception. But in more complex problems it is not always clear what bearing such exceptions to the generic reasoning might have on the conclusions drawn. Moreover, little attention was accorded to such matters.

During the early stages of the 19th century some mathematicians, concerned with raising the standards of rigor in analysis, rejected the legitimacy of generic reasoning. Thus CAUCHY insisted in the preface to his *Cours d'Analyse* [1821] that mathematical reasoning must not be based on arguments "drawn from the generality of algebra", arguments which "tend to attribute an indefinite scope to the algebraic formulas, whereas in reality the majority hold true only under certain conditions and for certain quantities of the variables involved." WEIERSTRASS fully accepted CAUCHY's program for making analysis more rigorous. Indeed, his two published contributions to the theory of matrices [1858, 1868] were an outgrowth of his acceptance of that program and of the inadequacy of generic reasoning. To see this, it is necessary to consider the work of LAGRANGE and LAPLACE.

In the course of applying the method of analysis to various problems in terrestrial and celestial mechanics, LAGRANGE and LAPLACE were led to the consideration of systems of linear differential equations with constant coefficients.

---

[2] Lest there be any misunderstanding, I wish to make it clear that my remarks about the generic reasoning of LAGRANGE and LAPLACE are not intended to be derogatory. At that moment in the history of mathematics the path they chose was the wise one. Like EULER before them, they grasped the revolutionary significance of the method of analysis and in their work they exploited it as best they could at the time. A good illustration of the fruitfulness of generic reasoning is given by LAGRANGE's justly celebrated memoir on the algebraic resolution of equations [1772/73] which gave birth to the theory of groups and to GALOIS' theory of equations.

For example, the system

$$A\ddot{Y}=BY. \tag{2.1}$$

In (2.1) A and B denote symmetric matrices with unspecified coefficients which are, respectively, positive and negative definite, and Y denotes a column matrix. The solution to (2.1) was known to involve the solutions to the system of linear equations

$$BX=\lambda AX. \tag{2.2}$$

The λ for which a non-trivial solution X exists were known to be the roots of an n^{th} degree polynomial with coefficients formed from those of A and B. Stability considerations required that these roots be real, but this seemed difficult to prove given the complicated nature of the polynomial. For a long time LAGRANGE and LAPLACE had to resort to an appeal to the physical context: since the mechanical system under consideration was stable, the solutions to (2.1) must be bounded as functions of t and hence the roots to (2.2) must be real. In the case of the planetary system the validity of such an argument was especially dubious.

A great advance was finally made by LAPLACE when he discovered (1787) that properties of the roots λ could be deduced from the symmetry properties of the coefficients of the differential equations, something we now take for granted. LAPLACE's discovery was based upon the following idea which I shall apply to the system (2.1), rather than the one actually studied by LAPLACE, because (2.1) is the focus of WEIERSTRASS' attention. If T and V are defined by

$$T=\sum_{i,j}A_{ij}\dot{y}_i\dot{y}_j, \qquad V=\sum_{i,j}B_{ij}y_iy_j, \tag{2.3}$$

then using the symmetry of A and B, one can deduce from $A\ddot{Y}=BY$ that $\dot{V}=\dot{T}$. Hence

$$T=V+C, \tag{2.4}$$

where C is a constant. From the positive and negative definiteness, respectively, of T and V, it follows that the solutions to (2.1) cannot be unbounded as functions of t. Hence the roots λ, which determine the solutions, must be real.

By virtue of the boundedness of the solutions, LAGRANGE and LAPLACE claimed that the characteristic roots λ, which, of course, occur in the solutions, must be real. But neither of them had pursued the study of the solution of systems of linear differential equations with sufficient care to justify their claim. They had no difficulty treating such a system when the characteristic roots are distinct, but their analysis of the case of multiple roots was inadequate. Given the generic tendency of their analytical methods, it is noteworthy that they considered the case at all. To the best of my knowledge, there is only one publication by either in which the solution to a system of linear differential equations is worked out under the hypothesis of a multiple root. It occurs in a memoir by LAGRANGE [1766, 520ff] where he solves the system

$$\ddot{Y}+BY=0 \tag{2.5}$$

under the assumption that the coefficients of B and the initial conditions are arbitrary. Since the views on multiple roots expressed by LAGRANGE and LAPLACE without demonstration in their subsequent publications agree with the conclusions reached in this memoir, it will serve as the basis of my discussion.

Most of LAGRANGE's elegant derivation of the solution to (2.5) is completely on the generic level. It was, however, manifest that the derivation and the final form of the solution depended upon the distinctness of the characteristic roots associated with the system

$$BX+\rho^2X=0. \tag{2.6}$$

LAGRANGE consequently asked how the form of the solution is affected by the assumption of a double root [1766, 528]. To obtain an answer, he assumed two roots of the form ρ_1 and $\rho_2=\rho_1+w$, where w is a "vanishing quantity". His idea was to substitute $\rho_2=\rho_1+w$ in the solution formulas obtained for distinct roots and then to use the vanishing nature of w to determine the form of solution when $\rho_1=\rho_2$. Except for the eventual assumption that $w=0$, LAGRANGE's approach remained entirely on the generic level and naturally led to a generic conclusion: "the terms of the value of $y^{(s)}$ [the s^{th} component of the solution Y to (2.5)] which correspond to equal roots will always contain the angle t" [1766, 529]. Hence in order for the solutions to be bounded functions of t for all initial conditions, it was deemed necessary that the roots ρ^2 of (2.6) be real, negative and unequal [1766, 530].

It is unclear whether LAGRANGE and LAPLACE were aware of the "general" nature of the above conclusion, but in their treatises on mechanics they affirmed without qualification the incompatibility of multiple roots and stable solutions to the corresponding system of differential equations. WEIERSTRASS' recognition of the questionable nature of their claims formed the starting point of the investigations that culminated in his theory of elementary divisors [1868]. In [1858] he obtained the general solution to the differential equations (2.1) and demonstrated that the existence of multiple roots is neither impossible nor problematical. His method of overcoming the algebraic difficulties presented by the possibility of multiple roots involved the key ideas subsequently developed further in his theory of elementary divisors.

3. Weierstrass' Paper of 1858

WEIERSTRASS' success in removing the misconceptions about multiple roots was due in part to the fact that he was able to draw upon mathematical developments that occurred in the 19th century and were consequently not available to LAGRANGE. WEIERSTRASS' work was especially influenced by that of three mathematicians of the preceding generation: CAUCHY, JACOBI and DIRICHLET. A work of fundamental significance in this connection is CAUCHY's memoir [1829] on the transformation of quadratic forms in any number of variables.

CAUCHY's memoir [1829] was motivated by the theory of quadric surfaces which was a standard component of the curriculum at the Ecole Polytechnique where CAUCHY taught. Although the results in CAUCHY's memoir were not inspired by the work of LAGRANGE and LAPLACE on systems of linear differential equations, he gave the first purely algebraic and completely rigorous proof that

the characteristic roots λ of the system $AX=\lambda X$, A symmetric, are real. CAUCHY's proof was essentially that given by LAGRANGE for $n=3$ in his analysis of the principal axis theorem of mechanics. The differences are, however, significant. CAUCHY succeeded in generalizing LAGRANGE's proof to n variables by using the theory of determinants, a new mathematical tool he had brought to perfection in an earlier memoir [1815]. One reason CAUCHY probably decided to generalize the principal axis theorem of mechanics and quadric surfaces was that it provided an occasion for him to apply his theory of determinants. A second difference from LAGRANGE's proof is that, whereas his is generic, CAUCHY's is not. By means of a few simple modifications the generic character of LAGRANGE's proof was avoided.

CAUCHY's generalized, rigorized version of LAGRANGE's proof was, however, not the centerpiece of his memoir. He reserved that position for the following two theorems:

Theorem I [1829, 187]. Let $S=|A-sI|$ and let $R, Q, \ldots$ denote the minors of S obtained by deleting, respectively, the first row and column, the first two rows and columns, etc. Then all the roots of $S=0$, $R=0$, $Q=0$, ... are real, and if the roots of $R=0$ are $r_1 \leqq r_2 \leqq \ldots \leqq r_{n-1}$, then the roots of $S=0$ are comprised, respectively, between the limits $-\infty, r_1, r_2, \ldots, r_{n-1}, \infty$.

Theorem II [1829, 194]. Given a homogeneous function of second degree

$$A_{xx}x^2+A_{yy}y^2+A_{zz}z^2+\cdots+2A_{xy}xy+2A_{xz}xz+2A_{yz}yz+\cdots$$

one can always substitute other variables $\xi, \eta, \zeta, \ldots$ which are related to $x, y, z, \ldots$ by linear equations such that $x^2+y^2+z^2+\cdots=\xi^2+\eta^2+\zeta^2+\cdots$ and the above homogeneous function becomes $s_1\xi^2+s_2\eta^2+s_3\zeta^2+\cdots$, where $s_1, s_2, s_3, \ldots$ are the roots of $S=0$.

CAUCHY was especially pleased with Theorem I because it provided information on the location of the roots s_i of $S=0$. As he himself noted (1829, 187] the simple algebraic proof he had sketched was valid only when no two successive equations in the series

$$S=0, \quad R=0, \quad Q=0, \ldots$$

have a root in common. His proof had, in effect, covered the generic case, and CAUCHY was fully aware of this limitation. In this respect his work contrasts with that of CHARLES STURM [1829] who had also discovered what amounts to Theorem I.

The objective of STURM's work was a new, strictly algebraic demonstration of the propositions about characteristic roots presented by LAPLACE and LAGRANGE. He considered the system of differential equations $A\dot{Y}+BY=0$, where A and B are symmetric. Under the hypothesis that A or B is definite he proposed to prove that the roots of $|sA+B|=0$ are real and distinct. STURM did not employ determinants and therefore limited his considerations to the case of five variables. By successively eliminating the components $x_1, x_2, \ldots, x_5$ of X in the equations $(sA+B)X=0$, he arrived at the analogues of CAUCHY's equations $S=0$, $R=0$, $Q=0, \ldots$.

STURM's analysis also proceeded on the generic level, although he exhibited a greater awareness of this fact than had LAGRANGE or LAPLACE. Regarding the

analogues of CAUCHY's functions S, R, Q, ..., which STURM writes in reverse order, he concluded that:

> The functions L, M, N, P, Q will have all their roots real and unequal.... The roots of each of these will comprise in their intervals the roots of the preceding function. It could happen, however, that for particular numerical values of the coefficients [of A and B] ... two consecutive functions may have one or more common roots; but this case does not make exception to the theorem. [1829, 317]

If the case to which STURM refers occurs, the conclusion about the distinctness of the characteristic roots is thrown in doubt. STURM returned to the matter further on:

> I must add that some of these theorems could be subject to exception if two or more consecutive functions among $LMNPQ$ should have common factors, a situation which cannot occur as long as the coefficients [of A and B] ... are indeterminate. [1829, 320]

These remarks exemplify the ambiguity accompanying generic analysis, although STURM's explicit admission of the generic nature of his theorems is not typical.

The admission may in fact have been prompted by a meeting he had with CAUCHY at which they learned they had discovered similar results. I have already suggested [1975, 21-22] that CAUCHY went away from this meeting aware of the relevance of his work to the theory of the secular perturbations of the planetary motions and that as a result came up with the title for his [1829]. It appears equally likely that CAUCHY called STURM's attention to the problems that can arise when successive equations $S=0$, $R=0$, $Q=0$, ... have a common root and prompted STURM's remarks.

In order to establish Theorems I and II while at the same time avoiding STURM's arguments based upon "the generality of algebra", CAUCHY proceeded, much as he had in *Cours d'analyse* [1821], to introduce limit considerations. His proof of Theorem II is especially interesting in connection with WEIERSTRASS' own approach in [1858]. Let

$$K=\prod_{i=1}^{n-1} S(r_i),$$

where S and the r_i are defined in Theorem I. When A is such that $K\neq 0$, $S=0$ and $R=0$ have no common roots and all the roots of A are distinct. CAUCHY gave a simple algebraic proof of Theorem II under the condition that A satisfy $K\neq 0$. In this case he specified formulas which determine the coefficient p_{ij} of the orthogonal substitution $\xi=PX$ up to sign, *viz.*

$$p_{ij}^2=F_{ij}^2\Big/\left(\sum_{j=1}^{n} F_{ij}^2\right) \tag{3.1}$$

where $F_{ij}=F_{1i}(s_j)$ and $F_{1i}(s)$ is the minor of $S=|A-sI|$ corresponding to the $(1, i)$ coefficient. Now K is, as CAUCHY observed, a symmetric function of $r_1, r_2, \ldots, r_{n-1}$ and by the properties of symmetric functions it follows that K is a polynomial in the coefficients A_{ij} of A. Hence "in general" $K\neq 0$. To cover the non-generic

case $K=0$, CAUCHY simply remarked that if the A_{ij} satisfy the condition $K=0$ then "it would suffice, in order to make it cease to hold, to attribute to one of the coefficients $[A_{ij}]$ in question an infinitely small increment ε; and since ε could be made to converge to zero without Theorem II ceasing to hold, it is clear that it still holds at the moment when ε vanishes" [1829, 195].

CAUCHY did not clarify his proof any further, but what he had in mind seems to be as follows. Suppose A is such that $K=0$. Then (a) let $A=\lim_{n\to\infty} A_n$ where $K(A_n)\neq 0$. For each A_n an orthogonal matrix P_n exists such that $X^tA_nX=\xi^tP_n^tA_nP_n\xi=\xi^tD_n\xi$ where D_n is diagonal. Thus (b) by taking limits one should obtain $P^tAP=D$, D diagonal, P orthogonal. CAUCHY's proof is in principle correct, but it is unclear whether he could have rigorously proved (b). Clearly (b) follows once it is established that: (c) the A_n may be chosen so that $P=\lim_{n\to\infty} P_n$ exists. But (c) does not immediately follow from CAUCHY's formulas (3.1). Probably for these reasons WEIERSTRASS in [1858] dismissed CAUCHY's proof as convincing only in the generic case.

WEIERSTRASS' attitude towards CAUCHY's proof is especially interesting because the approach is valid and easily justified by means of the BOLZANO-WEIERSTRASS Theorem that every bounded infinite subset of R_k has a limit point.[3] Since the orthogonal matrices form a bounded, closed subset of R_k, $k=n^2$, (c) follows readily from the Theorem. It is very doubtful that CAUCHY had this technique of a modern proof in mind. Indeed it was not until after WEIERSTRASS composed his [1858] that he began to develop systematically the foundations of analysis and make the BOLZANO-WEIERSTRASS Theorem one of the building blocks of the theory.[4] This turned out to be fortunate for WEIERSTRASS, because, failing to see the viability of CAUCHY's approach to the proof of Theorem II in the non-generic case, he sought to devise his own and, in the process, introduced the considerations that opened the way to his theory of elementary divisors.

WEIERSTRASS was unimpressed by CAUCHY's invocation of limits and impressed by JACOBI's elegant reworking of CAUCHY's Theorem II [1834]. JACOBI was attracted to CAUCHY's paper [1829] because the algebraic device employed, namely determinants, appealed to his love of elegant formalism. The reasoning in JACOBI's memoir [1834] was indeed formally elegant, but it was also generic. The apparent tediousness of specifically treating all singular cases no more appealed to JACOBI than the tediousness of reasoning with specific numbers had appealed to VIÈTE. In [1858], however, WEIERSTRASS showed that both generic reasoning *and* the tediousness of handling a myriad of special cases could be avoided without abandoning (as had CAUCHY) purely algebraic methods. The

---

[3] I am grateful to Professor RICHARD BEALS of the University of Chicago for calling this simple fact to my attention.

[4] According to DUGAC [1973, 56] WEIERSTRASS first began to study the foundations of analysis in his lectures in the fall semester 1858–1859, six months after his memoir [1858] was presented to the Berlin Academy. The earliest evidence of a systematic presentation of the foundations of analysis is his lectures as recorded by H. A. SCHWARZ during the Spring of 1861 at the *Gewerbeinstitut*[1973, 63]. The BOLZANO-WEIERSTRASS Theorem, stated for R_n and C_n, is basic to his lectures *Einleitung in die Theorie der analytischen Functionen* (1874) [1973, 77]. The proof for R_1 employs ideas used by BOLZANO in 1817 to prove that a set of real numbers with an upper bound has a least upper bound [1973, 74/75].

discovery of this fact is not only fundamental to [1858], it also involves the seminal ideas of the subsequent theory of elementary divisors.

WEIERSTRASS' interest in the transformation of quadratic forms, which lay outside his main area of research, may have been stimulated through his friendship with CARL BORCHARDT. BORCHARDT had written his doctoral dissertation under JACOBI's supervision and was primarily interested in the theory and application of determinants, a subject that JACOBI had popularized after acquiring an interest in it through CAUCHY's memoir [1829]. BORCHARDT and WEIERSTRASS first met in Braunsberg, East Prussia, in 1854 when WEIERSTRASS was still a Gymnasium professor. An account of that meeting was given by WEIERSTRASS in a letter to H. A. SCHWARZ shortly after BORCHARDT's death in 1880:

> At the time he was still a Privatdocent [in Berlin] but had already drawn to himself the attention of colleagues through two solid works. I still recall with great joy the pair of days we passed together. The pleasure they provided me can only be grasped by one who knows that from 1841 until then I had not spoken a mathematical word with anyone. (DUGAC, [1972, 158])

One of BORCHARDT's "solid works" was his paper "Neue Eigenschaft der Gleichung, mit deren Hälfe man die seculären Störungen der Planeten bestimmt" [1846] in which he used determinants to give a new, direct proof of the reality of the characteristic roots of a symmetric matrix. It seems likely that through that meeting and their subsequent contact in Berlin, where WEIERSTRASS moved the following year, WEIERSTRASS became interested in the theory of determinants and in the work of CAUCHY and JACOBI on its application to quadratic forms. (He had already studied, and been impressed by, LAPLACE's *Mécanique Céleste* during his years as a student in Bonn.) WEIERSTRASS adopted the elegant methods and notation of JACOBI but infused them with his critical approach to mathematics.

In [1858] WEIERSTRASS considered quadratic forms $\Phi = \sum_{ij} A_{ij} x_i x_j$, $\Psi = \sum B_{ij} x_i x_j$ where Φ is definite. CAUCHY and JACOBI had restricted their attention to the case $A = I$, but WEIERSTRASS treated the more general case so that his results would apply to the system of linear differential equations $A\dot{Y} = BY$. The main theorem WEIERSTRASS will prove is, stated for Φ positive definite, that a linear transformation $X' = PX$ exists such that $\Phi = \sum_{i=1}^{n} (x_i')^2$ and $\Psi = \sum_{i=1}^{n} s_i (x_i')^2$, where $s_1, s_2, \ldots, s_n$ are the roots of $f(s) = |sA - B|$. According to WEIERSTRASS, "this transformation of Φ, Ψ is one of the most interesting and important algebraic problems arising in diverse investigations." The proofs given by CAUCHY, JACOBI and others, he admitted, left nothing to be desired as long as the roots s_i are all unequal. But:

> it does not appear that special attention has been given to peculiar circumstances that arise when the roots of the equation $f(s) = 0$ are not all different; and the difficulties which they present — of which I was made aware by a question to be discussed more fully later — were not properly cleared up. I also at first believed that this would not be possible without extensive discussions in view of the large number of different cases that can occur. I was all the more pleased to find that the solution to the problem given by the

above-named mathematicians could be modified in such a way that it does not at all matter whether some of the quantities $s_1, s_2, \ldots, s_n$ are equal. [1858:234].

What was the question that, according to WEIERSTRASS, called to his attention the difficulties presented by multiple roots? He was referring to the question of the relation between multiple roots and stability, for he concluded his paper with a demonstration that when Φ and Ψ are positive and negative definite, respectively, the solutions to $A\ddot{Y}=BY$ are always stable, even though $f(s)$ may have multiple roots. Indeed I will show how this question probably led WEIERSTRASS to his realization that a simple modification of the CAUCHY-JACOBI proof is possible. Furthermore I will argue that WEIERSTRASS' approach to the transformation of quadratic forms, *i.e.* via the question about multiple roots and stability, involved some of the seminal ideas that he subsequently developed into his theory of elementary divisors [1868].

WEIERSTRASS' critical attitude towards the claims of LAGRANGE and LAPLACE was probably inspired, or at least reinforced, by a paper written by DIRICHLET [1846] which became an appendix to the third edition (1853) of LAGRANGE's *Mécanique analytique*. DIRICHLET was one of the earliest German mathematicians to accept CAUCHY's approach to analysis. He had also acquired an interest in mathematical physics through his contact with FOURIER in Paris in the 1820's. Combining his interest in mathematical physics with CAUCHY's critical approach to analysis, he produced several papers dealing with the mathematical foundations of physics, such as his paper [1829] on the convergence of FOURIER series. In [1846] DIRICHLET turned his critical eye upon a key proposition in LAGRANGE's *Mécanique analytique:* A state of equilibrium in a conservative mechanical system is stable if the potential function assumes a strict maximum value.[5] He pointed out that LAGRANGE's proof was decidedly circular and replaced it with a rigorous one which utilizes CAUCHY's conception of continuity.

According to WEIERSTRASS, in order to be convinced that the claims of LAGRANGE and LAPLACE are unfounded, "one need only recall Dirichlet's proof of the fundamental theorem of this theory" [1858, 244]. What he had in mind seems to be the following. DIRICHLET's proof naturally involved consideration of the potential and kinetic energy functions of the system, but the reasoning was applicable to a large class of continuous functions. For example it implied that if T and V are defined by $T=\sum_{ij} A_{ij}\dot{y}_i\dot{y}_j$, $V=\sum_{i,j} B_{ij}y_iy_j$ (equation (2.3)), if $T=V+C$ (equation (4)), and if T and V are positive and negative definite, respectively, then the y_i must be bounded as functions of t. DIRICHLET's proof therefore shows that the solutions to the differential equation $A\ddot{Y}=BY$ (equation (2.2)) studied by LAGRANGE [1788, Part II, Section V] must be bounded provided $T=V+C$. Since WEIERSTRASS was familiar with LAPLACE's discussion of systems of linear differential equations and stability, it is reasonable to assume he realized that $T=V+C$ could be deduced from the symmetry of the coefficients by a simple

[5] LAGRANGE's potential function is the negative of DIRICHLET's so that LAGRANGE spoke of a minimum [LAGRANGE, 1853, Pt. I, Sect. III, § V].

modification of LAPLACE's reasoning (as noted in connection with equations (2.2)–(2.4).[6]

DIRICHLET's proof thus made it clear that stability is not affected in this case by multiple roots. Now LAGRANGE and LAPLACE had deduced the incompatibility of multiple roots and stability from the nature of the solutions to systems of linear differential equations. Their understanding of the form of the solutions was therefore incorrect. What, then, did the solutions to (2.2) look like for multiple roots? To answer this question WEIERSTRASS, I suspect, turned to a memoir by CAUCHY [1840] (referred to by WEIERSTRASS [1858, 245]) on systems of linear differential equations, which contains what appears to be the first systematic study of the solutions to such equations since LAGRANGE's memoir [1766]. The main feature of CAUCHY's analysis, which he emphasized in the introduction, was that his method of solution applied equally well to the case of multiple roots—precisely the case of interest to WEIERSTRASS.

When CAUCHY had composed his memoir on the transformation of quadratic forms [1829] he was not particularly interested in systems of linear differential equations with constant coefficients. His paper was motivated by the principal axis theorem of mechanics and the theory of quadric surfaces. In the 1830's, however, he became increasingly interested in the problem of deriving the properties of light from a theory of the small vibrations of a solid elastic medium. (See WHITTAKER [1951, Ch. 5].) The groundbreaking work of FRESNEL in the 1820's was undoubtedly a factor behind CAUCHY's interest in this approach to the theory of light.

In developing the mathematics of his approach CAUCHY was naturally led to systems of linear differential equations, which in the simplest cases could be assumed to have constant coefficients [1839a]. At first CAUCHY limited his treatment to the case of distinct characteristic roots while briefly alluding to the "very simple" LAGRANGIAN method of handling the case of multiple roots. [1839a, 211–212, 220]. A few months later, he revised his opinion about the simplicity of that method [1839b]. As he now admitted, it was really a method designed to deal with distinct characteristic roots (a term introduced by CAUCHY in [1839b]). To cover the cases of multiple roots it had to be modified in an *ad hoc* manner for each possibility. CAUCHY could afford to be critical because he had discovered a uniform, completely general method of expressing the solutions. Deeming an exposition of his method to be valuable "not only to geometers but to physicists as well", he presented it in the pages of the *Comptes rendus* [1839b] and again in his *Exercises d'Analyse* [1840], where WEIERSTRASS read it.

CAUCHY's method utilized his calculus of residues – a calculus for which he was constantly seeking new applications.[7] I shall indicate the method as applied to the differential equation

$$\dot{Y}+BY=0 \tag{3.2}$$

---

[6] See pp. 13–14 of my [1975a].

[7] The pages of CAUCHY's private journal, *Exercises de mathématique* (1826–1830), are filled with such applications, including an application to the solution of the n^{th} order differential equation with constant coefficients [1826].

which CAUCHY himself considers first and with the greatest detail and generality. Let $f(s)=|sI+B|$ and let $f_{ij}(s)$ denote the minor corresponding to the (i,j) entry of $f(s)$. If $s_1, s_2, \ldots, s_n$ denote the characteristic roots then the solution to (3.2) satisfying $y_i(0)=a_i$ is given by

$$y_i(t)=\sum_{j,\,k=1}^{n} a_j \operatorname*{Res}_{s=s_k}\left[\frac{f_{ji}(s)}{f(s)}\,e^{st}\right], \quad i=1, 2, \ldots, n. \tag{3.3}$$

Although CAUCHY himself never pointed it out, his mode of expressing the solution reveals that the boundedness of the solution (3.3), insofar as it is affected by the occurrence of powers of t in the solution, depends upon the nature of the poles of the rational functions $f_{ji}(s)/f(s)$. Suppose, for example, that such a function has a pole of order m at $s=s_k$. By CAUCHY's own residue formula [1826, 28]:

$$\operatorname*{Res}_{s=s_k}\left[\frac{f_{ji}(s)}{f(s)}\,e^{st}\right]=\frac{1}{(m-1)!}\,\frac{d^{m-1}}{ds^{m-1}}\left[(s-s_k)^{m-1}\,\frac{f_{ji}(s)}{f(s)}\,e^{st}\right]_{s=s_k}.$$

Powers of t will thus occur in the solution from the differentiation of e^{st} unless $m=1$, *i.e.* unless all poles are simple. Similar remarks apply to $A\ddot{Y}=BY$ with $f(s)=|sA-B|$.

The above implication of CAUCHY's solution (3.3) can be expressed in another manner: powers of t never arise to destroy stability provided every root of $f(s)$ of multiplicity m is a root of multiplicity $\geqq m-1$ of each minor $f_{ji}(s)\not\equiv 0$. The same phenomenon, as WEIERSTRASS discovered, presents itself in connection with CAUCHY's Theorem II as treated by JACOBI [1834]. JACOBI showed that the coefficients p_{ij} of the orthogonal linear transformation of Theorem II satisfy the identities [1834, eqn. (36)]

$$p_{ik}p_{jk}=-\frac{g_{ij}(s_k)}{f'(s_k)} \tag{3.4}$$

where $f(s)=|B-sI|$ and $g_{ij}(s)$ is the (i,j) coefficient of the adjoint matrix corresponding to $B-sI$. (Hence $g_{ij}(s)$ is a constant multiple of the minor $f_{ji}(s)$.) Formula (3.4), with $i=j$, yields the squares of the desired coefficients p_{ik}. Of course, (3.4) becomes meaningless if s_k is a multiple root of $f(s)$ since $f'(s_k)=0$. As noted earlier, JACOBI was operating exclusively on the generic level and paid no attention to this possibility.

Some attention was paid to it by V.A. LEBESQUE in an exposition [1837] of the work of CAUCHY [1829], STURM [1829] and JACOBI [1834]. In a brief section on the non-generic case, LEBESQUE showed, using the symmetry properties of the coefficients, that if s_k is a multiple root of $f(s)$, then $f_{ji}(s_k)=0$ for all i and j. Thus formula (3.4) "can occur in the form 0/0 but not in the form $m/0$" [1837, 352]. Apparently satisfied with this observation, LEBESQUE chose to pursue the matter no further. The 0/0 form is of course not in itself reassuring. If s_k is a root of multiplicity m in $f(s)$, then it will be a root of multiplicity $m-1$ in $f'(s)$. Hence for 0/0 to be reassuring, it is necessary that each $f_{ji}(s)\not\equiv 0$ have s_k as a root of multiplicity $\geqq m-1$. In other words the poles of $f_{ji}(s)/f(s)$ for all i,j must be simple.

WEIERSTRASS grasped this implication of JACOBI's formula (3.4).[8] He saw in effect that *all* the problems posed by multiple roots would disappear if $f(s) = |sA - B|$ (A, B symmetric and A definite) has the property that for all i and j the poles of $f_{ji}(s)/f(s)$ are always simple. The demonstration [1858, 240] that this "remarkable circumstance" is always present under the above conditions is the key to WEIERSTRASS' paper.

In the terminology of the Introduction, WEIERSTRASS' "remarkable circumstance" is that if s_k is a root of multiplicity m of $D(s) = |sA - B|$ then it must be a root of multiplicity $\geqq m-1$ in $D_1(s)$. Hence

$$E_1(s) = D(s)/D_1(s) = \gamma_1 \prod_{i=1}^{n} (s - s_i)^{\varepsilon_i}, \qquad \varepsilon_i = 0 \quad \text{or} \quad 1.$$

Whether or not the characteristic equation $D(s)$ possesses multiple roots is irrelevant; what matters is that $E_1(s)$ not possess multiple roots. In effect, WEIERSTRASS' theorem on the transformation of the quadratic forms Φ and Ψ hinges upon the linearity of the elementary divisors of $s\Phi - \Psi$. The theorem is derived from the linearity of the elementary divisors of $E_1(s)$ and the linearity of those corresponding to $E_i(s)$ for $i > 1$ follows from the theorem itself. The seeds of WEIERSTRASS' theory of elementary divisors were already present in his paper of 1858. In the following section I will suggest the circumstances that led him to bring them to fruition in his paper of 1868.

4. The Theory of Elementary Divisors

By treating the transformation of the family of real quadratic forms $s\Phi - \Psi$ with Φ definite WEIERSTRASS had resolved the problems raised by the work of LAGRANGE and LAPLACE and by CAUCHY's Theorem II. But in his own mind the study of $s\Phi - \Psi$ must have raised further questions. For if Φ is not definite, the situation is completely different. The roots of $f(s) = |s\Phi - \Psi|$ need not be real and the poles of the $f_{ij}(s)/f(s)$ need not be simple. Simple examples can be given to illustrate these facts. The simultaneous transformation of Φ and Ψ into sums of squares would therefore not hold in general and the question arises as to when two families, $s\Phi - \Psi$ and $s\Phi' - \Psi'$, can be transformed into one another.

The question can also be posed for families of bilinear forms. Such forms had not been studied in the first half of the 19th century, but WEIERSTRASS had a good reason to be interested in them. His good friend BORCHARDT had recently arranged the posthumous publication of a work by JACOBI on bilinear forms [1857]. JACOBI suggested that theorems on the transformation of quadratic forms could be generalized to bilinear forms. His main theorem [1857, 589] presents a generic formula for the transformation of $f = \sum \alpha_{ik} x_i y_k$ into $f = \sum a_i u_i v_i$ which reduces to a known one for quadratic forms when $\alpha_{ik} = \alpha_{ki}$. JACOBI's work thus posed the challenge of a non-generic theory of the transformation of bilinear forms.

These problems were treated by WEIERSTRASS in his memoir "Zur Theorie der bilinearen und quadratischen Formen" [1868] in which he presents his theory of

[8] See especially p. 237 of [1858]. It is unclear whether WEIERSTRASS was familiar with LEBESQUE's paper [1837].

elementary divisors. In a footnote to the title he wrote that "[t]his work relates to one read in the Academy in March 1858 and is to be considered as a reworking and further development of it."[9] The work read in March of 1858 is, of course, the memoir [1858], which is the sole work WEIERSTRASS cites in [1868]. But why did a decade intervene between the publication of [1858] and its sequel?

An answer is supplied by WEIERSTRASS' student FROBENIUS, who was at the University when WEIERSTRASS' memoir [1868] appeared. According to FROBENIUS [1893, 719–720], WEIERSTRASS did go on in 1858 to consider the more general problem of the transformation of families of quadratic or bilinear forms:

> WEIERSTRASS solved this problem in the year 1858 by transforming the family of forms into an aggregate of elementary families which did not admit a further decomposition. Since the manuscript of his work was lost on a trip, however, he became so disgusted with the consideration of the subject that he first returned to it ten years later but then pursued the investigation in a completely different, entirely direct manner.

WEIERSTRASS' health was probably also a factor behind the termination of his research. Ill health forced him to cancel his lectures for the summer semester of 1859. The title of one of his proposed courses, incidentally, was "Die homogenen Functionen zweiten Grades mit beliebig vielen Veränderlichen".

What induced WEIERSTRASS to resume his study of forms after such a long hiatus? I suggest that the stimulation came from the work of two mathematicians, CHRISTOFFEL and KRONECKER, with whom WEIERSTRASS had contact. In the 1860's they published work on bilinear forms which renewed WEIERSTRASS' interest in the subject.

E. B. CHRISTOFFEL came to Berlin as Privatdocent in 1859 and remained there until 1862 when, upon the recommendation of WEIERSTRASS and KUMMER, he became Professor at the Polytechnikum in Zürich. He had also been a student at the University of Berlin, where he attended DIRICHLET's lectures on mathematical physics and was so impressed that he resolved to concentrate his own research in that area. His doctoral dissertation [1856] was on a problem in the theory of electricity. During his tenure as Privatdocent at the University he became interested in CAUCHY's work on the theory of light, referred to in Section 3. CAUCHY's mechanical model for light propagation led to a system of linear differential equations $\ddot{Y} = CY$. The coefficients of C were not necessarily constant, but CAUCHY limited his attention to that case, except for a few hints as to how to proceed more generally. CHRISTOFFEL therefore set himself the task of dealing more fully with the case of non-constant coefficients. His investigation was aided by the important discovery, made independently by HERMITE [1855] and CLEBSCH [1860] that certain properties of real symmetric systems and quadratic forms can be extended to forms with complex coefficients. In view of the importance of "HERMITEAN symmetry" in modern mathematics and physics as well as for CHRISTOFFEL's work the considerations that motivated its introduction are worth discussing.

[9] This footnote is missing from the version of [1868] in WEIERSTRASS' *Werke*. Other changes were also made by WEIERSTRASS, who oversaw the publication of the early volumes of his *Werke*. See notes 12 and 26.

HERMITE's work [1855] was an outgrowth of his research on the theory of numbers. The introduction of complex quantities into the theory of numbers was initiated by GAUSS who showed in his work on biquadratic residues [1832] that many arithmetical notions carry over to complex numbers $a+bi$ where a, b are ordinary integers. (*Cf.* Section 6.) It was natural to ask whether the arithmetical theory of forms could also be extended to GAUSSIAN integers. An answer was given by DIRICHLET [1842], who showed that many of GAUSS' results can be extended to the representation of GAUSSIAN integers by forms $f=\alpha z^2+2\beta zw+\gamma w^2$, where α, β, γ, z, w are GAUSSIAN integers and equivalence is defined using linear substitutions with GAUSSIAN integers as coefficients. In this manner the idea of considering forms with complex coefficients was introduced. HERMITE, however, showed that there was another, less straightforward, mode of generalization. That is, he indicated that many of the propositions in *Disquisitiones arithmeticae* have analogs for quadratic forms

$$f=\alpha zz+\beta z\bar{w}+\bar{\beta}\bar{z}w+\gamma w\bar{w},$$

where α and γ are real and β, z, w are GAUSSIAN integers.

HERMITE explained that he was led to study such forms through his interest in the representation of an integer as a sum of four squares [1854b:234]. The fact that any integer can be expressed as the sum of 4 squares had already been proved by LAGRANGE in 1770, but in the 19th century new proofs of this result were of interest, especially if they could shed light on the number of representations of an integer as a sum of four squares.[10] A HERMITEAN form is especially suitable for consideration of the four-square theorem since, as HERMITE stressed, it is actually a real quadratic form in the four variables x, y, u, v where $z=x+iy$ and $w=u+iv$. Perhaps HERMITE was led to introduce his HERMITEAN forms by observing that a sum of four squares can be regarded as a sum of two GAUSSIAN integer "squares", *i.e.* $x^2+y^2+u^2+v^2=z\bar{z}+w\bar{w}$. In any case, he obtained the following proof that any odd integer m is representable as a sum of four squares [1854b: 258–9]. Consider the form

$$f=mz\bar{z}+(a+bi)z\bar{w}+(a-bi)\bar{z}w+\left[\frac{a^2+b^2+1}{m}\right]w\bar{w},$$

where $a^2+b^2+1\equiv 0 \pmod{m}$. (HERMITE had proved that a and b exist in connection with another proof of the four-square theorem for odd integers [1853].) The form f represents m when $z=1$ and $w=0$. Furthermore, since its determinant is $+1$, it followed from the arithmetical theory as developed by HERMITE that f is arithmetically equivalent to $Z\bar{Z}+W\bar{W}$. Hence m is representable by this form and therefore as a sum of four squares.

Although HERMITE's discussion of HERMITEAN forms in [1854b] was limited to the case of two variables, he was familiar with CAUCHY's paper on the transformation of quadratic forms [1829] and interested in the n-variable theory. (See in that connection HERMITE [1854a].) Thus he later noted that the characteristic roots remain real in the n-variable HERMITEAN symmetric case [1855] and

[10] See DICKSON [1919, v. 2, 168].

that $\Phi = \sum_{i,j=1}^{n} a_{ij} x_i \bar{x}_j$, $a_{ji} = \bar{a}_{ij}$, can be transformed into a sum of squares [1856]. These theorems on HERMITEAN forms also interested HERMITE because of their relation to theorems of CAUCHY and STURM in the theory of equations. (See in this connection HERMITE [1856, 412] and [1912].)

CLEBSCH also discovered that the characteristic roots of a HERMITEAN symmetric system of equations are real, and it is likely that his work impressed CHRISTOFFEL more than that of HERMITE, for CLEBSCH was investigating CAUCHY's theory of light when he made the discovery. CAUCHY's method of integrating the differential equation derived from his model was first to obtain particular solutions of a specific form. (See CAUCHY [1850].) This in turn required consideration of a system of equations $CX = \lambda X$, where the coefficients of C are in general complex. (CAUCHY made extensive use of complex numbers in his work on optics.) In working out the details of CAUCHY's theory in a special case, CLEBSCH ended up with a 3×3 matrix of coefficients of the form $C = A + \sqrt{-1}\,B$ where A and B have real coefficients, A is symmetric and B is skew symmetric ($B^t = -B$). He discovered that the characteristic equation "has remarkable properties whereby it and equation (4) [$CX = \lambda X$] appear as extensions of those equations that arise in the theory of secular perturbations and in so many other investigations" [1860, 326].

CLEBSCH proved the reality of the characteristic roots of $C = A + \sqrt{-1}\,B$ for the n-variable case. Probably for the benefit of mathematicians not interested in optics, he also published a generalized version of his results in [1863]. (Here he also noted that the fact that the characteristic roots of C are real implies, with $A = 0$, that the roots of a real skew symmetric system B are pure imaginaries.) HERMITEAN symmetric systems were thus linked with physical problems long before the advent of matrix mechanics. A further link was supplied by CHRISTOFFEL [1864b]. In seeking to justify CAUCHY's brief remarks about the integration of $\ddot{Y} = CY$ when C is not constant, he showed that by using the idea that the solutions to $\ddot{Y} = CY$ should resemble averages, the solution of that system can be reduced to solving a system with constant coefficients of the form

$$A\ddot{Y} + BY = 0, \tag{4.1}$$

where $A = (\rho_i \delta_{ij})$, $\rho_i > 0$, and the coefficients of B are, in general, complex. CHRISTOFFEL showed that the coefficients of B have the symmetry property called to attention by HERMITE and CLEBSCH.

The problem of integrating (4.1) was therefore analogous to the problem studied by WEIERSTRASS in [1858], and the analogy prompted CHRISTOFFEL to devote a separate, purely mathematical, paper [1864a] to the generalization of WEIERSTRASS' results to the context of bilinear forms $\Phi = \sum \alpha_{gh} u_g v_h$, $\Psi = \sum \beta_{gh} u_g v_h$ where $\alpha_{hg} = \bar{\alpha}_{gh}$, $\beta_{hg} = \bar{\beta}_{gh}$, $v_g = \bar{u}_g$, and Φ is definite. WEIERSTRASS was familiar with CHRISTOFFEL's work, and it undoubtedly helped reawaken his interest in his own abandoned investigation of generalizations of the results in [1858], especially since he found CHRISTOFFEL's methods of proof to be neither purely algebraic nor sufficiently rigorous (WEIERSTRASS [1879, 140n]). Additional stimulation was provided by WEIERSTRASS' colleague, KRONECKER.

KRONECKER had obtained his doctorate from Berlin in 1845 with a dissertation done under the supervision of KUMMER. By the time WEIERSTRASS had been appointed to a professorship at the University in 1856, KRONECKER had settled in Berlin as an independently wealthy scholar-in-residence. When WEIERSTRASS was working on [1858], KRONECKER was occupied with the problem of obtaining a positive result about the solvability of the general quintic by means other than radicals. His work led to the consideration of the complex multiplication of elliptic functions. As KRONECKER explained in his inaugural address before the Berlin Academy [1861: 389]: "The fact that at the time when I was concerned with that chapter of the theory of elliptic functions, I had already established friendly relations with Herr Weierstrass, who is master of this area of analysis, had a significant influence on my work. . . ."

WEIERSTRASS continued to influence KRONECKER's work by suggesting to him the possibility of algebraic problems related to the theory of theta functions in several variables. He was not disappointed for he was led, by an ingenious chain of reasoning, to an interesting problem on the transformation of bilinear forms [1866]: Given a bilinear form in $4n$ variables, $M(x,y)=\sum_{i,j=1}^{2n} a_{ij}x_iy_j$, determine a non-singular linear transformation of the variables $x_i=\sum p_{ij}x'_j$, $y_i=\sum c_{ij}y'_j$, such that $M(x',y')$ vanishes identically when $x'_{n+r}=y'_{n+r}=0$ for $r=1,2,\ldots,n$. KRONECKER observed that

$$J(x',y')=\sum_{i=1}^{n}\lambda_i x'_i y'_{n+i}+\sum_{i=1}^{n}\lambda_{n+i}x'_{n+i}y'_i$$

satisfies the above requirement so that if $M(x,y)$ could be transformed into this "normal form" (as he called it) the problem would be solved.

Unfortunately such a transformation is not always possible. From the viewpoint of WEIERSTRASS' memoir [1868], $sI-M$ and $sI-J$ would have to have the same elementary divisors, but all the elementary divisors of $sI-J$ are linear whereas this need not be the case for $sI-M$. Expressed in the terms of WEIERSTRASS' memoir [1858], the poles of $f_{ij}(s)/f(s)$, where $f(s)=|sI-J|$, have the special property of being simple, and this property is preserved under linear transformations by a theorem of CAUCHY's on determinants (see below). KRONECKER, however, proposed to solve the problem by proving that, under certain conditions, the transformation was possible. His approach is of interest relative to WEIERSTRASS' work [1868].

KRONECKER pointed out that the problem of transforming $M=X^tAY$ into J by "cogredient transformation" $X=PX'$, $Y=PY'$ was closely related to the general problem of the cogredient transformation of M into $N=(X')^tBY'$. In terms of matrix symbolism (which KRONECKER did not use[11]), the problem is when are A and B congruent: $B=P^tAP$, $|P|\neq 0$. KRONECKER observed, in effect, that $B=P^tAP$ implies $B^t=P^tA^tP$ so that, if u and v are variables,

$$uB+vB^t=P^t(uA+vA^t)P. \tag{4.1}$$

Thus the problem reduces to the problem of when two families of bilinear forms are equivalent under cogredient transformation of the variables.

[11] It was in [1866], however, that KRONECKER first introduced the "KRONECKER delta" notation δ_{ij} [1866, 150].

From (4.1) KRONECKER deduced

$$|B||uA+vA^t|=|A||uB+vB^t|, \tag{4.2}$$

which thus expresses a necessary condition for the transformation to be possible. To solve his problem he therefore sought to establish the sufficiency of (4.2). Perhaps under WEIERSTRASS' influence, he pointed out that his proof of sufficiency was made under the condition that the determinants $|uA+vA^t|$, $|uB+vB^t|$ factor into *distinct* linear factors [1866, 149]. KRONECKER's work thus involved a question on the equivalence of families of bilinear forms which, from WEIERSTRASS' point of view, he had resolved to exactly the same extent as CAUCHY and JACOBI had resolved their problem on the transformation of quadratic forms. In his inaugural speech KRONECKER had referred to WEIERSTRASS as the master of elliptic functions, but he was also the master of multiple roots as he showed once again in [1868].

In [1868] WEIERSTRASS considered the general problem of determining the necessary and sufficient conditions under which one pair of bilinear forms $P=\sum A_{\alpha\beta}x_\alpha y_\beta$, $Q=\sum B_{\alpha\beta}x_\alpha y_\beta$ can be transformed into another pair P', Q' by means of non-singular linear transformations $X=HX'$, $Y=KY'$. Following KRONECKER's lead, or so it would seem, he introduced the family $pP+qQ$, where p and q are variables, and the determinant $D(p,q)=|pP+qQ|$. The case in which $D(p,q)$ vanishes identically is excluded because "I knew that while I was writing this work, Herr Kronecker had undertaken an investigation of it."[12] WEIERSTRASS observed that the necessary condition for the equivalence of $pP+qQ$ and $pP'+qQ'$ is

$$|pP'+qQ'|=|H|\ |K|\ |pP+qQ|. \tag{4.3}$$

Equation (4.3) is the analogue of KRONECKER's (4.2), but the work he had done ten years earlier enabled him to look more deeply into the structure of $D(p,q)=|pP+qQ|$ for the sufficient condition for equivalence that obtains regardless of whether multiple linear factors exist.

The sufficient condition is that the two determinants $|pP+qQ|$, $|pP'+qQ'|$ have the same elementary divisors. The elementary divisors of $D(p,q)=|pP+qQ|$ are defined similarly to those of $D(s)=|s\Phi-\Psi|$. (See the Introduction.) In fact WEIERSTRASS expresses $pP+qQ$ as $s\Phi-\Psi$, $|\Phi|\neq 0$, and works with the latter family. I shall therefore discuss his work in the context of the families $s\Phi-\Psi$.

Exactly when WEIERSTRASS formulated the concept of elementary divisors is unclear, although obviously its origins are to be found in [1858]. Once he had formulated it, it would have been a simple matter to see that the determinants $D_i(s)$, and hence the elementary divisors, are unchanged by linear transformation of the variables. This follows from the theorem, widely known by 1868, that if $C=AB$, then any $k\times k$ minor of $|C|$ is a sum of products of $k\times k$ minors of $|A|$ and $|B|$.[13] From this theorem it follows that when $s\Phi-\Psi$ and $s\Phi'-\Psi'$ are equi-

[12] WEIERSTRASS [1895, 19n]. This footnote is not in the original version of [1868]. KRONECKER's paper [1868] appeared immediately following WEIERSTRASS' in the Berlin *Sitzungsberichte*.

[13] CAUCHY proved this generalization of the multiplication theorem of determinants in his memoir on determinants [1815, 164], and it was presented in the leading German-language texts on determinants such as BRIOSCHI [1856, 55] and BALTZER [1864, 43]. (The first edition of BALTZER [1857, 24] states the result for $(n-1)\times(n-1)$ minors.) The result also follows from a theorem due to JACOBI [1841, Prop. IV], although the implication is not immediate. SYLVESTER [1851b] rediscovered the result in the course of his study of the contact of quadric surfaces. (See Section 5.)

valent, every $k \times k$ minor of $|s\Phi' - \Psi'|$ is a linear combination of $k \times k$ minors of $|s\Phi - \Psi|$ and conversely. Thus a necessary condition for equivalence is the equality of the elementary divisors.

The real problem was to prove the sufficiency of the condition and thereby a theorem "whereby a principal question in the theory of forms is settled" [1868, 33]. The details of the demonstration are difficult to follow because they involve considerable analytic formalism and many properties of determinants. But the underlying idea is simple enough: show that $s\Phi - \Psi$ can be put in a normal form by a linear change of variables. In describing the normal form I shall follow WEIERSTRASS' notation. Let $(s-c_\lambda)^{e_\lambda}$, $\lambda = 1, 2, \ldots, \rho$, denote the elementary divisors of $s\Phi - \Psi$. Then the original x and y variables can be transformed linearly into new variables $X_{\lambda\mu}$, $Y_{\lambda\nu}$, $\mu, \nu = 0, 1, \ldots, e_\lambda - 1$. The double index notation is used to indicate that each elementary divisor determines an independent set of variables. In terms of the new variables,

$$\Phi = \sum_{\lambda=1}^{p} (X_\lambda Y_\lambda)_{e_\lambda},$$

$$\Psi = \sum_{\lambda=1}^{p} c_\lambda (X_\lambda Y_\lambda)_{e_\lambda} + (X_\lambda Y_\lambda)_{e_\lambda - 1}$$

where for any integer e, $(X_\lambda Y_\lambda)_e$ denotes the bilinear form

$$(X_\lambda Y_\lambda)_e = \sum_{\mu+\nu=e-1} X_{\lambda\mu} Y_{\lambda\nu}.$$

WEIERSTRASS' normal form is essentially the JORDAN Canonical Form. For example, if we represent the λ^{th} bilinear form occurring in the decomposition of Ψ as $\sum_{\mu,\nu} w_{\mu\nu} X_{\lambda\mu} Y_{\lambda\nu}$, then the matrix $(w_{\mu\nu})$ has the form:

$$\begin{pmatrix} 0 & \ldots & 0 & 0 & 1 & c_\lambda \\ 0 & \ldots & 0 & 1 & c_\lambda & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ c_\lambda & \ldots & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4.4}$$

WEIERSTRASS concluded his paper by specializing his results to quadratic forms P, Q and $H = K$. He showed that the necessary and sufficient condition for two families of quadratic forms (with real or complex coefficients) to be transformable into one another is that they possess the same elementary divisors. Finally, he turned to the question raised by his results in [1858]: When can P and Q be transformed simultaneously into sums of square terms? The answer could now be stated simply: If and only if the elementary divisors of $pP + qQ$ are linear. The last theorem in [1868] is a reformulation of the main theorem of [1858]: If P and Q are real quadratic forms such that $pP + qQ$ is definite for some p and q, then the elementary divisors of $pP + qQ$ are linear and P and Q can be transformed simultaneously by a real linear transformation into sums of square terms.

KRONECKER was greatly impressed by WEIERSTRASS' memoir [1868], which inspired him to compose a number of papers dealing with the transformation

of quadratic and bilinear forms and, in particular, with the singular case $|pP+qQ| \equiv 0$. Whereas KRONECKER had remained content to treat the "general case" of distinct linear factors in [1866], WEIERSTRASS' work showed him that it was possible — and necessary — to extend the methods of analysis beyond the generic case. As KRONECKER explained with WEIERSTRASS' memoirs expressly in mind:

> It is common — especially in algebraic questions — to encounter essentially new difficulties when one breaks away from those cases which are customarily designated as general. As soon as one penetrates beneath the surface of the so-called generality that excludes every particularity into the true generality, which comprises all singularities, the real difficulties of the investigation are usually first encountered but, at the same time, also the wealth of new viewpoints and phenomena which are contained in its depths.[14]

KRONECKER's remarks form a fitting conclusion to the discussion of WEIERSTRASS's work and its background, for they capture the spirit in which it was undertaken and the message which it conveyed to his colleagues and students.

5. Related Developments in England and France

WEIERSTRASS' theory of elementary divisors is closely related to certain discoveries made in England by J. J. SYLVESTER, ARTHUR CAYLEY and H. J. S. SMITH and in France by C. JORDAN. These discoveries must be taken into account in order to clarify WEIERSTRASS' role in the history of the theory of matrices. I shall begin with the English work because it was done before 1868.

SYLVESTER's work grew out of his interest in the contact of conics which began with his paper [1850]. The contact of conics had already been extensively discussed from the standpoint of analytic geometry by PLÜCKER [1828] and others, but SYLVESTER, apparently under CAYLEY's influence, began to approach the subject systematically from the standpoint of the theory of determinants. In homogeneous coordinates the equation of a non-degenerate conic is given by a real quadratic form

$$U=\sum_{i,j=1}^{3} a_{ij}x_ix_j=0, \qquad |a_{ij}|\neq 0.$$

If $V=0$ is a second non-degenerate conic, the points of intersection of the two lie on all conics of the pencil $U+\lambda V=0$. In particular, they must lie on the conics $U+\lambda V=0$ which are degenerate in the sense that $D(\lambda)=|U+\lambda V|=0$. Contact between the conics $U=0$ and $V=0$ occurs when $D(\lambda)$ has a multiple root. Four different types of contact are possible.

All of the above facts were known in one form or another, but SYLVESTER appears to have been the first to express them systematically in terms of deter-

[14] KRONECKER [1874a, 405]. The above quotation appears as a caption to EDUARD WEYR's *O theorii forem bilinearných* [1889] which is missing from the German version [1890]. WEYR's work, first announced in the *Comptes rendus* of the Paris Academy [1885a, 1885b], was done in response to the inconclusive, generic remarks about matrix algebra made by SYLVESTER. The use of KRONECKER's remarks as a caption indicates how much WEYR's work, which involves introduction of the WEYR characteristics and a somewhat more geometrical (*i.e.* vector space) approach to WEIERSTRASS' theory, was influenced by the viewpoint of the Berlin school. *Cf.* note 15 of my [1977]. I am grateful to Professor HANS SCHNEIDER of the University of Wisconsin for showing me a copy of WEYR [1889].

minants and homogeneous coordinates. PLÜCKER, for example, used neither in [1828] even though he was one of the first to introduce homogeneous coordinates into analytic geometry. The superiority of SYLVESTER's approach is evident in the simple algebraic characterizations he obtained for the known types of contact. Two types of contact occur when $D(\lambda)$ has a double root and two when $D(\lambda)$ has a triple root. To distinguish the possibilities further he found that it was necessary to consider $D_1(\lambda)$, the greatest common divisor of the first minors (*i.e.* 2×2 minors) of $D(\lambda)$. The two types of contact occurring when $D(\lambda)$ has a double root $\lambda = -a$ are distinguished by the fact that for one type $D_1(\lambda) = (\lambda + a)$ and for the other $D_1(\lambda) = 1$. Likewise the types of contact when $D_1(\lambda) = (\lambda + a)^3$ depend on whether $D_1(\lambda) = (\lambda + a)$ or $D_1(\lambda) = 1$. (The possibility that $D(\lambda) = (\lambda + a)^3$ and $D_1(\lambda) = (\lambda + a)^2$ is ruled out by a relation discovered by SYLVESTER and discussed below.)

Encouraged by the elegance of his algebraic characterizations of the types of contact of conics, SYLVESTER decided to apply it to the study of the contact of quadric surfaces. Expressed in homogeneous coordinates, the equations of quadric surfaces $U = 0$, $V = 0$ are quadratic functions of four variables. No one had made an exhaustive study of the types of contact possible between quadric surfaces. SYLVESTER made the assumption that, by analogy with the contact of conics, there would be a one-to-one correspondence between types of contact and possible factorizations of $D(\lambda) = |U + \lambda V|$, $D_1(\lambda)$ and $D_2(\lambda)$, where $D_2(\lambda)$ is the greatest common divisor of the second (or 2×2) minors of $D(\lambda)$. This assumption was supported by his discovery that, for quadratic forms U, V in any number of variables, the $D_i(\lambda)$ are invariant with respect to linear transformations.

To support further his assumption and the feasibility of his approach SYLVESTER obtained simple expressions for U and V called "condensed forms" to illustrate each possible factorization of the $D_i(\lambda)$. Working with these condensed forms, he discussed the geometrical nature of the contact. For example, for $D(\lambda) = (\lambda + a)^2(\lambda + c)^2$, $D_1(\lambda) = (\lambda + a)$, $D_2(\lambda) = 1$, the corresponding condensed forms are

$$U = ax^2 + ay^2 + bz^2 + czt, \qquad V = x^2 + y^2 + z^2 + zt.$$

The 4 types of contact for conics and the 12 types for quadric surfaces are systematically studied from the above point of view in SYLVESTER's memoir [1851a], which represents an original and impressive contribution to analytical geometry.

It should be clear that the validity of SYLVESTER's method of classifying types of contact depends upon the theorem that two pairs of quadratic forms with the same $D_i(\lambda)$ are simultaneously transformable into one another and hence represent the same pair of quadric surfaces in two different coordinate systems. An explicit statement of the theorem is missing from his work, although its converse (the invariance of the $D_i(\lambda)$) is stressed and a separate paper [1851b] devoted to the proof. Although the lack of an explicit statement of the theorem is somewhat surprising, the lack of a proof is not. In all their work pertaining to the theory of matrices both SYLVESTER and CAYLEY exhibited a naive conception of mathematical demonstration. In particular, they frequently relied upon the evidence of incomplete induction when a general proof was not readily forthcoming.[15] SYLVESTER

[15] See my [1977], especially Section 4 and 6.

appears to have relied upon such inductive evidence to justify the validity of his method for quadric surfaces. Thus after treating the contact of conics he wrote: "Fortified by the successful application of our general principles to the preceding more familiar case of contact [of conics], we are now in a position to apply with greater confidence the same *a priori* method to the exhaustion and characterization of all the varied species of contact possible between surfaces of the second order..." [1851a, 226–227].

SYLVESTER made one further discovery that facilitated his method of classification. If d_i denotes the multiplicity of $\lambda+a$ in $D_i(\lambda)$, $i=0,1,2,\ldots$, then $d_{i-1} \geqq 2d_i - d_{i+1}$ [1851a, 222]. The importance of the inequality was that it limited the possibilities for the factorizations of the $D_i(\lambda)$ and hence the types of contact. SYLVESTER had no occasion to introduce the polynomials $E_i(\lambda)=D_{i-1}(\lambda)/D_i(\lambda)$, but in terms of the $E_i(\lambda)$ his identity implies the important relation $E_{i+1}(\lambda)|E_i(\lambda)$. His study of contact therefore involved the equivalent of the concept of elementary divisors and their fundamental properties. What was lacking was a theory of elementary divisors — a careful articulation and proof of the underlying algebraic theorem on the transformation of quadratic forms.

Although SYLVESTER restricted his attention in [1851a] to "the space cognizable through our senses", and thus to quadratic forms in at most 4 variables, he realized the entire theory of contact could be generalized to higher dimensions, and he even indicated the number of possibilities for the $D_i(\lambda)$ for 5 through 8 variables. He also indicated another avenue of generalization. Since determinants are not restricted to symmetric arrays of numbers, "[t]he theory of quadratic functions merges in a larger theory of binary functions... [The determinant of a binary function] would correspond to the square matrix[16]

$$\begin{matrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c''. \end{matrix}$$

Almost all the properties of quadratic apply, with slight modification, to binary functions" [1851a, 222n].

CAYLEY had personally witnessed and encouraged the development of SYLVESTER's ideas on contact, and in a note [1855c] he briefly commented upon his friend's allusions to a more general theory. The note was one of a series [1855a, 1855b] which foreshadowed CAYLEY's memoir on matrix algebra [1858].[17] In [1855c] CAYLEY considered $D(\lambda)=|a_{ij}(\lambda)|$ and the $D_i(\lambda)$ for a matrix $(a_{ij}(\lambda))$ where $a_{ij}(\lambda)$ is linear in λ. After showing how to compute the number of possible factorizations of the $D_i(\lambda)$ and correcting the values given by SYLVESTER, he pointed out that the binary or non-symmetric case alluded to by SYLVESTER also had geometrical applications:

> All this, however, applies to another geometrical theory, namely the theory of homographic figures. To fix the ideas, I consider only figures in the plane. If

[16] SYLVESTER introduced the term "matrix", one of the few new mathematical terms devised by him that has been generally adopted. Of course, Continental mathematicians had been calling the same things "systems" or "tableaux" before SYLVESTER conceived his own word for it. See *e.g.* CAUCHY [1815, 1829].

[17] See Section 3 of my [1977].

> x,y,z are the [homogeneous] coordinates of a point and $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ are linear functions of x,y,z, the coordinates of a point *homographic* to the point (x,y,z) will be $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$. By seeking the points homographic to themselves, one is led to the equations $\boldsymbol{x}-sx=0$, $\boldsymbol{y}-sy=0$, $\boldsymbol{z}-sz=0$. The quantities on the left $\boldsymbol{x}-sx$, $\boldsymbol{y}-sy$, $\boldsymbol{z}-sz$ are linear functions of x,y,z having linear functions of s as coefficients. One thus has a matrix all the terms of which are linear functions of s; the entire theory is connected with the properties of this matrix [1855c, 219].

CAYLEY then proceeded to note that the case of "ordinary homography" corresponds to $D(\lambda)$ ($D(s)$ in the notation of the quotation) having distinct roots, that $D(\lambda)=(\lambda-a)^2(\lambda-b)$, $D_1(\lambda)=(\lambda-a)$ corresponds to homology and that the other possibilities correspond to "less general cases". Thus CAYLEY could conclude that for the plane there are 6 types of homography and for space 14 types.

Both CAYLEY & SYLVESTER promised to return to the subjects of their respective papers and to discuss the n-variable case, but they never did. Expressed in modern notation, CAYLEY's problem would have been to study homographies (or collineations) $X'=AX$ in n-dimensional space where each type is assumed to correspond to a factorization of the invariants $D_i(\lambda)$ of $D(\lambda)=|A-sI|$. Perhaps, following SYLVESTER's lead, he would have attempted to specify a "condensed form" for each type. The condensed forms would correspond to a change of variables $X=PY$, $|P|\neq 0$, so that $Y'=BY$ is the "condensed form" corresponding to $X'=AX$. Since $B=P^{-1}AP$ CAYLEY would have been concerned with matrix similarity, whereas SYLVESTER's work involved (implicitly) matrix congruence. A few years later H. J. S. SMITH of Oxford University introduced [1861a] the arithmetical analog of the $D_i(\lambda)$ within the context of matrix equivalence ($PAQ=B$).

SMITH was concerned with the arithmetical problem of integral solutions to a system of congruences with integral coefficients

$$\sum_{j=1}^{n} A_{ij}x_j \equiv A_{i,n+1} \pmod{M}, \qquad i=1,2,\ldots,n.$$

His main result is the following. Let $\nabla_n,\nabla_{n-1},\ldots,\nabla_1,\nabla_0$ denote, respectively, the $n\times n$ determinant $|A_{ij}|$, the greatest common divisor of the first minors, *etc.*, and $\nabla_0=1$. Define $D_n,D_{n-1},\ldots,D_0$ analogously for the augmented matrix obtained by addition of the column $A_{i,n+1}$ $i=1,2,\ldots,n$. Then the necessary and sufficient condition that a solution should exist is that $m=\mu$, where $m=\prod_{i=1}^{n} d_i$, $\mu=\prod_{i=1}^{n} \delta_i$, and d_i and δ_i are, respectively, the greatest common divisors of M and D_i/D_{i-1} or M and ∇_i/∇_{i-1}. Also $m=\mu$ specifies the number of solutions.

To prove the above result SMITH established the following lemma, which is presented in his own notation: If $||A||$ is any square matrix of integers, there exist two matrices, $||\alpha||$ and $||\beta||$, of determinant 1 such that

$$||A|| = ||\alpha|| \times \begin{Vmatrix} \nabla_n/\nabla_{n-1} & 0 & \cdots & 0 \\ 0 & \nabla_{n-1}/\nabla_{n-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \nabla_1/\nabla_0 \end{Vmatrix} \times ||\beta||. \tag{5.1}$$

The diagonal entries in (5.1) are the arithmetical analogues of the $E_i(\lambda)=D_{i-1}(\lambda)/D_i(\lambda)$ of WEIERSTRASS' theory. (See Section 6.)

As the notation and terminology of the above lemma suggest, SMITH was familiar with CAYLEY's memoir on matrices [1858],[18] but nowhere does he refer to the earlier papers of CAYLEY and SYLVESTER where analogues of the ∇_i are introduced. Either he was unaware of them or he simply failed to envision the possibility of applying his results to supply the missing theoretical foundation for their work. SMITH returned to the subject of his memoir [1861a] again in [1873], but there is no awareness of a connection with the subject of WEIERSTRASS' memoir [1868].

In the writings of SYLVESTER, CAYLEY and SMITH between 1851 and 1861 it is thus possible to perceive most of the essential ingredients of the theory of matrices: the equivalent of the concept of elementary divisors and their properties, the elements of matrix algebra (CAYLEY [1858]), geometrical applications which involved problems in matrix congruence and similarity and, finally, the arithmetical foundations of the entire theory. But it would be wrong to conclude that the theory of matrices originated in England in the decade 1851–1861. Such a conclusion is historically untenable for two reasons. Briefly stated, they are as follows: (1) It is *in retrospect* that the essential ingredients of the theory of matrices are visible in their work. At the time these "ingredients" were not related and developed as parts of a unified theory. (2) Their work exerted little, if any, influence upon the emergence of the theory of matrices.

In connection with point (1) I have already noted SMITH's failure to relate his results to the work of CAYLEY and SYLVESTER. Conversely, it is doubtful that CAYLEY and SYLVESTER could have appreciated the significance of SMITH's work because they were insensitive to the need of a firmer theoretical foundation for their ideas. They also failed to relate their work on the invariants $D_i(\lambda)$ to their investigation of matrix algebra. They never formulated the problems on matrix congruence and similarity implicit in their geometrical work in the language of the calculus of matrices. Nor did they appreciate the relevance of the invariant factors $D_i(\lambda)$ to the problems posed by matrix algebra such as that of determining the matrices which commute with a given matrix. They tended to study matrix algebra on the generic level.

Thanks to his involvement with the contact of conics, SYLVESTER was made aware of the limitations of the customary generic approach to the transformation of quadratic forms. Thus he commenced his paper [1851a] with the following observations:

> It is well known that in general any two homogeneous quadratic functions of the same system of variables may be simultaneously transformed, so as to be expressed each of them as pure quadratic functions of a new system of variables equal in number and linearly connected with the original ones; a pure quadratic function meaning one in which only squares of the variables are retained. ... In order that the reduction of any such system, say U and V, to a pure quadratic form may be possible (and it generally is) it is necessary that none of the roots of the complete determinant of $U+\lambda V$ be equal; if any relation of equality

[18] He refers to it in his report on the theory of numbers to the BAAS [1861b, 167].

exists between these roots, the *general* reduction is *generally* no longer possible. . . ."

SYLVESTER also perceived the relevance of the $D_i(\lambda)$ for the determination of when the generic theorem is valid. But when it came to treating matrix algebra CAYLEY and SYLVESTER resorted to the generic mode of reasoning. Even after SYLVESTER discovered the limitations of CAYLEY's solution to the above problem on commuting matrices, he failed to see the relevance of the invariant factors.[19] Furthermore he never relinquished the generic approach to matrix algebra. He was made aware of some of the limitations of that approach to algebra but, unlike WEIERSTRASS, indisposed to wrestle with the algebraic problems of the non-generic case.

The explanation of point (2) is simple. Everything of relevance to the theory of matrices that was discovered or adumbrated in the work of SYLVESTER, CAYLEY and SMITH was rediscovered and more profoundly and coherently developed by the mathematicians of the Berlin school (WEIERSTRASS, KRONECKER, FROBENIUS) in the period 1868–1878. Furthermore it was through the work and influence of the Berlin school that the theory of matrices emerged at the turn of the century as a coherent, independent branch of mathematics, readily accessible to the mathematicians of the 20th century (Section 7).

One basic ingredient of the theory of matrices not even adumbrated by the English mathematicians is the canonical form (4.4) used by WEIERSTRASS in [1868]. Such a form was not compatible with the context of SYLVESTER's geometrical problem which permitted only quadratic forms. It would have been appropriate within the context of CAYLEY's proposed study of types of collineations but there is no evidence CAYLEY ever pursued these ideas that far. In this connection it is of interest to note that CAYLEY's problem of determining the matrices which commute with a given one led an aspiring young mathematician, CAMILLE JORDAN, to introduce the same canonical form used by WEIERSTRASS.

JORDAN was unaware of CAYLEY's work and introduced his canonical form in the course of his efforts to develop the profound but sketchy ideas on the solvability of equations that GALOIS had left to posterity.[20] The aspect of GALOIS' work which was the most fundamental and which has received the greatest attention by historians is his discovery that every polynomial equation has associated with it a group of permutations and that the solvability of the equation by radicals depends upon whether the associated group possesses (in modern terminology) a composition series with factor groups of prime order. But GALOIS also devoted considerable attention to the problem of using his criterion for solvability to determine when various types of equations are solvable by radicals. His approach involved representing the permutation group of the equation by what he called linear substitutions.

For example, GALOIS sought to determine necessary and sufficient conditions for an irreducible equation of prime degree p to be solvable. He showed that when the equation is solvable its permutation group G has the following form: Let $x_0, x_1, \ldots, x_{p-1}$ denote the p roots of the equation. Then the permutations of the x_i which constitute G will be of the form $x_k \to x_{k'}$, where $k' \equiv ak + b \pmod{p}$, $a \not\equiv 0$

[19] See Section 6 of my [1977].

[20] Nineteenth-century mathematicians developed GALOIS' work as it was known through [1846].

(mod p). Hence the permutations of G can be represented by linear substitutions $k' \equiv ak+b$. GALOIS also sought to characterize those primitive equations which are solvable [1846, 26–7, 51–61], and in this connection he made two important discoveries (asserted without proof): (1) the degrees of such solvable equations are of the form p^n, p prime; (2) the corresponding group G of permutations must have the following form analogous to that for a solvable equation of prime degree: If the p^n roots of the equation are written in the form $a_{k_1 \ldots k_n}$ then the permutations $a_{k_1 \ldots k_n} \to a_{k_1' \ldots k_n'}$ which constitute G are such that

$$k_i' \equiv m_{i1}k_1 + m_{i2}k_2 + \cdots + m_{in}k_n + a_i \pmod{p}, \qquad i = 1, 2, \ldots, n \tag{5.2}$$

where $|m_{ij}| \not\equiv 0 \pmod{p}$. GALOIS observed that all linear substitutions of type (5.2) form a group so that the problem was to determine all maximal solvable subgroups of this group.

GALOIS went no further with the general problem, but for $n=2$ he made some additional observations of interest. He pointed out that the "circular substitutions" $k_i' \equiv k_i + a_i$ form (in modern terminology) a normal subgroup. Thus he was led to consider the factor group

$$k_i' \equiv m_{i1}k_1 + m_{i2}k_2 \pmod{p}, \qquad i = 1, 2. \tag{5.3}$$

Since the group of substitutions (5.3) is not primitive, it did not figure prominently in GALOIS' analysis. Instead he focused on the factor group defined by

$$k' \equiv \frac{m_{11}k + m_{12}}{m_{21}k + m_{22}} \pmod{p} \tag{5.4}$$

i.e., the projective linear group.

During the period 1846–1866 some attempts were made to clarify GALOIS' ideas and establish his assertions, but significant progress was first made by JORDAN. Prior to JORDAN, GALOIS' successors, following his lead, concentrated upon projective linear substitutions (5.4). JORDAN, however, preferred to make the consideration of homogeneous linear substitutions (5.3) fundamental, and in [1867] he indicated their important role in the problem of determinating all the irreducible equations of a given degree which are solvable by radicals [1867: 132–133]. In connection with this problem he sought to determine the solvable subgroups of the group of linear substitutions in two variables (5.3) [1868]. Contrary to GALOIS' opinion, JORDAN showed that there exist three general types of solvable subgroups; to do it he utilized the fact that by a linear change of variables a linear substitution could be put in one of a limited number of "canonical forms" depending on the nature of the roots of $|S - kI| \equiv 0 \pmod{p}$ [1868: 111–113]. His method of constructing solvable subgroups was to build them up from their composition series, and this involved determining all linear substitutions which commute with a given substitution S. To this end, he introduced the possible canonical forms for S.

Two years later JORDAN published his *Traité des substitutions* [1870] which was largely a compendium and exposition of his own work during the 1860's. The study of linear homogeneous substitutions of the form (5.3) naturally occupied a prominent position in the *Traité* [1870, Ch. II]. Among other things, JORDAN generalized his canonical forms from 2 to n variables (or indices as he continued

to term them) [1870, 114ff]. That is, if S is a linear homogeneous substitution with defining equations $X' \equiv AX \pmod{p}$, he showed that a suitable linear change of variables $Y = PX$ can be determined so that the equations of S are $Y' \equiv JY \pmod{p}$, where J is the canonical form. The only difference between JORDAN's J and the familiar JORDAN canonical form is that the coefficients in the "JORDAN blocks" corresponding to the root k are written as

$$\begin{pmatrix} k & 0 & 0 & 0 & . & . & . \\ k & k & 0 & 0 & . & . & . \\ 0 & k & k & 0 & . & . & . \\ . & . & . & . & . & . & . \end{pmatrix},$$

where $k \not\equiv 0 \pmod{p}$ since $|S| \not\equiv 0 \pmod{p}$. Using his canonical form, JORDAN determined the linear substitutions which commute with a given substitution [1870, 128ff]. He also considered the problem of determining all substitutions S, T with the same canonical form and showed that they are related by $S = U^{-1}TU$ for some substitution U [1870, 136ff].

When JORDAN wrote the *Traité* he was unaware that WEIERSTRASS had introduced essentially the same canonical form for ordinary complex coefficients in [1868]. JORDAN was himself stimulated to introduce his canonical form within that context by the appearance of a note in the *Comptes rendus* by an engineer, YVON-VILLARCEAU [1870]. The object of the note was to call attention to the fact that LAGRANGE was mistaken when he claimed that multiple roots always produce unstable solutions in the associated linear differential equations.[21] WEIERSTRASS, of course, had clarified this matter in his [1858], but VILLARCEAU was unaware of it. His discovery was based upon a detailed analysis of $\ddot{Y} = AY$ in the case of two variables.

At the conclusion of his note VILLARCEAU raised the question of when a system $d^n Y/dt^n = AY$ can be resolved into separate equations each of which can be directly integrated. Since any such system can be reduced to $\dot{Y} = BY$, VILLARCEAU studied this case for 2 variables and announced that "it was easy to establish that, if, when the characteristic equation has equal roots, these equations can nevertheless be integrated by means of exponential and trigonometric functions unaffected by algebraic factors containing the independent variable t, the proposed system resolves into two equations which can be integrated separately" [1870, 766]. VILLARCEAU had raised some interesting questions, although he evidently did not possess the algebraic tools necessary to deal with the case of n variables. JORDAN responded with a note [1871] in which he pointed out that the question raised by VILLARCEAU can be easily resolved in general "by a procedure identical to that which we have used in our Traité des substitutions to reduce a linear substitution to its canonical form" [1871: 313].

Thus, JORDAN showed that $\dot{X} = AX$ could be transformed by a linear change of variables into $\dot{Y} = JY$, in which form the equations separate into subsystems

$$\frac{dy_1}{dt} = \sigma y_1, \frac{dy_2}{dt} = \sigma y_2 + y_1, \ \ldots \ \frac{dy_r}{dt} = \sigma y_r + y_{r-1}$$

[21] This fact was also realized by J. SOMOF [1859] and E. J. ROUTH [1877, Ch. I].

that can be integrated directly to yield $y_r=\psi(t)e^{\sigma t}$, $y_{r-1}=\psi'(t)e^{\sigma t}$, ... $y_1=\psi^{(r-1)}(t)e^{\sigma t}$, where $\psi(t)$ is any polynomial of degree $r-1$.[22] Hence VILLARCEAU's conclusion for 2 variables was generally true: the solutions involve no algebraic factors $\psi(t)$, $\psi'(t)$, ... if and only if the system reduces to equations $dy_i/dt=\sigma_i y_i$, $i=1, 2, \ldots, n$. JORDAN included this method of solution in his *Cours d'analyse* [1887, 173–175].

6. Frobenius

By 1874 the connection between WEIERSTRASS' theory of elementary divisors and JORDAN's work [1870, 1871] was recognized in Paris and in Berlin. Writing from Berlin, MEYER HAMBURGER pointed out [1873] that JORDAN's reduction procedure for systems of linear differential equations [1871] could be used to advantage in L. FUCHS' theory [1866] of the differential equation

$$\frac{d^n y}{dz^n}+p_1\frac{d^{n-1}y}{dz^{n-1}}+\cdots+p_{n-1}\frac{dy}{dz}+p_n=0 \tag{6.1}$$

with variable coefficients $p_i=p_i(z)$. If $y_1(z)$, $y_2(z)$, ..., $y_n(z)$ is a fundamental system of solutions to (6.1) and if z_0 is a singular point, then when z makes a complete revolution about z_0 the value of $y_i(z)$ becomes $\bar{y}_i(z)$, where $\bar{y}_i=\sum_{j=1}^{n}a_{ij}y_j$ and $|a_{ij}|\neq 0$. FUCHS showed that if w_1, w_2, ... denote the distinct roots of the "fundamental equation" $|a_{ij}-w\delta_{ij}|=0$ and if μ, μ', ... denote their multiplicities, then linear combinations $u_1, \ldots, u_\mu, v_1, \ldots, v_{\mu'}, \ldots$ of $y_1, y_2, \ldots, y_n$ may be chosen to form a new fundamental system with the property that

$$\begin{aligned}
&\bar{u}_1=w_1u_1\\
&\bar{u}_2=b_{21}u_1+w_1u_2\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\bar{u}_\mu=b_{\mu 1}u_1+b_{\mu 2}u_2+\cdots+w_1u_\mu
\end{aligned} \tag{6.2}$$

and similarly for $v_1, v_2, \ldots, v_{\mu'}$, etc.

In [1873] HAMBURGER pointed out that FUCHS' form (6.2) can be further simplified by JORDAN's reduction procedure, thereby pushing the analysis of the solutions to (6.1) further. He also noted the connection with the WEIERSTRASSIAN theory of elementary divisors [1873, 124], although it is unclear whether he realized that JORDAN's canonical form is actually in WEIERSTRASS' paper [1868]—a point subsequently emphasized by KRONECKER [1874a, 369]. HAMBURGER also presented [1873, 117–121] a lucid exposition of JORDAN's reduction to canonical form, which involves reasoning of an elementary nature in comparison with that of WEIERSTRASS, who employed sophisticated theorems from the theory of determinants.

---

[22] The same result on the integration of $A\dot{X}=BX$ was communicated by WEIERSTRASS to the Berlin Academy in 1875, but apparently first published in his *Werke* [1903]. No reference is made to JORDAN's note [1871]. It is likely WEIERSTRASS was aware of it—see the beginning of Section 6—but that he had realized this application of his theory of elementary divisors long before 1871, especially in view of the role of differential equations in [1858].

Perhaps HAMBURGER's paper called JORDAN's attention to the work of WEIERSTRASS and KRONECKER on quadratic and bilinear forms.[23] In a note inserted in the *Comptes rendus* for 22 December 1873, JORDAN proposed to study the theory of forms because "the solutions given by the eminent geometers of Berlin are incomplete in that they leave aside certain exceptional cases which are nonetheless of interest. Furthermore their analysis is fairly difficult to follow, especially that of M. Weierstrass" [1873, 7]. The details of JORDAN's study were subsequently published in LIOUVILLE's journal [1874a]. His provocative remarks about the "eminent geometers of Berlin" evoked more of the same from the pen of KRONECKER, who also pointed out an error in JORDAN's treatment.[24]

One consequence of the ensuing polemic between KRONECKER and JORDAN was that the connection between the theory of elementary divisors and JORDAN's reduction of linear substitutions to canonical form was well-advertised. It is likely that JORDAN found WEIERSTRASS' paper [1868] difficult because of its use of the theory of determinants. JORDAN pointed out [1874a, 35–36, 40–42] that his method of reduction can be used to derive WEIERSTRASS' theory. As noted in Section 4, WEIERSTRASS' proof of his main theorem had hinged upon showing the family of forms $s\Phi - \Psi$, $|\Phi| \not\equiv 0$ is equivalent to $sI - J$ where J is in JORDAN canonical form. By JORDAN's theorem, applied to the linear substitution $\Psi\Phi^{-1}$, there exists a non-singular substitution U such that $U^{-1}(\Psi\Phi^{-1})U = J$. Thus $H(s\Phi - \Psi)K = sI - J$, where $H = U^{-1}$ and $K = \Phi^{-1}U$. JORDAN did not give the argument in the compact, symbolical form just indicated, but such notation was employed by him and his proof does indicate that the transformation of forms can be reduced to the composition of linear substitutions. This idea became the starting point of FROBENIUS' important memoir [1878] to be discussed below, although FROBENIUS seems to have got the idea from CAYLEY [1855a, 1855b] rather than JORDAN. In any case, JORDAN himself failed to develop it.

Thus by 1874 it was recognized that, starting from JORDAN's theorem, WEIERSTRASS' theory of elementary divisors could be derived and that conversely, WEIERSTRASS' theory immediately implied JORDAN's theorem (consider $s\Phi - \Psi$ with $\Phi = I$). Their work was therefore mathematically equivalent. But in terms of the history of what was to become the theory of matrices, WEIERSTRASS' work was far more important because it exerted considerably more influence on the developments that comprise that history. The extent of his influence and its causes will be discussed in the following section. For the present I wish to concentrate upon a single manifestation of his influence: The fundamental and influential contributions to the theory of matrices made by his student, GEORG FROBENIUS.

FROBENIUS enrolled at the University of Berlin the year before WEIERSTRASS' memoir [1868] appeared. After six semesters of instruction there, he received his doctorate *summa cum laude* with a dissertation done under WEIERSTRASS' direction. His family resided in Berlin and family matters caused him to decline a handsome offer to become a *Privatdocent* at the University of Freiburg.[25] Instead he remained

[23] In [1874c] JORDAN applied his canonical form in the course of a group-theoretic study of the solution to (6.1).

[24] KRONECKER [1874a, 406]. See also DIEUDONNÉ [1962, viii].

[25] This information is from WEIERSTRASS' glowing recommendation (1872) of FROBENIUS for a position of *ausserordentlicher Professor* at Berlin. This and many other interesting documents have been published by BIERMANN in [1973]. See pp. 190–191 of [1973] for WEIERSTRASS' recommendation.

in Berlin until 1875 where he held various teaching positions at the high school level and then, in 1874, became *ausserordentlicher Professor* at the University. In 1875 he left Berlin to become *ordentlicher Professor* at the Polytechnikum in Zürich (now the ETH), where he remained until he returned to Berlin as KRONECKER'S successor at the University.

FROBENIUS' early years within the circle of Berlin mathematicians (1867–1875) thus coincided with the period in which his teachers WEIERSTRASS and KRONECKER were actively engaged in the creation of the theory of quadratic and bilinear forms. His own research at that time was in the area of differential equations, but in his first publication from Zürich, "Über das Pfaffsche Problem" [1876], he was already applying the results of both WEIERSTRASS and KRONECKER on the transformation of forms. His interest in the theory of forms was also encouraged by his friend and colleague at the Polytechnikum, LUDWIG STICKELBERGER. STICKELBERGER was also WEIERSTRASS' student, and his doctoral dissertation [1874] dealt with a problem in the theory of elementary divisors. Matters related to that theory continued to dominate STICKELBERGER'S research interests during his tenure at the Polytechnikum (1874–1879), during which time FROBENIUS composed two memoirs [1878, 1879] of fundamental importance for the evolution of the theory.

The memoir [1876] on the PFAFFIAN problem exemplifies FROBENIUS' penchant for taking a problem and reworking and extending it in his own original and lucid manner so as to produce a polished, definitive "monograph" on the subject. Eight months after submitting [1876], he submitted another "monograph" to CRELLE'S journal, "Über lineare Substitutionen und bilineare Formen" [1878], in which he created the equivalent of the calculus of matrices. The objects of his calculus are bilinear forms $A=\sum a_{\alpha\beta}x_\alpha y_\beta$, $B=\sum b_{\alpha\beta}x_\alpha y_\beta$. The addition and scalar multiplication of such forms was already implicit in the work of WEIERSTRASS and KRONECKER where families $pA+qB$ and $sA-B$ are studied. To these operations FROBENIUS added that of multiplication: $AB=\sum\frac{\partial A}{\partial y_\gamma}\frac{\partial B}{\partial x_\gamma}$. The form AB is, of course, $C=\sum c_{\alpha\beta}x_\alpha y_\beta$ where $c_{\alpha\beta}=\sum_\gamma a_{\alpha\gamma}b_{\gamma\beta}$. The transformation of the form A by means of linear substitutions $x_\alpha=\sum_\beta p_{\alpha\beta}X_\beta$, $y_\alpha=\sum_\beta q_{\alpha\beta}Y_\beta$ yields the form $P'AQ$ if the above substitutions are identified with $P=\sum p_{\alpha\beta}x_\alpha y_\beta$ and $Q=\sum q_{\alpha\beta}x_\alpha y_\beta$. ($P'$ is FROBENIUS' notation for the transpose of P which I have denoted by P^t.)

The problems which prompted FROBENIUS to create his calculus of forms and the significance of his memoir [1878] for the development of the symbolical algebra of matrices have been discussed in my [1977]. Here I simply wish to call attention to some of my conclusions that relate to the theme of the present essay. In the first place, the problems that motivated FROBENIUS to create the calculus of forms attracted his attention because they were good "WEIERSTRASSIAN problems". That is, they had previously been treated on the generic level and FROBENIUS saw in them the same sort of challenge that inspired WEIERSTRASS' papers [1858] and [1868] as well as KRONECKER'S. (Recall the quotation at the end of Section 4.) Of course, FROBENIUS was greatly aided in his own efforts by the discoveries of his teachers. His memoir [1878] brilliantly demonstrated the value of combining their discoveries with a symbolical calculus. His success in developing

matrix algebra far beyond the rudimentary, error-prone generic level found in CAYLEY's memoir [1858] is a tribute to his training at Berlin as well as to his own mathematical talent.

Whereas FROBENIUS' memoir [1878] involved the application of theory of elementary divisors, his memoir [1879] presented a radically different approach to the entire theory. WEIERSTRASS' own approach raised some questions. For example, as WEIERSTRASS realized [1868, 33] his work showed that two non-singular families of forms rA_1+A_2, rB_1+B_2 are equivalent if and only if the corresponding polynomials $D_i(r)$, $D_i'(r)$ are identical (except for constant factor). (The $D_i(r)$ are defined in the Introduction.) Therefore in order to determine whether rA_1+A_2 and rB_1+B_2 are equivalent it is not necessary to know the elementary divisors of $D(r)=|rA_1+A_2|$ and $D'(r)=|rB_1+B_2|$. As FROBENIUS observed [1879, 483]: "By means of rational operations one can therefore determine whether given forms $[A=rA_1+A_2,\ B=rB_1+B_2]$ are equivalent or not.... By contrast, all proofs known to me for the theorem of Herr Weierstrass involve irrational operations; for they are based on the transformation of A into a reduced form, whose coefficients depend upon the roots of the equation $|A|=0$. Hence I long ago had proposed to myself the problem of finding a proof for that theorem in which only rational operations occur."

Another problem was raised by WEIERSTRASS' approach. His reduction to a canonical form rested upon a lemma and his proof of the lemma was based on a property of determinants that he subsequently discovered was insufficiently proved.[26] STICKELBERGER's dissertation [1874] dealt with this problem. Since a direct proof of the lemma as originally attempted by WEIERSTRASS was not forthcoming, he obtained the desired reduction by an indirect argument. Both FROBENIUS and KRONECKER were interested in the problem. KRONECKER believed he had a method for resolving it directly, which, on several occasions, he sought to defend in the face of the criticisms offered by STICKELBERGER and FROBENIUS.[27] The apparent difficulty of patching up WEIERSTRASS' argument along the lines originally intended finally prompted STICKELBERGER [1879] to derive WEIERSTRASS' main theorem by an ingenious rearrangement of the parts of the original proof, although at the expense of diminishing its formal appeal.

The lacuna in WEIERSTRASS' proof involved the aspect of his paper [1868] that appears most foreign to modern readers: the theory of determinants. WEIERSTRASS was simply following the lead of JACOBI. KRONECKER also relied upon properties of determinants, although his many extensive memoirs on singular families of forms involved other types of reasoning as well. But KRONECKER appears to have been dissatisfied with his methods. In November of 1874 he composed a work on forms that contained a significant extension of WEIERSTRASS' reduction procedure to the singular case, but he decided against publication at that time "since as I composed it I vividly sensed the deficiencies of the analytic-algebraic methods applied therewith and, as a consequence, the need for a purely arith-

---

[26] See pp. 321–324 of the original version of [1868]. In the edition for his *Werke*, WEIERSTRASS had that part omitted (§ 4) and referred to FROBENIUS' subsequent proof [1894].

[27] According to FROBENIUS [1894, 578]. FROBENIUS' direct algebraic proof of WEIERSTRASS' lemma employed a determinant identity discovered by KRONECKER in 1870.

metical treatment of the problem."[28] Early in his career KRONECKER's dual interests in the theory of numbers and the theory of equations caused him "to keep especially in mind the arithmetical side of algebra" and to stress the arithmetical viewpoint in algebra in his publications and lectures.[29] Thus by 1874 KRONECKER sensed the need for an arithmetical approach to the theory of forms, but it was FROBENIUS, probably aided by KRONECKER's viewpoint, who first succeeded in providing one in [1879].

FROBENIUS' success was due to two factors: his rediscovery of SMITH's Lemma (Section 5) and his realization that much of the arithmetic of ordinary integers carries over to polynomials $f(r)=\sum_{k=0}^{n} a_k r^k$ with coefficients a_k that are complex numbers, or elements of an algebraic number field, or integers taken modulo a prime number p. Such a realization was a by-product of the arithmetizing tendency reflected in the above quotation from KRONECKER, a tendency which was mostly inspired by the work of GAUSS, especially his introduction of GAUSSIAN integers to establish the law of biquadratic reciprocity [1832]. GAUSS' work revealed that many properties of ordinary integers, including the EUCLIDEAN algorithm and uniqueness of factorization into prime numbers, carry over to other quantities. As DIRICHLET emphasized, the significance of GAUSSIAN integers transcended their use in dealing with biquadratic residues; they opened a vast new domain to higher arithmetic in which "almost every property of real integers finds its analogue" [1841, 507].

GAUSS' work led mathematicians to study the arithmetic of other types of "integers" as well. Although the emphasis was upon the integers of algebraic number fields, the tendency to view objects arithmetically that were not originally regarded as within the province of arithmetic made it clear that polynomials in one or several variables possessed arithmetical properties such as a EUCLIDEAN algorithm (one variable) and uniqueness of factorization. For example, in [1857] DEDEKIND proposed to develop a "simple, coherent presentation" of the theory of higher congruences

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \equiv 0 \pmod{p}$$

"which adheres rigorously to the analogy with the elements of the theory of numbers." The analogy was so "sweeping" that, " with the exception of some investigations peculiar to our subject, only a change of words is required in the proofs of the theory of numbers" [1857, 40]. The objects of DEDEKIND's study are polynomials with integral coefficients taken modulo a prime p, and the first result he establishes is, of course, that these polynomials permit a EUCLIDEAN algorithm from which the basic arithmetical properties are then derived.

The arithmetical aspect of polynomials was especially important in KRONECKER's work. He was interested in the solvability of equations and arithmetical approach to polynomials is basic to his conception of a field *(Rationalitätsbereich)*

[28] KRONECKER [1890b, 141]. The results and methods in [1890b] were discovered in 1874 but not published until 1890 because of KRONECKER's dissatisfaction with the methods used.

[29] See KRONECKER [1882, 245].

and field extensions.[30] Furthermore, KRONECKER was a leader in the application of the new arithmetical ideas and results to problems regarded as algebraic, such as the solvability of equations.[31] FROBENIUS' arithmetical approach to the theory of elementary divisors in [1879] was thus in harmony with the work of several mathematicians of the period, including in particular his teacher KRONECKER.

FROBENIUS' dissatisfaction with the irrational element in previous renditions of WEIERSTRASS' theory and his appreciation of the arithmetical approach to algebraic problems would never have produced a positive result without his rediscovery of SMITH's Lemma. Ignorant of SMITH's work [1861],[32] he decided to investigate the arithmetical theory of linear equations and congruences. Most of [1879] and its sequel [1880] is, in fact, devoted to such matters. He naturally realized that the subject is related to the transformation of bilinear forms $\sum a_{\alpha\beta}x_\alpha y_\beta$ with integral coefficients by means of linear substitutions with integral coefficients $p_{\alpha\beta}$, $q_{\alpha\beta}$ satisfying $|p_{\alpha\beta}| = \pm 1$, $|q_{\alpha\beta}| = \pm 1$ (unimodular substitutions). Two forms are defined to be equivalent if they are transformable into one another by unimodular substitutions. In FROBENIUS' notation, $A = \sum a_{\alpha\beta}x_\alpha y_\beta$ is equivalent to B if $PAQ = B$ where $|P| = \pm 1$, $|Q| = \pm 1$.

Within this context he discovered a slightly more general version of SMITH's Lemma: If the $a_{\alpha\beta}$ associated with A has rank l, where $1 \leq \alpha \leq m$, $1 \leq \beta \leq n$, then A is equivalent to

$$F = f_1 x_1 y_1 + f_1 f_2 x_1 y_2 + \cdots + f_1 f_2 \ldots f_l x_l y_l$$

or, with $e_\lambda = f_1 f_2 \ldots f_\lambda$,

$$F = e_1 x_1 y_1 + e_2 x_2 y_2 + \cdots + e_l x_l y_l. \tag{6.3}$$

SMITH's Lemma covered the case $l = m$. (It is, of course, the case $l = m = n$ that applies to non-singular families of biliniear forms.) Let $d_i = e_1 e_2 \ldots e_i$; then d_i and $e_i = d_i/d_{i-1}$ are the analogues of $D_i(r)$ and $E_i(r)$ of the theory of elementary divisors. (See the Introduction.) Undoubtedly the discovery of the reduced form (6.3) suggested to FROBENIUS the possibility of using it in place of WEIERSTRASS' in order to obtain a rational derivation of WEIERSTRASS' Theorem for families of forms $A = rA' + A''$ with coefficients which are (first degree) polynomials in r and can be treated, by analogy, by arithmetical methods.

The derivation of WEIERSTRASS' Theorem from (6.3) is, however, not immediate. SMITH's Lemma implies that if A has coefficients $a_{\alpha\beta}$ that are polynomials in r with complex coefficients then $PAQ = F$ where P and Q have coefficients which are polynomials in r and $|P|$, $|Q|$ are non-zero complex numbers (the analogue of ± 1). It is therefore necessary to prove that P and Q can be replaced by forms P_0, Q_0 with complex coefficients when A is a polynomial matrix of first degree in r. FROBENIUS proved the existence of P_0, Q_0. His proof [1879, 539–541] is of interest because it makes extensive use of his calculus of forms—matrix algebra. In [1878] he had demonstrated the value of WEIERSTRASS' theory of elementary

[30] See KRONECKER [1882].

[31] See KRONECKER [1853, 1856] and KUMMER's remarks [1975, 583] about it.

[32] According to FROBENIUS [1879, 483n] he first became aware of SMITH's work after completion of his paper [1879].

divisors for resolving problems in matrix algebra whereas in [1879] he demonstrated the value of matrix algebra for a simple and elegant derivation of the theory itself.

The reduced form (6.3) corresponding to $A=rA'+A''$ is not generally of the first degree in r. FROBENIUS therefore considered whether first degree reduced forms exist with prescribed elementary divisors. The answer is, of course, affirmative when complex numbers are allowed, for then the JORDAN form $rI-J$ can be used. But FROBENIUS realized his arithmetical theory applied in many other contexts such as when the $a_{\alpha\beta}$ are polynomials with coefficients that are algebraic numbers in some algebraic number field. (DEDEKIND's work on algebraic numbers and number fields was well-known to FROBENIUS at this time.) Therefore: "If, more generally, $\phi(r)$, $\phi_1(r)$, ... are polynomials in r whose coefficients are algebraic numbers of a certain field, and which are irreducible in this field, the question arises whether a form of first degree exists whose coefficients belong to the same field and which possess the simple elementary divisors[33] $(\phi(r))^{\varepsilon}$, $(\phi_1(r))^{\varepsilon'}$, ..." [1879, 542]. He constructed such a form as a sum of forms, each with its own set of variables, ε of them corresponding to $\phi(r)$, ε' to $\phi_1(r)$, *etc.* Corresponding to $\phi(r)=r^{\alpha}+a_1r^{\alpha-1}+\cdots+a_{\alpha}$ are forms

$$r(x_1y_1+\cdots+\gamma_{\alpha}y_{\alpha})+y_1(a_1x_1+\cdots+a_{\alpha}x_{\alpha})-(x_1y_2+x_2y_3+\cdots+x_{\alpha-1}y_{\alpha})$$

with determinant

$$\begin{vmatrix} r+a_1 & -1 & 0 & 0\ldots0 \\ a_2 & r & -1 & 0\ldots0 \\ a_3 & 0 & r & -1\ldots0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{\alpha} & 0 & 0 & 0\ldots r \end{vmatrix} = \phi(r).$$

In this manner FROBENIUS introduced the now-familiar rational canonical form.

FROBENIUS concluded his memoir by observing that "the principles developed also remain applicable in the case where the elements $a_{\alpha\beta}$ of the system A are polynomials in r with integral coefficients and two such functions are not considered to be distinct when their corresponding coefficients are congruent relative to a prime number modulus p." Thus a rational canonical form for $A=rA'+A''$ exists and "if use is made of the complex numbers introduced by Galois," the canonical form of JORDAN's *Traité des substitutions* [1870] is obtained.

From the standpoint of modern algebra, the reduction to the normal form (6.3) is valid for any EUCLIDEAN ring—or even a principal ideal ring—, and since the polynomials over any field form a EUCLIDEAN ring, FROBENIUS' derivation of elementary divisor theory is valid for any field. As the above quotations suggest, FROBENIUS recognized the applicability of his method to any particular field, but in characteristically 19th-century fashion he preferred to remain on the level of analogies instead of introducing abstract structures — the one notable exception being the relatively simple case of finite groups. Nonetheless his arithmetical

[33] FROBENIUS used the term "simple elementary divisor" for what WEIERSTRASS referred to as an elementary divisor. See the Introduction (Section 1). He reserved the term "elementary divisor" for the $E_i(r)$.

derivation of WEIERSTRASS' theory became the paradigm for the development of the theory of matrices within the framework of modern algebra. Even before the advent of modern algebra, FROBENIUS' approach became the standard one and remained so through the mid-1930's.[34]

One of the reasons FROBENIUS' "arithmetization" of WEIERSTRASS' theory was appealing was that it avoided reliance upon the theory of determinants. This is somewhat ironic because FROBENIUS maintained a special fondness for that theory throughout his life. As for WEIERSTRASS, he apparently never fully appreciated FROBENIUS' achievements in [1878] and [1879]. For in his lectures at Berlin on the theory of quadratic and bilinear forms during the winter semester of 1886/87 (HEFFTER [1886/87]) he made no use of FROBENIUS' symbolical calculus nor of his arithmetical approach to the theory of elementary divisors.

Throughout the rest of his career FROBENIUS continued to make other important contributions to the theory and application of matrices, which are too extensive to relate here. One paper by him [1896] is worth mentioning, for it forms a fitting conclusion to this section because it relates to the work of both WEIERSTRASS and KRONECKER and to the problem posed by KRONECKER in [1866].

In [1866] KRONECKER had posed the question of when two bilinear forms are equivalent by means of cogredient transformation. In the notation of FROBENIUS, the problem is: given bilinear forms A, B when does R exist, $|R| \neq 0$, such that $R'AR = B$? (R' is the transpose of R.) KRONECKER had observed that $R'AR = B$ implies

$$R'(uA + vA')R = uB + vB'$$

which, of course, implies the two families are equivalent. Equivalence of these families is therefore a necessary condition for the congruence of A and B, but it is not at all clear that it is actually sufficient as well. This followed from the investigations of KRONECKER in 1874. WEIERSTRASS had taken the first step in [1868] by proving that non-singular families of quadratic forms are equivalent if and only if they are congruent, thereby relating his general theory to his earlier results [1858]. But the extension to the singular case and to families $uA + vA'$ was a consequence of KRONECKER's efforts.

---

[34] See *e.g.* MUTH [1899], BÔCHER [1907], CULLIS [1913–25], VAN DER WAERDEN [1931], SCHREIER & SPERNER [1932], TURNBULL & AIKEN [1932], MACDUFFEE [1933] and WEDDERBURN [1934]. It is now fashionable in algebra texts to derive what amounts to WEIERSTRASS' theory of elementary divisors (invariant factors) as an application of the fundamental theorem on finitely-generated modules over a principal ideal ring. Such an approach seems to go back to KRULL [1926] and involves an emphasis upon vector space notions. (See note 46.) But it also owes much to the arithmetic approach of FROBENIUS and his successors. In fact, VAN DER WAERDEN [1931, § 107] proves the above fundamental theorem as an application of the theory of elementary divisors which he develops [1931, § 106] along the lines indicated by FROBENIUS. (VAN DER WAERDEN recently [1975, 36] specified FROBENIUS as his source for his treatment of this part of linear algebra in his [1931].) That the theory of elementary divisors, especially as treated by FROBENIUS, was a factor in the development of the theory and application of modules is also suggested by KRULL's own background. His doctoral dissertation was *Über Begleitmatrizen und Elementartheilertheorie* [1922]. (Companion matrices [Begleitmatrizen] are the matrix blocks of FROBENIUS' rational canonical form.) The term seems to go back to KRULL's teacher at Freiburg, A. LOEWY, whose research was primarily involved with the theory and applications of matrices and who was strongly influenced by the Berlin school. (See LOEWY [1910] and the many references between 1897 and 1930 in WEDDERBURN's bibliography [1934].)

The problem of studying the singular case is a difficult one, and it is not surprising that KRONECKER's work is extremely complicated and difficult to read. As already noted, he was himself dissatisfied with his algebraic methods. In [1896] FROBENIUS showed that the above theorem (A and B are congruent if and only if $uA+vA'$ and $uB+vB'$ are equivalent) can be derived in a simple and elegant fashion by means of his symbolical calculus which also serves to reveal the "true basis" of the theorem. Finally, FROBENIUS also encouraged and assisted P. MUTH in the preparation of the latter's book, *Theorie und Anwendungen der Elementartheiler* [1899] and, in particular, provided MUTH with several simpler proofs, using matrix algebra, of theorems due to KRONECKER & STICKELBERGER.[35]

7. Weierstrass and the Rise of the Theory of Matrices

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 was absorbed into the mainstream of mathematics, it was a creation of the 19th century. During that fertile epoch in the history of mathematics, basic theorems on matrix similarity, congruence and equivalence were established and canonical matrix forms introduced. The properties of the characteristic roots of various types of matrices (symmetric, HERMITEAN, orthogonal, unitary, *etc.*) were also discovered.[36] In addition, the symbolical algebra of matrices was conceived by several mathematicians (EISENSTEIN (1852), CAYLEY (1858), LAGUERRE (1867), FROBENIUS (1878), SYLVESTER (1881)) and extensively developed by FROBENIUS, who was also responsible for the introduction of essentially modern proof techniques through the use he made of arithmetical notions and matrix algebra. Many applications of these results were also made — applications to problems in physics, algebra, the theory of numbers, geometry and differential equations.

At the close of the 19th century the first systematic and comprehensive account of the above developments was presented by P. MUTH in his treatise *Theorie und Anwendungen der Elementartheiler* [1899]. Then M. BÔCHER of Harvard made the theory and application of matrices a major component of his textbook on algebra [1907], which was soon translated into German [1910]. The books by MUTH and BÔCHER played a key role in the dissemination of the ideas and results of the theory by making them available in a more accessible form. They became the standard references on the subject during the early 20th century.[37]

As the century progressed, interest in the subject increased. It was recognized as a fundamental part of modern algebra, a recognition facilitated by the common arithmetical nature of modern algebra and the theory of matrices after FROBENIUS.

---

[35] See *e.g.* pp. 135–142 of MUTH [1899].

[36] BRIOSCHI [1854] showed the characteristic roots of an orthogonal substitution are of the form $e^{i\theta}$. In [1878] FROBENIUS reproved this result and also showed the elementary divisors are all linear (so that the JORDAN canonical form is diagonal). In [1883] FROBENIUS observed that the proofs and results carry over to what are now called unitary matrices. The theory of HERMITEAN and unitary matrices was extensively studied by AUTONNE [1902], whose work was strongly influenced by FROBENIUS' paper [1878], especially with regard to notation.

[37] Other early 20th century syntheses were LOEWY [1910] and CULLIS [1913/25].

Interest was also heightened by the advent of matrix mechanics.[38] The result was a spate of treatises devoted to the theory of matrices in the early 1930's, such as TURNBULL & AIKEN [1932], SCHREIER & SPERNER [1932], MACDUFFEE [1933] and WEDDERBURN [1934]. Although most of these books reflected the abstract viewpoint of modern algebra and some, such as SCHREIER & SPERNER [1932] and WEDDERBURN [1934], made the concept of a vector space fundamental,[39] all of them clearly owed a great deal to the earlier books by MUTH and BÔCHER. The influence of FROBENIUS is also evident in all of them, including the books by MUTH and BÔCHER; for they all employ what is essentially FROBENIUS' arithmetical approach to the theory of elementary divisors. And that theory is the centerpiece of all of them. Indeed in the foreword to *Vorlesungen über Matrizen* [1932] SPERNER explained that "This little volume contains the lectures of Otto Schreier on the theory of elementary divisors." By the early 1930's the theory of matrices had thus been completely assimilated into the mainstream of mathematics. In one form or another, however, the basic results of the theory as presented in the 1930's had been discovered in the course of the 19th century.

It has been customary to regard CAYLEY as the founder of the theory whose progress has just been sketched. I have suggested [1977] that such a characterization of CAYLEY's work is historically misleading even though CAYLEY did play a special role, through his work prior to 1858, in the introduction of one aspect of the theory (matrix algebra). There is an obvious danger in bestowing such a title upon anyone since it fosters a simplistic interpretation of history. With that caveat in mind, I would suggest that, insofar as anyone deserves the title of founder of the theory of matrices, it is WEIERSTRASS. As I have already tried to indicate in the preceding pages, he is the central figure in the developments occuring in the 19th century. His theory of elementary divisors provided a theoretical core, a substantial foundation, upon which to build. His work demonstrated the possibility of dealing by the methods of analysis with the non-generic case, thereby opening up a whole new world to mathematical investigation, a world that his colleagues and students proceeded to explore. JORDAN's work revealed the same world of possiblities and, in general, he approached mathematics on the same critical level

---

[38] HEISENBERG [1925], BORN-JORDAN [1925], and BORN-HEISENBERG-JORDAN [1926]. For English translations of these papers and commentary, see VAN DER WAERDEN [1968]. For the further history of matrix mechanics consult also HUND [1974], JAMMER [1966] and REICHE [1922].

[39] The concept of a vector space is basic to GRASSMAN's *Ausdehnungslehre* [1844] and in the second version of this work [1862] he formulated results about quadratic forms and linear substitutions within the framework of his *Ausdehnungslehre* [1862, n: 377-391]. For example, if $f(\rho)=(\rho-\alpha)^a(\alpha-\beta)^b\ldots$ is the factorization of the characteristic equation of a linear transformation T—GRASSMAN has a different name for T—then GRASSMAN proves (in effect) that the n-dimensional vector space V upon which T acts has a basis such that $V=V_\alpha\oplus V_\beta\oplus \ldots$ where the matrix of equations of T on each of V_α, V_β, ... is triangular. This result may have been generally known at the time. It is employed by FUCHS in [1866] (see Section 6) and according to LIE [1888, 585], CAUCHY employed a reduction to triangular form in his work on differential equations. (I have not been able to determine the work of CAUCHY's to which LIE is referring.) The concept of a vector space was only gradually adopted as basic and not generally until the 20th century, especially after the geometrical formulation of HILBERT's theory of integral equations by E. SCHMIDT and the importance of covariance, and hence abstract vector spaces, in the general theory of relativity as presented by WEYL [1918]. WEYL developed [1923, 88ff] an approach to the theory of elementary divisors essentially similar to KRULL's (note 41).

as WEIERSTRASS. But in terms of influence, WEIERSTRASS' work was far more significant.

WEIERSTRASS' memoirs inspired KRONECKER's extensive study of singular families of forms and, in fact, several papers on quadratic and bilinear forms by JORDAN himself.[40] Although WEIERSTRASS made no direct contribution to the development of matrix algebra, his student FROBENIUS, very much under his influence, rediscovered the contents of CAYLEY's memoir [1858] (insofar as they were correct) and pushed the theoretical development of matrix algebra considerably further by uniting it with the results of WEIERSTRASS and KRONECKER and their critical, non-generic approach to mathematics. Motivated by questions raised by WEIERSTRASS' treatment of the theory of elementary divisors, FROBENIUS also introduced the arithmetical approach to the theory which, as noted, facilitated the incorporation of the theory into the modern algebra of the 1920's.

Most of the applications of the theory of matrices that were discovered in the 19th century were also discovered as applications of WEIERSTRASS' theory of elementary divisors. The earliest geometrical application was in a sense made already by SYLVESTER in [1851]; but, as noted in Section 5, SYLVESTER's work on contact of quadric surfaces was an application of a non-existent theory. His work was also unknown, even by his fellow countryman H. J. S. SMITH. Independently of SYLVESTER, the geometry of families of quadric surfaces was thoroughly studied as an application of the results in WEIERSTRASS' memoir [1868] by his student W. KILLING [1872]. WEIERSTRASS' theory and its application to quadric surfaces also became widely known because S. GUNDELFINGER included both in the third edition of HESSE's popular *Vorlesungen über analytische Geometrie des Raumes* [1876]. Other geometrical applications of WEIERSTRASS' theory included F. KLEIN's inaugural dissertation [1868] on PLÜCKER-type geometry and the work of C.SEGRE [1884] on the classification of collineations, which involved introduction of the SEGRE characteristics. In the field of algebra, KILLING made crucial use of WEIERSTRASS' work in his ground-breaking investigation of the structure of LIE algebras [1888–1890]. Elementary divisors and the calculus of forms were also basic tools employed by FROBENIUS in his creation of the theory of group characters and representations (1896–1898).[41] Even in the theory of differential equations, where JORDAN was the first to indicate the relevance of his canonical form, the influence of WEIERSTRASS is evident in subsequent applications.[42]

Finally, the dominant WEIERSTRASSIAN influence upon the two key books of MUTH [1899] and BÔCHER [1907, 1910] is unmistakable. MUTH's treatise was expressly written to expound WEIERSTRASS' work and its subsequent development and application; and BÔCHER, who had applied WEIERSTRASS' theory of elementary divisors in his doctoral dissertation at Göttingen [1891], regarded it as "one of the most perfect and useful of mathematical theories" [1908, 194]. Furthermore WEIERSTRASS' colleagues and students greatly influenced the form of these books. MUTH acknowledged the encouragement and assistance of "outstanding experts on the theory of elementary divisors" such as FROBENIUS and KURT HENSEL, a

[40] For references and a summary of JORDAN's work in modern terminology, see DIEUDONNÉ [1962, vii–xiv].

[41] For references, consult my [1974].

[42] See, *e.g.*, L. SAUVAGE [1895] and J. HORN [1890, 16, n. 1].

student of both WEIERSTRASS and KRONECKER and FROBENIUS' colleague at Berlin in the 1890's. And BÔCHER wrote in the Preface of [1907]:

> In a work of this kind, it has not seemed advisable to give many bibliographic references, nor would an acknowledgement at this point of the sources from which the material has been taken be feasible. The work of two mathematicians, however, Kronecker and Frobenius, has been of such decisive influence on the character of the book that it is fitting that their names receive special mention here.

The above considerations, taken in conjunction with the detailed analysis of the preceding sections, constitute, I believe, a cogent argument for the central position of WEIERSTRASS and his work in the history of the theory of matrices. The question thus arises as to why his work should have been so influential. The answer is implicit in what has already been written. On the one hand, WEIERSTRASS' theory appealed to the mathematicians of the 19th century by virtue of its general scope and its expression in the language of forms. By contrast JORDAN had simply discovered a theorem on the reduction of a linear substitution which could be used to developed WEIERSTRASS' theory (as noted in Section 6). On the other hand, WEIERSTRASS' influence was also a manifestation of the fact that during the second half of the 19th century Germany was the unquestioned leader in the field of mathematical research. In terms of both quantity and quality of research the German university system excelled, and the University of Berlin was the most famous and productive of all.[43] WEIERSTRASS had colleagues and students to develop his ideas, and mathematicians from all over Germany and throughout the learned world looked to the work of the Berlin mathematicians for guidance and inspiration.

It is interesting to consider in retrospect the motivating forces which combined to produce the theory of matrices. One that comes immediately to mind is a concern with physical problems. Such a concern was, of course, primary in the work of the 18th-century geometers, but it continued to inspire the creation of mathematics in the 19th century. WEIERSTRASS' memoir [1858], the prelude to his theory of elementary divisors, was motivated by such a concern, as were the works of CAUCHY [1840] and DIRICHLET [1846] which aided his resolution of the problems posed in [1858]. Physical problems also produced the purely mathematical papers on HERMITEAN symmetric forms by CLEBSCH [1863] and CHRISTOFFEL [1866]. In the second half of the 19th century, however, the motivation for the further development of the theory came largely from within mathematics itself which, of course, was undergoing tremendous expansion. Thus KRONECKER's key paper [1866] was inspired by a question raised by the theory of theta functions. The work of HERMITE [1855] and FROBENIUS [1879] was a by-product of arithmetical investigations, while FROBENIUS' paper [1878] was prompted by a problem related to a question in the theory of numbers.[44] JORDAN's canonical form, of course, resulted from a group-theoretic problem raised by GALOIS' work and the papers

[43] A good idea of the central position of the University of Berlin in the 19th century can be gained from BIERMANN's book [1973].

[44] See Section 5 of my [1977].

of SYLVESTER [1851] and CAYLEY [1855c] were concerned with geometrical questions.

One motivational force common to the entire 19th century was a concern for a more rigorous level of reasoning in mathematics. It is sometimes suggested that rigor in mathematics is tantamount to *rigor mortis*, that when a mathematical theory is rigorized it has ceased to be productive. This essay provides a counter example to such a generality. A concern for higher standards of reasoning was a driving force behind WEIERSTRASS' work and also behind that of CAUCHY and DIRICHLET which preceded it and behind that of KRONECKER and FROBENIUS which succeeded it. (Recall KRONECKER's statement at the end of Section 4.) The rise of the theory of matrices was directly related to the fall of the generic approach to algebraic analysis. A concern for rigor did not mark the end of the creative development of the theory but its beginning.

8. Bibliography

AUTONNE, L., 1902. "Sur l'Hermitien", *Rendiconti circolo mat. di Palermo* **16**, 104–127.

BALTZER, R., 1857. *Theorie und Anwendungen der Determinanten*, Leipzig. 2nd ed. Leipzig, 1864.

BIERMANN, KURT-R., 1973. *Die Mathematik und ihre Dozenten an der Berliner Universität*, Berlin.

BÔCHER, M., 1891. *Ueber die Reihenentwicklungen der Potentialtheorie*, Göttingen. Expanded into a book and published by Teubner of Leipzig, 1894.

BÔCHER, M., 1907. *Introduction to Higher Algebra*, New York. 2nd ed., New York, 1922. German translation by H. BECK, Leipzig, 1910.

BÔCHER, M., 1908. [Review of BROMWICH, *Quadratic Forms*...], *Bull. Amer. Math. Soc.*, **14**, 194–195.

BORCHARDT, C., 1846. "Neue Eigenschaft der Gleichung, mit deren Hälfe man die saeculären Störungen der Planeten bestimmt," *Jl. für reine u. angew. Math.*, **30**, 38–45.

BORN, M., & P. JORDAN, 1925. "Zur Quantenmechanik," *Zeitschr. für Physik*, **34**, 858–888.

BORN, M., W. HEISENBERG & P. JORDAN, 1926. "Zur Quantenmechanik II," *Zeitschr. für Physik*, **35**, 577–615.

BRIOSCHI, F., 1854. "Note sur un théorème relatif aux déterminants gauches," *Jl. de. Math. pures et appl.*, **19**, 253–256 = *Opere*, **5** (Milan, 1909), 161–164.

BRIOSCHI, F., 1856. *Theorie der Determinanten und ihre hauptsächlichen Anwendungen*, Berlin.

CAUCHY, A., 1815. "Mémoire sur les fonctions qui ne peuvent obtenir que denx valeurs égales et de signes contraires par suite des transpositions opérées entre les variables qu'elles renferment," *Jl. Ec. Poly., IX, cah.* **17** = *Œuvres* **(2) 1** (Paris, 1905), 91–169.

CAUCHY, A., 1821. *Cours d'analyse de l'Ecole Royale Polytechnique*, Paris = *Œuvres* **(2) 3** (Paris, 1897).

CAUCHY, A., 1826. "Application du calcul des résidues à l'intégration des équations différentielles linéares et à coefficients constants," *Exer. de. math.*, **1** = *Œuvres* **(2) 6** (Paris, 1887), 252–255.

CAUCHY, A., 1829. "Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes," *Exer. de math.*, **4** = *Œuvres* **(2) 9**, 174–195.

CAUCHY, A., 1839a. "Méthode générale propre à fournier les équations de condition relatives aux limites des corps dans les problèmes de physique mathématique," *Comptes rendus Acad. Sci. Paris* **8** = *Œuvres* **(1) 4** (Paris, 1884), 193–227.

CAUCHY, A., 1839b, 1840. "Mémoire sur l'intégration des équations linéaires," *Comptes rendus Acad. Sci. Paris*, **8** = *Œuvres* **(1) 4** (Paris, 1884), 369–426. Reprinted in *Exer. d'analyse* **1** (1840) = *Œuvres* **(2) 11** (Paris, 1913), 75–133.

CAUCHY, A., 1850. "Mémoire sur les perturbations produites dans les mouvements vibratoires d'un système de molécules par l'influence d'un autre système," *Comptes rendus Acad. Sci. Paris* **30** = *Œuvres* **(1) 11** (Paris, 1899), 202–211.

CAYLEY, A., 1855a. "Remarques sur la notation des fonctions algébriques," *Jl. für reine u. angew. Math.*, **50**, 282–285 = *Papers* **2** (Cambridge, 1889), 185–188.

CAYLEY, A., 1855b. "Sur la transformation d'une fonction quadratique en elle-même par des substitutions linéaires," *Jl. für reine u. angew. Math.* **50**, 288–299 = *Papers* **2** (Cambridge, 1889), 192–201.

CAYLEY, A., 1855c. "Recherches sur les matrices dont les terms sont des fonctions linéaires d'une seule indéterminée," *Jl. für reine u. angew. Math.*, **50**, 313–317 = *Papers* **2** (Cambridge, 1889), 216–220.

CAYLEY, A., 1858. "A Memoir on the Theory of Matrices," *Phil. Trans. R. Soc. London*, **148** = *Papers* **2** (Cambridge, 1889), 475–496.

CHRISTOFFEL, E. B., 1856. *De motu permanenti electricitatis in corporibus homogeneis*, Berlin = *Abhandlungen*, **1** (Leipzig & Berlin, 1910), 1–64.

CHRISTOFFEL, E. B., 1864a. "Verallgemeinerung einiger Theoreme des Herrn Weierstrass," *Jl. für reine u. angew. Math.*, **63**, 255–272 = *Abhandlungen* **1** (Leipzig & Berlin, 1910), 128–145.

CHRISTOFFEL, E.B., 1864b. "Ueber die kleinen Schwingungen eines periodisch eingerichteten Systems materieller Punkte," *Jl. für reine u. angew. Math.*, **63**, 273–288 = *Abhandlungen* **1** (Leipzig & Berlin, 1910), 146–161.

CLEBSCH, A., 1860. "Theorie der circularpolarisirenden Medien," *Jl. für reine u. angew. Math.*, **57**, 319–358.

CLEBSCH, A., 1863. "Ueber eine Classe von Gleichungen, welche nur reelle Wurzeln besitzen," *Jl. für reine u. angew. Math.*, **62**, 232–245.

CULLIS, C. E., 1913–25. *Matrices and Determinoids*, 3 vols., Cambridge.

DEDEKIND, R., 1857. "Abriss einer Theorie der höheren Kongruenzen in bezug auf einen reellen Primzahl-Modulus," *Jl. für reine u. angew. Math.*, **54** = *Werke* **1** (Braunschweig, 1930), 40–67.

DIEUDONNÉ, J., 1962. "Notes sur les travaux de C. Jordan relatifs à l'algèbre linéaire et multilinéaire et la théorie des nombres," *Œuvres de Camille Jordan*, **3** (Paris, 1962), V–XX.

DICKSON, L. E., 1919. *History of the Theory of Numbers*, 3 vols., Washington, D.C.

DIRICHLET, P. G. L., 1842. "Recherches sur les formes quadratiques à coefficients et à indéterminées complexes," *Jl. für reine u. angew. Math.*, **24** = *Werke*, **1** (Berlin, 1889), 533–618.

DIRICHLET, P. G. L., 1846. "Ueber die Stabilität des Gleichgewichts," *Jl. für reine u. angew. Math.*, **32** = *Werke*, **2** (Berlin, 1897), 3–8.

DUGAC, P., 1972. *Eléments d'analyse de Karl Weierstrass*, Paris (mimeograph, Université de Paris VI).

DUGAC, P., 1973. "Eléments d'analyse de Karl Weierstrass," *Archive Hist. Ex. Sci.*, **10**, 41–176.

FROBENIUS, G., 1876. "Ueber das Pfaffsche Problem," *Jl. für reine u. angew. Math.*, **82** = [1968, v. **1**, 249–334].

FROBENIUS, G., 1878. "Ueber lineare Substitutionen und bilineare Formen," *Jl. für reine u. angew. Math.*, **84** = [1968, v. **1**, 343–405].

FROBENIUS, G., 1879. "Theorie der linearen Formen mit ganzen Coefficienten," *Jl. für reine u. angew. Math.*, **86** = [1968, v. **1**, 482–544].

FROBENIUS, G., 1880. "Theorie der linearen Formen mit ganzen Coefficienten (Forts.)," *Jl. für reine u. angew. Math.*, **88** = [1968, v. **1**, 591–611].

FROBENIUS, G., 1883. "Ueber die principale Transformation der Thetafunctionen mehrer Variabeln," *Jl. für reine u. angew. Math.*, **95** = [1968, v. **2**, 97–129].

FROBENIUS, G., 1893. "Gedächtnisrede auf Leopold Kronecker", *Abh. Akad. Wiss. Berlin*, = [1968, v. **3**, 705–724].

FROBENIUS, G., 1894. "Ueber die Elementartheiler der Determinanten," *S'ber. Akad. Wiss. Berlin* = [1968, v. **2**, 577–590].

FROBENIUS, G., 1896. "Ueber die cogredienten Transformationen der bilinearen Formen," *S'ber. Akad. Wiss. Berlin* = [1968, v. **2**, 695–704].

FROBENIUS, G., 1968. *Gesammelte Abhandlungen*, 3 vols., Berlin, 1968.

FUCHS, L., 1866. "Zur Theorie der linearen Differentialgleichungen mit veränderlichen Coefficienten," *Jl. für reine u. angew. Math.* **66** = *Werke*, **1** (Berlin, 1904), 159–202.

GALOIS, E., 1846. "Œuvres mathématiques d'Évariste Galois," *Jl. de math. pures et appl.*, **11**, 381–444. Reprinted as a book, *Œuvres mathématiques d'Evariste Galois*, Paris, 1897.

GAUSS, C. F., 1832. "Theoria residuorum biquadraticorum commentatio secunda," *Comm. soc. sci. Gottingenesis*, **7** = *Werke* **2** (Göttingen, 1863), 93–149.

GRASSMANN, H., 1844. *Die lineale Ausdehnungslehre*, Leipzig = *Werke*, v. **1**, Pt. 1 (Leipzig, 1894).

GRASSMANN, H., 1862. *Die Ausdehnungslehre*, Berlin = *Werke*, v. **1**, Pt. 2 (Leipzig, 1896).

HAMBURGER, M., 1873. "Bemerkung über die Form der Integrale der linearen Differentialgleichungen mit veränderlichen Coefficienten," *Jl. für reine u. angew. Math.*, **76**, 113–125.

HAWKINS, T., 1974. "New Light on Frobenius' Creation of the Theory of Group Characters," *Archive Hist. Ex. Sci.*, **12**, 217–243.

HAWKINS, T., 1975a. "Cauchy and the Spectral Theory of Matrices," *Historia Math.* **2**, 1–29.

HAWKINS, T., 1975b. "The Theory of Matrices in the 19th Century," *Proceedings of the International Congress of Mathematicians, Vancouver, 1974*, v. **2**, Canadian Math. Congress, 1975, 561–570.

HAWKINS, T., 1977. "Another Look at Cayley and the Theory of Matrices," *Archives int'l. hist. sci.* (to appear).

HEFFTER, L., 1886/87. *Theorie der bilinearen und quadratischen Formen von Prof. Weierstrass (Berlin, W. S. 1886/87), ausgearbeitet von Dr. L. Heffter*, Unpublished MS, Institut Mittag-Leffler, Djursholm, Sweden.

HEISENBERG, W., 1925. "Ueber quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen," *Zeitschr. für Physik*, **33**, 879–893.

HERMITE, C., 1853. "Sur la décomposition d'un nombre en quatre carrés," *Comptes rendus Acad. Sci. Paris*, **37** = *Œuvres*, **1** (Paris, 1905), 288–289.

HERMITE, C., 1854a. "Sur la théorie des formes quadratiques. Premier mémoire," *Jl. für reine u. angew. Math.* **47** = *Œuvres*, **1** (Paris, 1905), 200–233.

HERMITE, C., 1854b. "Sur la théorie des formes quadratiques. Seconde mémoire," *Jl. für reine u. angew. Math.*, **47** = *Œuvres*, **1** (Paris, 1905), 234–263.

HERMITE, C., 1855. "Remarque sur un théorème de M. Cauchy," *Comptes rendus Acad. Sci. Paris*, **41** = *Œuvres*, **1** (Paris, 1905), 479–481.

HERMITE, C., 1856. "Sur la nombre des racines d'une équation algébrique comprises entre des limites données," *Jl. für reine u. angew. Math.*, **52** = *Œuvres*, **1** (Paris, 1905), 397–414.

HERMITE, C., 1912. "Sur l'extension du théorème de M. Sturm à un système d'équations simultanées [Presented to Paris Academy on 12 July 1852], *Œuvres*, **3**, Paris, 1–34.

HESSE, O., 1876. *Vorlesungen über analytische Geometrie des Raumes, insbesondere über Oberfläche zweiter Ordnung*. Revidiert und mit Zusätzen versehen von Dr. S. Gundelfinger... 3rd edn., Leipzig.

HORN, J., 1890. *Ueber Systeme lineare Differentialgleichungen mit mehreren Veränderlichen. Beiträge zur Verallgemeinerung der Fuchs'schen Theorie der linearen Differentialgleichungen*, Berlin.

HUND, F., 1974. *The History of Quantum Theory*, London.

JACOBI, C. G. J., 1834. "De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis quae solis quadratis variabilium constant...," *Jl. für reine u. angew. Math.*, **12** = *Werke*, **3** (Berlin, 1884), 191–268.

JACOBI, C. G. J., 1857. "Ueber eine elementare Transformation eines in bezug auf jedes von zwei Variablen-Systemen linearen und homogenen Ausdrucks," *Jl. für reine u. angew. Math.*, **53** = *Werke*, **3** (Berlin, 1884), 583–590.

JAMMER, M., 1966. *The Conceptual Development of Quantum Mechanics*, New York.

JORDAN, C., 1867. "Mémoire sur la résolution algébrique des équations," *Jl. de math. pures et appl.*, (**2**) **12**, 109–157.

JORDAN, C., 1868. "Sur la résolution algébrique des équations primitives de degré p^2 (p étant premier impair), *Jl. de math. pures et appl.*, (**2**) **13**, 111–135.

JORDAN, C., 1870. *Traité des substitutions et des équations algébriques*, Paris.

JORDAN, C., 1871. "Sur la résolution des équations différentielles linéaires," *Comptes rendus Acad. Sci. Paris*, **73**, 787–791.

JORDAN, C., 1873. "Sur les polynômes bilinéaires," *Comptes rendus Acad. Sci. Paris*, **77**, 952–954.

JORDAN, C., 1874a. "Mémoire sur les formes bilinéaires," *Jl. de math. pures et appl.*, (**2**) **19**, 35–54.

JORDAN, C., 1874b. "Sur la réduction des formes bilinéaires," *Comptes rendus Acad. Sci. Paris*, **78**, 614–617.

JORDAN, C., 1874c. "Mémoire sur une application de la théorie des substitutions à l'etude des équations différentielles linéaires," *Bull. Soc. Math. France*, **2**, 100–127.

JORDAN, C., 1874d. "Mémoire sur la réduction et la transformation des systèmes quadratiques," *Jl. de math. pures et. appl.*, (**2**) **19**, 397–422.

JORDAN, C., 1874e. "Sur les systèmes de formes quadratiques," *Comptes rendus Acad. Sci. Paris*, **78**, 1763–1767.

JORDAN, C., 1887. *Cours d'analyse de l'Ecole Polytechnique*, v. 3, Paris. 2nd ed., Paris, 1896.

KILLING, W., 1872. *Der Flächenbüschel zweiter Ordnung*, Berlin.

KILLING, W., 1888–90. "Die Zusammensetzung der stetigen, endlichen Transformationsgruppen," *Math. Ann.*, **31**, 252–290, **33**, 1–48, **34**, 57–122, **36**, 161–189.

KLEIN, F., 1868. *Ueber die Transformation der allgemeinen Gleichung des zweiten Grades zwischen Linien Coordinaten auf eine canonische Form,* Bonn. Reprinted with comments and corrections in *Math. Ann.*, **23** (1884), 539–578 = *Abhandlungen* **1** (Berlin, 1921), 5–49.

KRONECKER, L., 1853. "Ueber die algebraisch auflösbaren Gleichungen," *M'ber. Akad. Wiss. Berlin* = *Werke* **4** (Leipzig & Berlin, 1929), 1–11.

KRONECKER, L., 1856. "Ueber die algebraisch auflösbaren Gleichungen," *M'ber. Akad. Wiss. Berlin* = *Werke* **4** (Leipzig & Berlin, 1929), 25–37.

KRONECKER, L., 1861. "Antrittsrede von L. Kronecker bei der Aufnahme in die Akademie der Wissenschaften am 4 Juli 1861...," *M'ber. Akad. Wiss. Berlin* = *Werke* **5** (Leipzig & Berlin, 1930), 385–393.

KRONECKER, L., 1866. "Ueber bilineare Formen," *M'ber. Akad. Wiss. Berlin* = *Werke,* **1** (Leipzig, 1895), 145–162.

KRONECKER, L., 1868. "Ueber Schaaren quadratischer Formen," *M'ber. Akad. Wiss. Berlin* = *Werke,* **1** (Leipzig, 1895), 163–174.

KRONECKER, L., 1874a. "Ueber Schaaren von quadratischen und bilinearen Formen," *M'ber. Akad. Wiss. Berlin* = *Werke,* **1** (Leipzig, 1895), 349–413.

KRONECKER, L., 1874b. "Sur les faisceaux de formes quadratiques et bilinéaires," *Comptes rendus Acad. Sci. Paris,* **78** = *Werke* **1** (Leipzig 1895), 415–419.

KRONECKER, L., 1874c. "Ueber die congruententen Transformationen der bilinearen Formen," *M'ber. Akad. Wiss. Berlin* = *Werke* **1** (Leipzig 1895), 421 – 483.

KRONECKER, L., 1882. *Grundzüge einer arithmetischen Theorie der allgebraischen Grössen,* Berlin = *Werke,* **3** (Leipzig 1897), 239–401.

KRONECKER, L., 1890a. "Reduction der Systeme von n^2 ganzzahligen Elementen," *Jl. für reine u. angew. Math.,* **107** = *Werke* **32** (Leipzig & Berlin, 1931), 123–126.

KRONECKER, L., 1890b. "Algebraische Reduction der Schaaren bilinearen Formen," *S'ber. Akad. Wiss. Berlin* = *Werke,* **32** (Leipzig & Berlin, 1931), 139–155.

KRULL, W., 1922. *Ueber Begleitmatrizen und Elementartheilertheorie,* Freiburg, 1922.

KRULL, W., 1926. "Theorie und Anwendung der verallgemeinten Abelschen Gruppen," *S'ber. Akad. Wiss. Heidelberg* (5 Feb., 1926).

KUMMER, E. E., 1975. *Collected Papers,* **1** (A. WEIL, ed.) Berlin, Heidelberg, New York.

LAGRANGE, J. L., 1766. "Solution de différents problèmes de calcul intégral...," *Misc. Taurinensia* **3** = *Œuvres* **1** (Paris, 1867), 471–668.

LAGRANGE, J. L., 1772–73. "Réflexions sur la résolution algébrique des équations," *Nouv. Mém. Acad. Sci. Berlin* = *Œuvres* **3** (Paris, 1869), 205–421.

LAGRANGE, J. L., 1788. *Méchanique analitique,* v. **1**, Paris.

LAGRANGE, J. L., 1853. *Mécanique analytique,* v. **1**, Paris. Reprinted as 4th edn. in *Œuvres,* **11** (Paris, 1888).

LAPLACE, P. S., 1799. *Traité de mécanique céleste,* v. **1**, Paris = *Œuvres* **1** (Paris, 1878).

LAPLACE, P. S., 1835. *Exposition du système du monde* 6th ed., Paris = *Œuvres,* **6** (Paris, 1884).

LEBESQUE, V. A., 1837. "Thèses de mécanique et d'astronomie," *Jl. de math. pures et appl.,* **2**, 337 – 355.

LIE, S., 1888. *Theorie der Transformationsgruppen. Erster Abschnitt,* Leipzig.

LOEWY, A., 1910. "Kombinatorik, Determinanten und Matrices," Ch. 2 of *Repertorium der höheren Analysis,* v. **1**, Part 1, P. EPSTEIN, ed., Leipzig & Berlin.

MACDUFFEE, C. C., 1933. *The Theory of Matrices,* Berlin.

MUTH, P., 1899. *Theorie und Anwendung der Elementartheiler,* Leipzig.

NOETHER, M., 1898. "James Joseph Sylvester," *Math. Ann.,* **50**, 133–156.

PLÜCKER, J., 1828. *Analytisch-Geometrische Entwicklungen. Erster Band,* Essen.

REICHE, F., 1922. *The Quantum Theory,* London.

ROUTH, E. J., 1877. *A Treatise on the Stability of a Given State of Motion, Particularly Steady Motion. Being the Essay to which the Adams Prize was Adjudged in 1877, in the University of Cambridge.* London. Reprinted with introduction by A. T. FULLER (New York, 1975).

SAUVAGE, L., 1895. *Théorie générale des systèmes d'équations différentielles linéaires et homogènes,* Paris.

SCHREIER, O., & E. SPERNER, 1932. *Vorlesungen über Matrizen,* Leipzig.

SEGRE, C., 1884. "Sulla teoria e sulla classificazione delle omografie in uno spazio lineare ad un numero qualunque di dimensioni," *Mem. Acc. Lincei Roma,* **(3) 19**, 127–148.

SMITH, H. J. S., 1861a. "On Systems of Linear Indeterminate Equations and Congruences," *Phil. Trans. R. Soc. London,* **151** = *Papers,* **1** (Oxford, 1894), 367–409.

SMITH, H. J. S., 1861b. "Report on the Theory of Numbers" [Part III], Report of the B. A. A. S., 1861 = *Papers* **1** (Oxford, 1894), 163–228. Report reprinted by Chelsea (New York, 1965).

SMITH, H. J. S., 1873. "Arithmetical Notes," *Proc. London Math. Soc.*, **4** = *Papers* **2** (Oxford, 1894), 67–85.

SOMOF, J., 1859. "Sur l'équation algébrique à l'aide de laquelle on determiné les oscillations très-petites d'un système de points matériels," *Mém. Acad. Sci. St. Pétersbourg*, (7) **1**, no. 14.

STICKELBERGER, L., 1874. *De problemate quodam ad duarum bilinearium vel quadraticarium transformationum pertinente*, Berlin.

STURM, C., 1829. "Extrait d'un mémoire sur l'intégration d'un système d'équations différentielles linéaires, présenté à l'Academie des Sciences le 27 Juillet 1829 par M. Sturm," *Bull. Sci. Math.*, **12**, 313–322.

SYLVESTER, J. J., 1850. "On the Intersections, Contacts, and Other Correlations of two Conics expressed by Indeterminante Coordinates," *Cambr. and Dublin Math. Jl.*, **5** = *Papers*, **1** (Cambridge, 1904), 119–137.

SYLVESTER, J. J., 1851a. "An Enumeration of the Contacts of Lines and Surfaces of the Second Order," *Phil. Mag.* (**4**) **1** = *Papers*, **1** (Cambridge, 1904), 219–240.

SYLVESTER, J. J., 1851b. "On the Relation between the Minor Determinants of Linearly Equivalent Quadratic Functions," *Phil. Mag.*, (**4**) **1** = *Papers*, **1** (Cambridge, 1904).

TURNBULL, H. W., & A. C. AITKEN, 1932. *An Introduction to the Theory of Canonical Matrices*, London & Glasgow.

VIÈTE, F., 1591. *In artem analyticen isagoge*..., Tours. References are to the English translation in J. KLEIN, *Greek Mathematical Thought and the Origin of Algebra*, Cambridge, Mass., 1968.

VAN DER WAERDEN, B. L., 1931. *Moderne Algebra*, v. **2**, Berlin.

VAN DER WAERDEN, B. L., 1968. *Sources of Quantum Mechanics*, Amsterdam, 1967. Reprint New York, 1968.

VAN DER WAERDEN, B. L., 1975. "On the Sources of My Book *Moderne Algebra*," *Historia Math.*, **2**, 31–40.

WEDDERBURN, J. H. M., 1934. *Lectures on Matrices*, New York.

WEIERSTRASS, K., 1858. "Ueber ein die homogene Functionen zweiten Grades betreffendes Theorem," *M'ber. Akad. Wiss. Berlin* = *Werke* **1** (Berlin, 1894), 233–246.

WEIERSTRASS, K., 1868. "Zur Theorie der quadratischen und bilinearen Formen," *M'ber. Akad. Wiss. Berlin*, 311–338 = *Werke* **2** (Berlin, 1895), 19–44.

WEIERSTRASS, K., 1879. "Nachtrag zu der am 4. März ... gelesenen Abhandlung: Ueber ein die homogenen Functionen zweiten Grades betreffendes Theorem," *S'ber. Akad. Wiss. Berlin* = *Werke* **3** (Berlin, 1903), 139–148.

WEIERSTRASS, K., 1895. "Bemerkungen zur Integration eines Systems linearer Differentialgleichungen mit constanten Coefficienten," [Communicated to Berlin Academy of Science on 28 October, 1875] *Werke* **2** (Berlin, 1895), 75–76.

WEYL, H., 1918. *Raum, Zeit, Materie*, Berlin.

WEYL, H., 1923. *Mathematische Analyse des Raumproblems*, Berlin.

WEYR, EDUARD, 1885a. "Sur la théorie des matrices," *Comptes rendus Acad. Sci. Paris*, **100**, 787–789.

WEYR, EDUARD, 1885b. "Repartition des matrices en espèces et formation de toutes les espèces," *Comptes rendus Acad. Sci. Paris*, **100**, 966–969.

WEYR, EDUARD, 1889. *O theorii forem bilinearnych*, Prague.

WEYR, EDUARD, 1890. "Zur Theorie der bilinearen Formen," *Monatsheft für Math. u. Phys.*, **1**, 163–236.

WHITTAKER, E., 1951. *A History of the Theories of Aether and Electricity*, Vol. 1, 2nd ed., London.

YVON-VILLARCEAU, A., 1870. "Note sur les conditions des petites oscillations d'un corps solide de figure quelconque et la théorie des équations différentielles linéaires," *Comptes rendus Acad. Sci. Paris*, **71**, 762–766.

Department of Mathematics
Boston University

(Received February 7, 1977)