Arch. Hist. Exact Sci. (2008) 62:23–57
DOI 10.1007/s00407-007-0006-6

Frobenius and the symbolical algebra of matrices

Thomas Hawkins

Received: 10 November 2006 / Published online: 22 May 2007

Contents

Communicated by J.J. Gray.

T. Hawkins (✉)
Department of Mathematics and Statistics, Boston University,
111 Cummington Street, Boston, MA 02215, USA
e-mail: twh@math.bu.edu

1 Introduction

Although bits and pieces of what would now be classified under the rubric of linear algebra can be traced back to ancient times, much of it originated in the late eighteenth and, especially, during the nineteenth century [21,23]. The manner in which it was practiced, i.e., the tools used in dealing with problems of a linear algebraic nature, was, however, quite different than at present. The primary tools were (1) ordinary Cartesian algebra as it applies, e.g., to the manipulation of systems of equations and (2) relations based upon determinant-theoretic constructs, which were developed extensively by Cauchy and Jacobi during the first half of the nineteenth century. Not only was the concept and viewpoint of a vector space, together with attendant notions, generally missing but even the tool of matrix algebra is rarely found. The absence of any widespread use of matrix algebra during the nineteenth and early twentieth centuries is particularly curious since, as I pointed out many years ago [22], the idea of a symbolical algebra of matrices was introduced independently by three mathematicians during the mid-nineteenth century: by Arthur Cayley in England in 1858 [2], by Edmond Laguerre in France in 1867 [45], and by Georg Frobenius in Germany in 1877 [8]. By the idea of a symbolical algebra of matrices I mean the now commonplace idea that coefficient arrays associated to such staples of nineteenth century mathematics as systems of linear equations, linear substitutions (i.e., transformations), and bilinear forms, can be conceived of as mathematical objects in their own right, denoted by a single symbol and added and multiplied so as to form what is now called a linear associative algebra.

After their initial explorations of the idea of symbolical algebra, neither Cayley nor Laguerre found any significant use for matrix algebra in their subsequent work. The same is not true of Frobenius. As I have already pointed out, unlike Cayley and Laguerre, who practiced linear algebra on the generic level (as was usual at the time), and developed matrix algebra accordingly, Frobenius fused matrix algebra with the nongeneric linear algebra that had been developed at Berlin by his mentors Weierstrass and Kronecker. The result was a substantial and fruitful tool for formulating and investigating linear algebraic problems, and in his work after 1877, when a problem he was studying involved linear algebra, he made good use of matrix algebra whenever appropriate.

In what follows I have focused on two particularly interesting problems that Frobenius solved with the aid of matrix algebra.[1] Both problems originated from the work of Kronecker and Weierstrass and their histories are interrelated. The first to be considered (in Sect. 3) involved a problem posed by Kronecker on the complex multiplication of abelian functions that Frobenius solved definitively with the aid of matrix algebra. (Readers take note: no knowledge of abelian functions is needed to understand Sect. 3; the presentation is self-contained and on an elementary level.) The discussion of this problem also shows that, although Laguerre developed matrix algebra on the generic level, his application of it to the work of Hermite on abelian functions was anything but routine and in fact seems to have inspired Frobenius' solution to the complex mul-

[1] The research on which the following account is based was done with the financial support of the NSF Science and Technology Studies program under grant SES-0312697.

tiplication problem. The second problem, which I call the congruent transformation problem (Sect. 4) arose from the Weierstrass–Kronecker theory of the transformation of bilinear forms, and Frobenius' solution to it illustrated the manner in which relatively brief matrix algebraic reasoning could replace copious equation manipulation and subtle determinant based transformations.

In order to understand Frobenius' solutions to these problems it is necessary to discuss some of the concepts and results in his 1877 paper on matrix algebra. This is done in Sect. 2, which also serves to convey a greater sense of how and in what sense Frobenius fused matrix algebra with the techniques of Berlin-style linear algebra. In the final section of the paper I consider the extent to which the example set by Frobenius' many applications of matrix algebra played a role in making matrix algebra a standard tool of linear algebra in the twentieth century.

2 Frobenius' 1877 paper on matrix algebra

Frobenius' paper, like the earlier ones by Cayley and Laguerre, was motivated by a specific problem that was posed by Hermite, whose formulation of and proposed solution to the problem were in turn inspired by earlier work of Cayley.[2] The problem was to determine all nonsingular linear transformations that leave a nonsingular quadratic form invariant. Expressed in modern notation the problem is to determine the transformations $\mathbf{x} = U\mathbf{y}$ that take the quadratic form $\Phi(\mathbf{x}) = \mathbf{x}^t A\mathbf{x}$, $A^t = A$, $\det A \neq 0$, into itself, i.e., $\Phi(U\mathbf{y}) = \Phi(\mathbf{y})$ so that

$$U^t AU = A. \tag{2.1}$$

Frobenius dealt with the problem in a manner that became his hallmark.[3] He made a careful study of all the relevant literature on the problem (albeit missing the papers by Cayley and Laguerre on matrix algebra[4]), he solved the problem by means of a novel approach (matrix algebra), and he then systematically developed the mathematics of the new viewpoint with the solution of the original problem as an application. The resulting paper reads like a lucid "Treatise on Matrix Algebra and its Applications" and took up 63 of the large quarto pages of Crelle's *Journal*.

2.1 Minimal polynomials

Frobenius' solution to Hermite's problem has been briefly indicated [22, p. 100] and will not concern us here. Two parts of Frobenius' "treatise" are, however, essential for

[2] For the history of the Cayley–Hermite problem and its role in the creation of matrix algebra, see [22].

[3] See in this connection my recent discussion of Frobenius' characteristic approach to creating solutions to mathematical problems, including the problem of Pfaff [25, p. 431ff.].

[4] The discussion of the Cayley–Hermite problem in the 1850s and its revival in the 1870s (by Bachmann, Hermite, and Rosanes) was all carried out in the pages of "Crelle's Journal," the *Journal für die reine und angewandte Mathematik*, whereas the papers by Cayley and Laguerre on matrix algebra were published, respectively, in the *Transactions* of the Royal Society of London and the journal of the École Polytechnique in Paris, neither of which was a publication routinely scoured by German mathematicians.

what is to follow. The first is related to the fact that once matrix algebra is introduced, the formalism suggests that if $\varphi(\lambda) = \det(\lambda I - A)$ is the characteristic polynomial of A, then $\varphi(A) = 0$. This is of course what is now usually called the Hamilton–Cayley Theorem.[5] Both Cayley and Laguerre posited this theorem in their respective papers and both of them simply verified it by direct computation with 2×2 matrices. (Cayley assured his readers that he had also verified it for 3×3 matrices.) Unaware of their work, Frobenius no doubt conjectured that $\varphi(A) = 0$ as well, but as a Berlin trained mathematician it would have been unacceptable to "prove" this by a computation with 2×2 matrices. Instead, he approached the matter by a way of thinking that was nongeneric and also, I suspect, encouraged by Weierstrass' theory of elementary divisors, published in 1868 [60].

If one is content to reason generically then the characteristic polynomial $\varphi(A) = \det(\lambda I - A)$ is "in general" the polynomial of minimal degree satisfied by A. Schooled at Berlin, Frobenius did not think this way, and so he introduced the concept of what is now called the minimal polynomial of A. As Frobenius defined it (without giving it a name) $\psi(\lambda)$ is a polynomial of minimal degree with the property that $\psi(A) = 0$. That such a polynomial must exist follows, he observed, from the fact that the maximum number of linearly independent $n \times n$ matrices is n^2; thus $I, A, \ldots, A^{n^2-1}$ must be linearly dependent and so $m \overset{\text{def.}}{=} \deg \psi \leq n^2 - 1$. In what follows I will assume $\psi(\lambda)$ has leading coefficient 1, although Frobenius did not.

Not only is a suspicion of the truth of the Hamilton–Cayley Theorem together with a nongeneric approach enough to suggest the idea of the minimal polynomial, the idea is also suggested by Weierstrass' theory of elementary divisors. For present purposes Weierstrass' theory may be summed up in the following theorem, which I state in modern notation.

Theorem 2.1 (Weierstrass, 1868) *If* $\beta(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t(\lambda A - B)\mathbf{y}$ *is a pencil of bilinear forms that is nonsingular (*$\det A \neq 0$*) then: (1) nonsingular transformations* $\mathbf{x} = P\mathbf{X}$, $\mathbf{y} = Q\mathbf{Y}$ *exist such that*

$$\beta(\mathbf{x}, \mathbf{y}) = \mathbf{X}^t(\lambda I - J)\mathbf{Y}, \quad \text{or, equivalently} \quad P^t(\lambda A - B)Q = \lambda I - J,$$

where J *is a matrix in Jordan canonical form. (2) Two nonsingular pencils have the same normal form* $\lambda I - J$*, and hence can be transformed into one another, if and only if they have the same elementary divisors.*

Weierstrass' definition of the elementary divisors of $\lambda A - B$ directly relates to the coefficients of $\lambda A - B$ by taking polynomial greatest common divisors of $n-k \times n-k$ minors for $k = 1, \ldots, n-1$ [23, p. 120]. It turns out that each elementary divisor is the characteristic polynomial of a Jordan block of $\lambda I - J$ and hence of the form $(\lambda - a)^e$

[5] Hamilton's name is apparently included because in his *Lectures on Quaternions* (1853) he considered an operation ϕ that in effect defines a linear transformation of $\mathbb{R}^3$ [18, pp. 566–567]. Expressed using present day notation for scalar and cross products ϕ may be defined as follows. Let $\mathbf{a}^{(i)}$, $\mathbf{b}^{(i)}$, $i = 1, 2, 3$, and $\mathbf{c}$ be fixed vectors and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$; then $\phi(\mathbf{r}) = \sum_{i=1}^{3} -(\mathbf{a}^{(i)} \cdot \mathbf{r})\mathbf{b}^{(i)} + \mathbf{c} \times \mathbf{r}$. Hamilton showed that ϕ satisfies a cubic equation, which turns out to be its characteristic equation. He also wrote down the characteristic polynomial for $\psi(\mathbf{r}) = \phi(\mathbf{r}) + g\mathbf{r}$, where g is a scalar.

where a is a root of $\varphi(\lambda)$ and the corresponding Jordan block is $e \times e$. It follows that $\varphi(\lambda)$ is the product of the elementary divisors. If $a_1, \ldots, a_d$ are the distinct roots of $\varphi(\lambda)$, then for each a_i there will be an elementary divisor of maximal degree e_i; and if we set $\psi^*(\lambda) = \prod_{i=1}^{d}(\lambda - a_i)^{e_i}$ then in general $\psi^*(\lambda)$ will be a proper factor of $\varphi(\lambda)$. In fact from Weierstrass' theory it followed that $\psi^*(\lambda) = \varphi(\lambda)/D_1(\lambda)$, where $\varphi(\lambda) = \det(\lambda I - A)$ is the characteristic polynomial of A and $D_1(\lambda)$ is the greatest common divisor of the $n - 1 \times n - 1$ minors of $\lambda I - A$.

In view of these considerations, it would have been reasonable for Frobenius to conjecture that $\psi^*(A) = 0$ and that in fact ψ^* is the minimal polynomial ψ. The following theorem verified this conjecture.

Theorem 2.2 (Frobenius, 1877) *Given any $n \times n$ matrix A, let $\varphi(\lambda) = \det(\lambda I - A)$ and let $\psi(\lambda)$ be the minimal polynomial of A. Then $\varphi(\lambda) = \pm\psi(\lambda)D_1(\lambda)$. Hence:* (i) *$m = \deg \psi = n - \deg D_1 \leq n$;* (ii) *$\varphi(A) = 0$;* (iii) *$\psi(\lambda)$ has the same distinct roots as $\varphi(\lambda)$;* (iv) *$\psi(\lambda)$ has no multiple roots if and only if all the elementary divisors of A are linear so that A is similar to a diagonal matrix.*

Part (ii) is of course the Hamilton–Cayley Theorem, here proved for the first time. From Weierstrass' definition of the elementary divisors of the special pencil $\lambda I - A$ it followed that $\psi(\lambda)$, being represented as the quotient of φ and D_1, equals $\prod_{i=1}^{d}(\lambda - a_i)^{e_i}$, where, as above, $a_1, \ldots, a_d$ are the distinct roots of $\varphi(\lambda)$ and $(\lambda - a_i)^{e_i}$ is the elementary divisor for that root of maximal degree. Thus Frobenius' theorem implies that $\psi(\lambda)$ is the product of the elementary divisors of maximal degree for each root a_i, from which (iii) follows. Part (iv) also followed from $\psi(\lambda) = \prod_{i=1}^{d}(\lambda - a_i)^{e_i}$: all the elementary divisors of A are linear precisely when all $e_i = 1$, i.e., when $\psi(\lambda)$ has no multiple roots. As Frobenius showed in his paper [8, p. 363] Weierstrass' theory as applied to the pencil $\lambda I - A$ implies that all the elementary divisors are simple if and only if A is *similar* to a diagonal matrix (its Jordan form J).[6] Frobenius defined the notion of similarity as we do today.

The minimal polynomial was a key tool in Frobenius' application of matrix algebra. For example, in his 1877 paper he used it to discuss the question of what matrices B commute with a given matrix A. In his 1858 paper on matrix algebra, Cayley had already considered this question [2, Sect. 33]. By reasoning generically and with 2×2 matrices, he concluded that the only such B are polynomials in A. In contrast, Frobenius showed that Cayley's conclusion (for matrices of any size) is true if and only if the characteristic and minimal polynomials are identical. He also showed that when the minimal polynomial has no multiple roots, then all B satisfying $AB = BA$ are expressible in the form $B = \sum_{i=1}^{d} \psi_k(A)C_k\psi_k(A)$, where $\psi_k(\lambda) = \psi(\lambda)/(\lambda - a_k) = \prod_{i \neq k}(\lambda - a_i)$ and $C_1, \ldots, C_d$ are arbitrary $n \times n$ matrices [8, p. 370].[7] The minimal polynomial was also used by Frobenius to prove that the only real division algebras of $n \times n$ matrices are the real numbers ($n = 1$), the complex numbers ($n = 2$) and the quaternions ($n = 4$) [8, Sect. 14].

[6] When all the elementary divisors of $\lambda I - A$ are simple, J is a diagonal matrix and P, Q exist so that $P^t(\lambda I - A)Q = \lambda I - J$, which implies that $P^t Q = I$ and so $RAR^{-1} = J$, $R = P^t$.

[7] He also gave a general formula for the number of linearly independent B that commute with A.

2.2 Orthogonal matrices

The other topic from Frobenius' 1877 paper that needs to be mentioned is that of orthogonal matrices. These originated in Lagrange's elegant non-trigonometric reformulation of the rotational transformations

$$\begin{aligned} x &= t\cos\ \xi + u\sin\ \xi\ \cos\eta - v\sin\ \xi\ \sin\eta \\ y &= -t\sin\ \xi + u\cos\ \xi\ \cos\eta - v\cos\ \xi\ \sin\eta \\ z &= u\sin\eta + v\cos\eta \end{aligned} \tag{2.2}$$

that had been introduced by Euler [21, p. 16ff.]. Cauchy extended Lagrange's formulation to n variables in his generalization of the principal axes theorem for quadric surfaces. Thus he considered transforming a real quadratic form in n variables $x_1, \ldots, x_n$ into a sum of squares $\sum_{i=1}^{n} \lambda_i y_i^2$ by means of transformations of the form $x_i = \sum_{j=1}^{n} u_{ij} y_j$, $i = 1, \ldots, n$, with real coefficients u_{ij} satisfying the generalized version of Lagrange's orthogonality relations, viz

$$\sum_{i=1}^{n} u_{ij}^2 = 1 \quad (j = 1, \ldots, n), \qquad \sum_{k=1}^{n} u_{ki} u_{kj} = 0 \quad (i \neq j). \tag{2.3}$$

The coefficient matrices $U = (u_{ij})$ of Cauchy's transformations satisfy $U^t U = I$, and Frobenius referred to them as orthogonal, as we still do today.[8] And, as we do today, Frobenius included possibly complex orthogonal matrices U, although for Lagrange and Cauchy only real orthogonal transformations were considered. The U of Euler's rotations (2.2) satisfy $\det U = 1$, whereas for general orthogonal transformations one has only $\det U = \pm 1$. When $\det U = 1$, I will follow Frobenius and call U a proper orthogonal matrix or transformation.

Orthogonal transformations were important in connection with Hermite's problem since they form the solutions to (2.1) when $A = I$, and in fact Hermite's formulation of the problem and the form of his solution owed much to a paper Cayley wrote in 1846 [22, p. 88–89]. There Cayley showed how proper orthogonal transformations U could be expressed as rational functions of the $n(n-1)/2$ parameters of a coefficient system (t_{ij}) with the skew symmetry property $t_{ji} = -t_{ij}$. In the notation of matrix algebra, Cayley's formula for U can be expressed as

$$U = \frac{I - T}{I + T}, \quad T^t = -T, \tag{2.4}$$

which is now called the Cayley transform. Actually Cayley left it unclear whether his formula gave all proper orthogonal transformations, and so Frobenius considered the matter. Using matrix algebra he showed easily that (2.4) represents all proper orthogonal transformations such that $\det(U + I) \neq 0$, i.e., such that $\lambda = -1$ is

[8] Strictly speaking Frobenius spoke of orthogonal forms and orthogonal substitutions, not orthogonal matrices, since his coefficient systems represented either bilinear forms or linear substitutions (i.e., linear transformations).

not a characteristic root of U. (His solution to Hermite's problem also implied that every proper orthogonal U for which $\det(U + I) = 0$ is expressible in the form $U = \lim_{h\to 0} U_h$, where each U_h is of the form (2.4).)

Thus Cayley's formula (2.4) did not represent all proper orthogonal U, let alone all orthogonal transformations. This posed another problem for Frobenius because in 1856 Brioschi had used Cayley's formula to say something about the characteristic roots of real transformations in n variables $\mathbf{x} = U\mathbf{y}$ satisfying the Lagrange–Cauchy orthogonality relations (2.3). Using Cayley's formula for U and reasoning generically, Brioschi had concluded that the roots of $\varphi(\lambda) = \det(\lambda I - U)$ all satisfy $|\lambda| = 1$. Frobenius thus explored the question of the extent of validity of Brioschi's theorem for all real orthogonal matrices, as well as the related matter of the nature of their elementary divisors. Combining matrix algebra with Weierstrass' theory he proved the following theorem [8, pp. 393–395].

Theorem 2.3 (Frobenius, 1877) *Let U be any real orthogonal matrix. Then:* (1) $|\lambda| = 1$ *for all characteristic roots λ of U;* (2) *all the elementary divisors of U are linear (so that U is similar to a diagonal matrix).*

Nowadays part (1) of Frobenius' theorem is proved by a simple "inner product" type argument. However, such arguments—which can be given without formally introducing an inner product—remained unfamiliar to most mathematicians in the nineteenth century. Christoffel and Clebsch were the first to introduce them in the early 1860s in order to establish the nature of the characteristic roots of hermitian symmetric coefficient systems,[9] but Frobenius, who frequently cited work by Christoffel, failed to see the possibility of proving (1) this way. His own proof is nonetheless fairly simple and indicates the manner in which he combined Weierstrassian techniques with matrix algebra. In 1858 Weierstrass had proved a theorem about certain pencils of quadratic forms $\lambda A - B$ (namely Theorem 3.3, which is given in Sect. 3.3) that prefigured his Theorem 2.1 on elementary divisors, and he had made critical use of Laurent expansions of the coefficients of what we may now write as $(\lambda A - B)^{-1}$. That is, in nineteenth century linear algebra the adjoined matrix Adj (M) associated to a given matrix M played a central role. The (i, j) coefficient of Adj (M) is the (j, i) cofactor of M and the fundamental property was the system of equations we can express as $M\,\mathrm{Adj}(M) = (\det M)\, I$. For the $n \times n$ system $\lambda A - B$ the coefficients of the adjoined system are thus polynomials $\varphi_{ij}(\lambda)$ of degree at most $n - 1$. Weierstrass considered Laurent expansions of the rational functions $\varphi_{ij}(\lambda)/\varphi(\lambda)$, $\varphi(\lambda) = \det(\lambda A - B)$, which are precisely the coefficients of $(\lambda A - B)^{-1}$.

Thus if $\lambda = a$ is a characteristic root of U, matrix algebra enabled Frobenius to express Weierstrass' idea in the succinct form of a matrix Laurent expansion

$$(\lambda I - U)^{-1} = A(\lambda - a)^{-k} + \text{higher powers of } (\lambda - a), \tag{2.5}$$

[9] Clebsch's proof [4, p. 327ff.] was fatally flawed, but in a subsequent paper [5, p. 233ff.] he used similar "inner product" considerations to proved correctly that the roots of a skew symmetric matrix are purely imaginary. Christoffel [3, p. 131] gave a correct and very simple "inner product" proof in the hermitian case.

valid in some region $0 < |\lambda - a| < r$.[10] Multiplying both sides of (2.5) on the right by $\lambda I - U$ and writing $\lambda I - U = (\lambda - a)I - (U - aI)$ on the resulting right-hand side yields

$$I = -A(U - aI)(\lambda - a)^{-k} + \text{higher powers of } (\lambda - a),$$

and since I has no pole at $\lambda = a$, it must be that $A(U - aI) = 0$, which may be written as $AU = aA$. Using conjugation and transposition and the reality of U, Frobenius observed that since $AU = aA$ and $UU^t = I$ it follows that

$$|a|^2 A\bar{A}^t = (aA)(\bar{a}\bar{A}^t) = (AU)(U^t\bar{A}^t) = A\bar{A}^t,$$

which implies $|a| = 1$, since $A\bar{A}^t \neq 0$.[11]

The same matrix Laurent expansion technique enabled Frobenius to give a simple proof that the elementary divisors of U are simple, i.e., that U can be diagonalized, something that is still not so easy to prove quickly nowadays. Starting with (2.5), differentiation with respect to λ gives

$$-(\lambda I - U)^{-2} = -kA(\lambda - a)^{-k-1} + \text{higher powers of } (\lambda - a), \tag{2.6}$$

whereas squaring both sides of (2.5) gives

$$(\lambda I - U)^{-2} = A^2(\lambda - a)^{-2k} + \text{higher powers of } (\lambda - a). \tag{2.7}$$

By means of additional matrix algebra, Frobenius showed that since $A \neq 0$, $A^2 \neq 0$ as well. Comparison of (2.6) and (2.7) then showed (by uniqueness of Laurent expansions) that $(\lambda I - U)^{-2}$ has a pole of order $2k = k + 1$ at $\lambda = a$, whence $k = 1$ in (2.5). To anyone familiar with elementary divisor theory, the fact that $(\lambda I - U)^{-1}$ has a pole of order $k = 1$ at every characteristic root $\lambda = a$ meant that the elementary divisors of U are all linear.[12]

About a year after Frobenius published his paper on matrix algebra he became interested in a particular aspect of the theory of abelian functions, and in the course of studying the relevant literature he came across Kronecker's complex multiplication problem, to which I now turn.

[10] Thus k is the largest order of a pole at $\lambda = a$ of some coefficient $\varphi_{ij}(\lambda)/\varphi(\lambda)$ of $\text{Adj}(\lambda I - U)/\det(\lambda I - U) = (\lambda I - U)^{-1}$.

[11] The fact that $A \neq 0$ implies that $A\bar{A}^t \neq 0$ is based on the hermitian symmetry of $A\bar{A}^t$. Frobenius simply referred to a similar argument when AA^t is real symmetric [8, p. 346].

[12] Since $(\lambda I - U)^{-1} = \text{Adj}(\lambda I - U)/\det(\lambda I - U) \overset{\text{def}}{=} \varphi_{ij}(\lambda)/\varphi(\lambda)$, all the coefficients $\varphi_{ij}(\lambda)/\varphi(\lambda)$ have poles of order at most 1 at $\lambda = a$, i.e., if $(\lambda - a)^q$ divides $\varphi(\lambda)$ then $(\lambda - a)^{q-1}$ divides $\varphi_{ij}(\lambda)$ for all i, j. This means that $D_1(\lambda)$, being the greatest common divisor of all the $\varphi_{ij}(\lambda)$, is also divisible by $(\lambda - a)^{q-1}$, and so $\varphi(\lambda)/D_1(\lambda)$ is the product of distinct linear factors. But from Weierstrass' definition of elementary divisors, $\varphi(\lambda)/D_1(\lambda) = \prod_{i=1}^{d}(\lambda - a_i)^{e_i}$, where $a_1, \ldots, a_d$ are the distinct characteristic roots of U and $(\lambda - a)^{e_i}$ is the elementary divisor for a_i of maximal exponent. Hence all $e_i = 1$ and all elementary divisors are linear.

3 Kronecker's complex multiplication problem

One of the highlights of early nineteenth century mathematics had been the discovery by Abel and Jacobi of the existence of nonconstant functions of a complex variable with two periods linearly independent over $\mathbb{R}$. The theory of these doubly periodic or elliptic functions became a major area of research throughout the nineteenth century. In the middle of the century some mathematicians began to explore the possibility of generalizing the theory to functions of $g > 1$ complex variables. For example, Hermite made important and influential contributions to the nascent theory in the case of $g = 2$ variables in a memoir of 1855 [27] that will be discussed below. Weierstrass and Riemann independently took up the development of the theory in any number g of variables, the theory of abelian functions as it eventually came to be called. This theory was central to Weierstrass' research program, and as we shall see, it was as a result of Kronecker's efforts to assist him on algebraic aspects of the theory that Kronecker formulated the complex multiplication problem. In order to state the problem, some definitions and historical background first need to be indicated.

3.1 Abelian matrices

Let $f(\mathbf{z})$ denote a meromorphic function of g complex variables $z_1, \ldots, z_g$, which I will denote as a column matrix: $\mathbf{z} = \begin{pmatrix} z_1 & \cdots & z_g \end{pmatrix}^t$. Then f has period $\mathbf{p} = \begin{pmatrix} p_1 & \cdots & p_g \end{pmatrix}^t \in \mathbb{C}^g$ if $f(\mathbf{z}+\mathbf{p}) = f(\mathbf{z})$. If f has $2g$ periods $\mathbf{p}_1, \ldots, \mathbf{p}_{2g}$ that are linearly independent over $\mathbb{R}$ it is called an abelian function. Thus an abelian function in $g = 1$ complex variable is an elliptic function. Given an abelian function f with periods $\mathbf{p}_1, \ldots, \mathbf{p}_{2g}$, then clearly for any integers $m_1, \ldots, m_{2g}$ we have $f(\mathbf{z} + m_1\mathbf{p}_1 + \cdots + m_{2g}\mathbf{p}_{2g}) = f(\mathbf{z})$. The points $m_1\mathbf{p}_1 + \cdots + m_{2g}\mathbf{p}_{2g}$ form what is now called a lattice in $\mathbb{C}^g$.

The $2g$ periods of $f(\mathbf{z})$, being expressed as column matrices, may be combined to form what I will call the $g \times 2g$ *period matrix of* f, namely $P = \begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_{2g} \end{pmatrix}$. In order for nonconstant abelian functions to exist, P must satisfy certain conditions (usually called Riemann's conditions).[13] A simple but representative example of a period matrix P satisfying these conditions is

$$P = \begin{pmatrix} I_g & T \end{pmatrix}, \quad T \in \mathfrak{H}_g. \tag{3.1}$$

In (3.1) I_g denotes the $g \times g$ identity matrix and $\mathfrak{H}_g$ denotes what is now called the Siegel half-space; it consists of all complex symmetric $g \times g$ matrices $T = \Phi + i\Psi$ with the property that the imaginary part Ψ, which is of course real symmetric, is also positive definite. Thus an abelian function f with period matrix (3.1) has the standard basis for $\mathbb{C}^g$ as g of its periods, and the remaining g periods form the complex symmetric matrix $T \in \mathfrak{H}_g$. In what follows P will be assumed in the form (3.1). A basic theorem of the theory is that abelian functions are quotients of theta functions. Theta

[13] Riemann's conditions in terms of period matrices are given by Frobenius [11, p. 102, (4) and (6)]. They are also given as part of a lucid modern exposition by Rosen [52, pp. 98–99], together with a simple example of a period matrix with no nonconstant abelian functions.

functions are generated by infinite series that involve a complex symmetric coefficient system $T = \Phi + i\Psi$ in their definition and Ψ must be positive definite in order to insure the uniform convergence of the theta series on compact subsets of $\mathbb{C}^g$. The condition that $T \in \mathfrak{H}_g$ in (3.1) was consequently essential.

Kronecker's problem was related to that aspect of the theory dealing with the transformation of abelian and theta functions by means of a variable change $\mathbf{z} = M\mathbf{w}$, where $\det M \neq 0$. The objective was to choose M so that if $f(\mathbf{z})$ is an abelian function with period matrix $P = \begin{pmatrix} I_g & T \end{pmatrix}$, $T \in \mathfrak{H}_g$, then $g(\mathbf{w}) = f(M\mathbf{w})$ should have a period matrix of the form $P' = \begin{pmatrix} I_g & T' \end{pmatrix}$, where $T' \in \mathfrak{H}_g$. This means that if $\mathbf{p}'_j$ denotes the jth column of P', $1 \leq j \leq 2g$, then $g(\mathbf{w} + \mathbf{p}'_j) = g(\mathbf{w})$, or equivalently $f(\mathbf{z} + M\mathbf{p}'_j) = f(\mathbf{z})$; and so $M\mathbf{p}'_j$ is a period of f and hence a $\mathbb{Z}$-linear combination of the $2g$ periods of f given by the $2g$ columns of $P = \begin{pmatrix} I_g & T \end{pmatrix}$. In other words M takes the lattice of periods associated to $P' = \begin{pmatrix} I_g & T' \end{pmatrix}$ into the corresponding lattice for P. Thus M must transform each column of P' into some integral linear combination of the columns of P. Since what M does to each column of P' is given by the columns of $MP' = \begin{pmatrix} M & MT' \end{pmatrix}$, the above integrality condition may be stated as

$$M = \mathrm{A} + T\Gamma, \quad MT' = \mathrm{B} + T\Delta, \tag{3.2}$$

where the capital Greek letters A, B, Γ, Δ stand for $g \times g$ matrices with integer coefficients. M can be eliminated from (3.2) to obtain

$$T' = (\mathrm{A} + T\Gamma)^{-1}(\mathrm{B} + T\Delta), \tag{3.3}$$

since $M = \mathrm{A} + T\Gamma$ is assumed invertible.

The $g \times g$ blocks of integers were frequently combined into a $2g \times 2g$ array, which may be represented by the block-partitioned matrix

$$\widetilde{A} = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ \Gamma & \Delta \end{pmatrix}. \tag{3.4}$$

In order that the integers comprising $\widetilde{A}$ define an invertible transformation $M = \mathrm{A} + T\Gamma$ that takes $P = \begin{pmatrix} I_g & T \end{pmatrix}$ into $P' = \begin{pmatrix} I_g & T' \end{pmatrix}$ with $T' \in \mathfrak{H}_g$ when $T \in \mathfrak{H}_g$ it is necessary that $\widetilde{A}$ have a special property that is easy to express in matrix notation: there is a positive integer n such that

$$\widetilde{A}^t J \widetilde{A} = nJ, \quad \text{where } J = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}. \tag{3.5}$$

From this relation it follows that $\widetilde{A}$ is invertible. For ease of reference, I will call any $2g \times 2g$ integral matrix that satisfies (3.5) an *abelian matrix of order n*.

To sum up what has been said thus far: If $\widetilde{A}$ is abelian then it determines a nonsingular variable change $\mathbf{z} = M\mathbf{w}$, $M = \mathrm{A} + T\Gamma$, that transforms any abelian function

with period matrix $P = \left(I_g \ T\right)$, $T \in \mathfrak{H}_g$, into one with period matrix $P' = \left(I_g \ T'\right)$, where also $T' \in \mathfrak{H}_g$, and the transformation $T \to T'$ is given by (3.3).[14]

The determination of abelian matrices $\widetilde{A}$ was central to the nineteenth century theory of abelian and theta functions. For $g = 2$ variables, Hermite, in the above-mentioned paper of 1855 [26], had shown how to classify all abelian matrices of prime order. The classification was based upon the fact (easily verified by matrix algebra) that if $\widetilde{A}_1$ and $\widetilde{A}_2$ are abelian of orders n_1 and n_2 then $\widetilde{A}_1\widetilde{A}_2$ is abelian of order n_1n_2, as is the reverse product. Thus two abelian matrices $\widetilde{A}_1$ and $\widetilde{A}_2$ of the same order n are said to be equivalent if there is an abelian matrix $\widetilde{L}$ of order 1 such that $\widetilde{A}_1\widetilde{L} = \widetilde{A}_2$.[15] This defines an equivalence relation that partitions the abelian matrices of order n into equivalence classes, which turn out to be finite in number. In the case $g = 2$ and for n prime, Hermite determined the number of equivalence classes and a representative for each. In 1858 Weierstrass asked Kronecker to investigate the classification problem of abelian matrices of order n for any g. He wished to include this topic in a planned work on abelian functions.[16] Kronecker obliged, wrote up his results the following year and gave them to Weierstrass, although he himself did not publish the manuscript he produced.[17]

3.2 Complex multiplication

Kronecker's involvement with abelian matrices for $g > 1$ led him to consider a generalization of a part of the elliptic theory ($g = 1$) that had not been considered by Hermite in 1855. It involved what came to be known as the complex multiplication of elliptic functions, a phenomenon that had been noted with interest by both Jacobi and Abel [33, p. 213] and then further studied by other mathematicians, primarily because of deep arithmetical connections. Kronecker became a leader in the development of the arithmetical aspects of complex multiplication,[18] and his first paper on the subject was published in 1857 [34], the year before Weierstrass approached him with the classification problem.[19] Before describing Kronecker's generalization of complex multiplication to abelian functions, it will be helpful to first describe the simpler elliptic case.

Let $f(z)$ be an elliptic function with periods 1 and τ, and so with period matrix $P = \left(1 \ \tau\right)$, where now $T = (\tau) \in \mathfrak{H}_g$ simply means that τ is in the upper half-plane. Then if m is any integer $g(z) = f(mz)$ has the same periods as f and so was known to

[14] The nonsingularity of $\mathrm{A} + T\Gamma$ for any abelian $\widetilde{A}$ follows from the fact that $T \in \mathfrak{H}_g$ [11, p. 105].

[15] Abelian matrices of order $n = 1$ are precisely the elements of the symplectic group $\mathbf{Sp}(2g, \mathbb{Z})$ defined with respect to the skew symmetric form $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^t J \mathbf{y}$.

[16] This according to Kronecker [35, p. 162].

[17] Some of Kronecker's results were described by him on pp. 158–161 of [35]. These have to do with what would now be described as giving a set of generators for the group of $2g \times 2g$ abelian matrices L of order 1, viz $\mathbf{Sp}(2g, \mathbb{Z})$, an important part of the classification problem not considered by Hermite. For further mathematical and historical details on the classification problem see [33, pp. 148–164].

[18] See in this connection the comments of H. Weber [58, pp. vi–vii], whose book [58] of 1891 and its second edition of 1908 expounded the theory as it had developed in the nineteenth century.

[19] For a description of Kronecker's paper and its relation to his *Jugendtraum*, see [56, p. 66ff.].

be algebraically related to f. This phenomenon was described by saying that f admits ordinary (or real) multiplication. Thus no matter what τ is, ordinary multiplication by any $m \in \mathbb{Z}$ is admitted. However, there are certain periods $1, \tau_0$ such that elliptic functions with those periods admit a "complex multiplication" as well. To see this, and the relation to abelian matrices, consider $\widetilde{A}_1 = \begin{pmatrix} 1 & 6 \\ -8 & 5 \end{pmatrix}$. It is easy to check that $\widetilde{A}_1$ is abelian of order $n = 53$. Thus, in the sense explained above, $\widetilde{A}_1$ induces a transformation of periods $\begin{pmatrix} 1 & \tau \end{pmatrix} \to \begin{pmatrix} 1 & \tau' \end{pmatrix}$, where τ' is given by (3.3), which for $g = 1$ becomes $\tau' = (1-8\tau)(6+5\tau)^{-1}$, and we can ask whether this equation has a suitable solution with $\tau' = \tau$, i.e., whether $\tau = (6+5\tau)^{-1}(1-8\tau)$ has a solution in the upper half-plane. By solving the quadratic equation involved, we see that $\tau_0 = (-1+\sqrt{11}\,\mathrm{i})/4$ is such a solution. This means that the corresponding transformation $\mathbf{z} = M\mathbf{w}$ given by (3.2) is in this case $z = Mw$ with $M = 1 + 8\tau_0 = 3 - 2\sqrt{11}\,\mathrm{i}$. Thus $g(w) = f(Mw)$ has the same periods as f, and so f and g are algebraically related. In this case f was said to admit a complex multiplication by $3 - 2\sqrt{11}\,\mathrm{i}$.[20] Unlike the above-defined ordinary multiplication, complex multiplication is only possible for special periods τ. Furthermore, not every abelian matrix gives rise to a complex multiplication. For example $\widetilde{A}_2 = \begin{pmatrix} 9 & -3 \\ -4 & 3 \end{pmatrix}$ is abelian of order $n = 15$. However in this case (3.3) with $T' = T$ is $\tau = (9 - 4\tau)^{-1}(-3 + 3\tau)$, and the corresponding quadratic equation is $4\tau^2 - 6\tau - 3 = 0$, which has positive discriminant and so has only real solutions. Thus for $\widetilde{A}_2$ there is no solution to (3.3) with $\tau' = \tau$ for a τ in the upper half-plane.

3.3 Kronecker's problem

The work classifying abelian matrices for arbitrary g led Kronecker to pose the problem of determining which abelian matrices $\widetilde{A}$ have the property that (3.3) has a solution with $T' = T$, i.e.,

$$T = (\mathrm{A} + T\Gamma)^{-1}(\mathrm{B} + T\Delta) \quad \text{for some } T \in \mathfrak{H}_g. \tag{3.6}$$

When an abelian matrix $\widetilde{A}$ has this property a T satisfying (3.6) yields a period matrix $P = \begin{pmatrix} I_g & T \end{pmatrix}$ with the property that the lattice Λ generated by the columns of P is taken into itself by $\mathbf{z} = M\mathbf{w}$, where $M = \mathrm{A} + T\Gamma$. I will adopt the terminology introduced later by Frobenius and call an abelian matrix $\widetilde{A}$ *principal* if (3.6) has a solution. Likewise, if $\widetilde{A}$ is principal, a solution $T = (\tau_{ij})$ to (3.6) is called a *singular parameter system* for $\widetilde{A}$. The problem Kronecker posed may now be stated as follows: *determine which abelian matrices are principal*.

In order to state the result Kronecker obtained by investigating this problem, it is necessary to observe that if T is a singular parameter system for a principal $\widetilde{A}$ then it satisfies (3.6), which, after multiplying through by $\mathrm{A} + T\Gamma$, may be rewritten in the form

[20] The connection of this notion with the modern one of abelian varieties admitting complex multiplication is indicated at the end of Sect. 3.4.

$$\mathrm{B} + T\Delta - \mathrm{A}T - T\Gamma T = 0. \tag{3.7}$$

Kronecker's theorem, which he published in 1866 [35], may be stated in the following form by using matrix algebra.

Theorem 3.1 (Kronecker, 1866) *Let $\widetilde{A}$ be an abelian matrix, and set $\widetilde{B} = -J\widetilde{A}$, where J is defined as in (3.5). Then if $\varphi(\lambda) = \det(\lambda\widetilde{B} - \widetilde{B}^t)$ has no multiple roots, there exists a complex symmetric matrix T satisfying (3.7).*

Kronecker's proof of his theorem was ingenious and will be discussed below. First, however, it is important to realize that the theorem, although relevant to the problem, does not provide even a partial solution by providing a sufficient condition that $\widetilde{A}$ be principal. That is, even if the T posited by the theorem is such that $\mathrm{A} + T\Gamma$ is invertible so that (3.6) holds, the theorem does not provide any information about whether or not $T \in \mathfrak{H}_g$, i.e., whether or not the imaginary part of T has the critically important property of being positive definite. Thus the T of Kronecker's theorem need not be a singular parameter system, i.e., $\widetilde{A}$ need not be principal.

Kronecker was well aware of this fact. In this connection he made two important observations: (1) abelian $\widetilde{A}$ exist such that no T satisfying (3.7) has positive imaginary part; (2) abelian $\widetilde{A}$ exist such that $\varphi(\lambda)$ has multiple roots and "the numbers τ remain partially undetermined, i.e., in this case there exist certain functions of one or more variables, which if set equal to the τ_{ik} solve the problem" [35, p. 157]. Kronecker's first observation is easy to see in the elliptic case $g = 1$. (Take the solutions to (3.7) corresponding to the example $\widetilde{A}_2$ considered above.) His observation (2) must have been based on examples with $g > 1$. What he meant by "solve the problem" is, however, not entirely clear. Did he simply mean he knew of examples where $\varphi(\lambda)$ has multiple roots and (3.7) has infinitely many solutions, or did he know of examples in which $\varphi(\lambda)$ has multiple roots and infinitely many T exist satisfying (3.7) *and* with $T \in \mathfrak{H}_g$? Whatever, he meant, his remarks raised the question as to whether an abelian $\widetilde{A}$ such that $\varphi(\lambda)$ has multiple roots can be principal and, if so, whether $\widetilde{A}$ can have infinitely many singular parameter systems associated to it.

Thus Kronecker's original problem, when combined with his remarks, suggests the following elaboration: *Without imposing any generic preconditions, determine which abelian $\widetilde{A}$ are principal and for principal $\widetilde{A}$ determine exactly when the associated singular parameter system $T = (\tau_{ij})$ is unique*. As we shall see, Frobenius gave a definitive solution to the elaborated version of Kronecker's problem in 1883, and matrix algebra played a key role in the solution. First, however, it is necessary to say something about Kronecker's proof because it too is relevant, historically, to the manner in which Frobenius developed matrix algebra—as well as to the congruent transformation problem of Sect. 4.

Kronecker's ingenious proof of Theorem 3.1 was based upon a theorem he proved about the transformation of certain pencils of bilinear forms, namely pencils of the form $\beta(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t(\lambda B - B^t)\mathbf{y}$, where B is $2g \times 2g$ and $\det B \neq 0$. (In the proof of Theorem 3.1 B is taken to be $\widetilde{B} = -J\widetilde{A}$, but here the coefficients of B can be any complex numbers.) It involved making the same linear transformation of both sets of variables so as to put B into the special "normal form" (as he called it)

$$N = \begin{pmatrix} 0 & D_1 \\ D_2 & 0 \end{pmatrix}, \tag{3.8}$$

where D_1, D_2 are diagonal matrices with nonzero entries $\rho_1, \ldots, \rho_g$ and $\rho_{g+1}, \ldots, \rho_{2g}$, respectively, along the diagonals. The transformation of $\mathbf{x}^t B\mathbf{y}$ to $\mathbf{z}^t N\mathbf{w}$ is by means of $\mathbf{x} = P\mathbf{z}, \mathbf{y} = P\mathbf{w}$, $\det P \neq 0$, which Kronecker later called a *congruent transformation* of the bilinear form $\mathbf{x}^t B\mathbf{y}$ [40, p. 424] since both $\mathbf{x}$ and $\mathbf{y}$ are subject to the same transformation. Clearly such a transformation exists if and only if $P^t B P = N$ and hence if and only if

$$P^t(\lambda B - B^t)P = \lambda N - N^t, \tag{3.9}$$

i.e., if and only if the pencil $\beta(\mathbf{x}, \mathbf{y})$ can be transformed congruently into the pencil $\mathbf{z}^t(\lambda N - N^t)\mathbf{w}$. By taking determinants in (3.9) and applying the product theorem, we see that $(\det P)^2\varphi(\lambda) = \det(\lambda N - N^t)$, and so the two pencils have the same characteristic roots. Kronecker showed this condition was also sufficient in the generic case in which $\varphi(\lambda)$ has no multiple roots:

Theorem 3.2 (Kronecker, 1866) *If the pencil of bilinear forms* $\mathbf{x}^t(\lambda B - B^t)\mathbf{y}$ *in* $2g$ *variables is such that* $\det B \neq 0$ *and* $\varphi(\lambda) = \det(\lambda B - B^t)$ *has no multiple roots, then there is a nonsingular congruent transformation,* $\mathbf{x} = P\mathbf{z}$ *and* $\mathbf{y} = P\mathbf{w}$ *such that* $\mathbf{x}^t(\lambda B - B^t)\mathbf{y} = \mathbf{z}^t N\mathbf{w}$, *where* N *is given by* (3.8) *with* $\rho_i = \lambda_i$, $i = 1, \ldots, 2g$. *Hence any two pencils of the form* $\mathbf{x}^t(\lambda B_i - B_i^t)\mathbf{y}$, $i = 1, 2$, *with* $\det B_i \neq 0$ *and* $\varphi_i(\lambda) = \det(\lambda B_i - B_i^t)$ *free from multiple roots can be transformed into one another by means of a congruent transformation if and only if* $(\det B_2)\varphi_1(\lambda) = (\det B_1)\varphi_2(\lambda)$.

The last equality implies that the characteristic polynomials $\varphi_i(\lambda)$ have the same roots and hence the same normal form N, whence the second part of the theorem follows from the first.

Kronecker chose the name "normal form" for $\mathbf{z}^t N\mathbf{w}$

> because *every* bilinear form can be transformed into it. This reduction of the bilinear form of $2n$ variables into the given normal form is of the greatest significance because not only is the above question about the the special values of the quantities τ resolved but also the general transformation of any bilinear form into another is thereby obtained. [35, p. 148]

This passage is noteworthy for several reasons. Although Kronecker's theorem is limited to the case of distinct characteristic roots, as he went on to mention explicitly, he evidently did not regard this as serious limitation at the time, as his italicization of "every" suggests. Indeed, generic results were the norm. For example, the elegant papers of Jacobi on the transformation of quadratic and bilinear forms were filled with generic reasoning [30, p. 212ff., p. 247ff.], [31]. Certainly Jacobi must have realized his calculations often lost their meaning for certain singular values of his variables—this fact was too obvious for a mathematician as capable as he to have overlooked—and he probably realized as well that some of his theorems were only true "in general". I suspect that mathematicians such as Jacobi believed that if generic reasoning were abandoned, the elegant general sort of symbolical mathematics that attended it would

also be lost and replaced by a tedious morass of special cases, each requiring separate consideration.

My suspicion is supported by Weierstrass, who in 1858 admitted that he had entertained such sentiments himself [59, p. 234] but had ended up rejecting the view that one should rest content with generic arguments when he discovered that it was possible to obtain nongeneric results without falling into a morass of special cases. The result in question was the following theorem.

Theorem 3.3 (Weierstrass, 1858) *If* $\Phi = \mathbf{x}^t A\mathbf{x}$ *and* $\Psi = \mathbf{x}^t B\mathbf{x}$ *are two real quadratic forms and* Φ *is positive definite, then:* (1) *the roots of* $\varphi(\lambda) = \det(\lambda A - B)$ *are real, and* (2) *a nonsingular linear transformation* $\mathbf{x} = P\mathbf{y}$ *exists such that in the* y*-variables* $\Phi = y_1^2 + \cdots + y_n^2$ *and* $\Psi = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$*, where* $\lambda_1, \ldots, \lambda_n$ *are the roots of* $\varphi(\lambda)$.

Jacobi had already given an elegant generic proof of (2) in the case $A = I_n$. The elegant determinant-based formula he gave for P became meaningless in certain singular cases. Weierstrass showed, however, that even in the more general context of any positive definite symmetric matrix A, the existence of a generalized version of Jacobi's P could be given by virtue of a key lemma, which showed that if $\lambda = \lambda_0$ is a root of $\varphi(\lambda)$ of multiplicity m, then each $n-1 \times n-1$ minor $\varphi_{ij}(\lambda) \not\equiv 0$ of $\lambda A - B$ has λ_0 as a root of multiplicity $m-1$. (In the light of his subsequent theory of elementary divisors, which was motivated by this result, the lemma shows that the elementary divisors of $\lambda A - B$ are all linear.)

Weierstrass gave Theorem 3.3 in a paper presented to the Berlin Academy four years before Kronecker became a member in 1862, and I doubt Kronecker was familiar with it when he presented his own Theorem 3.2 to the Academy in 1866. Kronecker's and Weierstrass' theorems are similar in that they both concern the transformation of a pencil of forms to a special, simple form, and I suspect that Weierstrass pointed to his own result to indicate the possibility of giving nongeneric proofs to such theorems. Kronecker's own enthusiasm for the theory of the transformation of forms is also manifest in the above quotation and in the fact that he entitled his paper "On bilinear forms" and not "On the transformation of abelian functions." That enthusiasm from his colleague was, I suspect, a major factor in Weierstrass' decision to recreate the theory of elementary divisors that he had originally conceived and written down in 1858 but then lost the manuscript while traveling [12, pp. 719–720]. Shortly thereafter he had suffered a nervous breakdown due to overwork, and he no doubt needed some prodding, once recovered, to devote time to a subject outside his main interest, the theory of abelian functions. Kronecker provided just that. Weierstrass went on to publish his theory of elementary divisors in 1868. He restricted his attention to nonsingular pencils ($\mathbf{x}^t(\lambda A - B)\mathbf{y}$ with $\det A \neq 0$ as in Theorem 2.1) because he knew, as he said, that Kronecker had set himself the task of dealing with singular families. Kronecker's first steps in that direction were taken in a paper presented at the same session of the Academy [36] and then further outlined in a series of papers presented to the Academy in 1874. In these papers Kronecker also articulated the disciplinary ideals exemplified by Weierstrass' papers [59,60], ideals that called for rigorous nongeneric reasoning in algebra.

3.4 Frobenius' solution

During the period 1868–1874 while Weierstrass and Kronecker were creating Berlin-style linear algebra, Frobenius was in Berlin, first as student, then as instructor and finally as teacher in a local high school. He left in 1875 to become a professor at the Polytechnicum in Zürich (now the ETH) and soon began to apply nongeneric Berlin-style linear algebra in his work on the problem of Pfaff,[21] and then on Hermite's problem (discussed above at (2.1)), where, as we have seen, he found it expedient to fuse Berlin-style linear algebra with the symbolical algebra of matrices in his paper of 1877 [8]. Not long after the publication of the 1877 paper, Frobenius' interests turned to the algebraic aspects of the theory of abelian functions, and in particular to the theory abelian matrices. The catalyst for the new research direction was apparently provided by a paper on the subject published by Heinrich Weber in 1878 [57]. Weber's paper was aimed at extending many of Hermite's results for $g = 2$ to the case $g = 3$, including the classification of abelian matrices of prime order along the lines discussed in Sect. 3.1. Inspired by Weber's observations and conjectures about the classification for $g = 3$, Frobenius showed how the algorithms he had recently introduced in order to lay an arithmetical foundation for Weierstrass' theory of elementary divisors [9] could be slightly modified so as to provide a new way of generating and classifying abelian matrices [10]. He also proved Weber's conjecture that if $\widetilde{A}$ is abelian of order n and if $n = p_1 p_2 \cdots p_m$ denotes the factorization of n into (possibly equal) prime numbers p_i, then abelian matrices $\widetilde{A}_i$ of respective orders $p_1, \ldots, p_m$ exist such that $\widetilde{A} = \widetilde{A}_1 \cdots \widetilde{A}_m$. The truth of this conjecture justified restricting the problem of determining abelian matrices to those of prime order.

Weber also considered Kronecker's complex multiplication problem (described at the beginning of Sect. 3.3), and it was probably through Weber's paper that Frobenius became interested in it. Weber assumed that $\widetilde{A}$ was a principal abelian matrix with singular parameter system T, and he sought to deduce properties of $\widetilde{A}$, i.e., necessary conditions that an abelian matrix be principal. He did not practice the generic mode of reasoning in linear algebra, but his linear algebraic tools at the time were the traditional ones, and before long he was forced to assume that the characteristic roots of $\widetilde{A}$ were all distinct.[22] The conclusions he reached are summarized in the following theorem.

Theorem 3.4 (Weber, 1878) *Let $\widetilde{A}$ be a principal abelian matrix with the property that the characteristic polynomial of $\widetilde{A}$ has no multiple roots. Then* (1) *$\widetilde{A}$ can have no real characteristic roots and* (2) *there is only one associated singular parameter system T.*

Part (1) of Weber's theorem generalized what was known in the elliptic case $g = 1$, since in that case $\widetilde{A}$ is 2×2 and the only principal $\widetilde{A}$ with multiple roots is $\widetilde{A} = mI_2$, $m \in \mathbb{Z}$. Part (2) confirmed that the phenomenon discovered by Kronecker, namely that when a principal $\widetilde{A}$ has multiple roots, the number of singular parameter systems can be infinite, is indeed limited to the multiple root case. Of course it remained moot

[21] See in this connection my recent paper [25].

[22] Incidentally, Weber's assumption is weaker than Kronecker's assumption that the roots of $\varphi(\lambda) = \det(\lambda\widetilde{B} - \widetilde{B}^t)$ be distinct, i.e., $\varphi(\lambda)$ can have multiple roots when $\det(\lambda I - \widetilde{A})$ does not, but not conversely.

whether or not a principal $\widetilde{A}$ with multiple roots necessarily gives rise to more than one singular parameter system T.

Frobenius was familiar with Weber's paper by the spring of 1879,[23] but it was not until January 1883 that he was able to submit a paper containing a definitive solution to Kronecker's problem. In the interim he had discovered not only that Laguerre had anticipated his idea of a symbolical algebra of matrices a decade earlier, but that he had applied it to Hermite's results on abelian matrices, not only translating them into the notation of matrix algebra but also extending them in the process from the case $g = 2$ to any integer $g \geq 1$ [45, Sect. V]. As indicated in Sect. 2.1, Laguerre developed matrix algebra exclusively on the formal, generic level, but, as we shall now see, what he did on that level in reformulating some of Hermite's results was substantial and provided Frobenius with the key to resolving Kronecker's problem.

In his pioneering memoir of 1855 establishing the theory of abelian matrices Hermite needed to prove if $\widetilde{A}$ is abelian and if $T \to T'$ by virtue of $\widetilde{A}$ in the sense of (3.3), then when $T = \Phi + \mathrm{i}\Psi \in \mathfrak{H}_g$, the same is true of $T' = \Phi' + \mathrm{i}\Psi'$. The proof that T' is complex symmetric was relatively easy, but proving that Ψ' is also positive definite was more difficult [26, pp. 456–458]. To that end Hermite used the coefficients of Φ and Ψ to define a real quadratic form $f(\mathbf{x}) = \mathbf{x}^t H \mathbf{x}$ in 4 (=$2g$) variables, which he showed to be positive definite and to have special properties under transformations of the form $\mathbf{x} = \widetilde{A}\mathbf{y}$, properties he had singled out earlier in his paper [26, pp. 450–452]. Hermite used these properties to prove that Ψ' is positive definite.

Hermite had developed the theory of abelian matrices based on the identity $\widetilde{A}^t J_h \widetilde{A} = n J_h$, where

$$J_h = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \quad \text{rather than} \quad J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Laguerre chose to work with J instead of J_h. If one thinks of J_h as defining the alternating form $a(\mathbf{x}) = \mathbf{x}^t J_h \mathbf{x}$—a type of association made by Laguerre—then if $\mathbf{x} = P_{34}\mathbf{y}$, where P_{34} is the permutation matrix obtained from I_4 by transposing rows 3 and 4, then $a(\mathbf{y}) = \mathbf{y}^t P_{34} J_h P_{34} \mathbf{y} = \mathbf{y}^t J \mathbf{y}$. Probably this variable change was made by Laguerre because he realized that Hermite's results would then take on a more succinct and satisfying symbolical matrix form that is more readily extended to any value of g. In particular, Hermite's real quadratic form is expressible as $f(\mathbf{y}) = \mathbf{y}^t L \mathbf{y}$ where $L = P_{34} H P_{34}$ and L has a particularly enlightening form as a block-partitioned matrix, viz

$$L = \begin{pmatrix} \Psi_0 & \Psi_0 \Phi \\ \Phi \Psi_0 & \Phi \Psi_0 \Phi + (\det \Psi)\Psi \end{pmatrix}, \tag{3.10}$$

[23] Frobenius' first paper inspired by Weber's, namely [10], was submitted in May 1879.

where Ψ_0 is Laguerre's notation for the adjoined matrix of Ψ, i.e., the transposed matrix of cofactors of Ψ [45, pp. 263, 265].[24] Laguerre seems to have been the first to exploit the block multiplication of partitioned matrices, which, of course, enters naturally into the theory of abelian matrices since from the outset they were regarded as partitioned into the four $g \times g$ blocks denoted in (3.4) by A, B, Γ, Δ. Expressing Hermite's form by the matrix L enabled him to make effective use of block multiplication.

To show $s'(x, y) = \begin{pmatrix} x & y \end{pmatrix} \Psi' \begin{pmatrix} x & y \end{pmatrix}^t$ is also positive definite, Hermite obtained a remarkable relationship between s', f, and $\widetilde{A} = (a_{ij})$, namely

$$s'(x, y) = \mu f(a_{21}x - a_{11}y, a_{22}x - a_{12}y, a_{23}x - a_{13}y, a_{24}x - a_{14}y), \tag{3.11}$$

where $\mu = (\det \Psi')/(n \det \Psi)$ [26, p. 457]. In developing Hermite's theory using the matrix representations J and L, Laguerre obtained a new relation, valid for any g, which involves the matrix L and the matrix L' corresponding to $T' = \Phi' + i\Psi'$ [45, p. 264, Eq. (10)]:

$$L' = \mu \widetilde{A}^t L \widetilde{A}, \quad \mu = (\det\ \Psi')/(n \det\ \Psi). \tag{3.12}$$

Laguerre showed that an analog of (3.11) could then be deduced from (3.12) [45, pp. 265–266], which readily implies that Ψ' is positive definite since both Ψ and L are.

Frobenius read Laguerre's paper with an interest in abelian matrices that are principal. When $\widetilde{A}$ is principal, it has a singular parameter system $T = \Phi + i\Psi$ for which $T' = T$ in (3.3). Thus when $\widetilde{A}$ is principal, $\Phi' = \Phi$, $\Psi' = \Psi$, $L' = L$, and μ simplifies to $1/n$ so that Laguerre's relation (3.12) becomes $L = (1/n)\widetilde{A}^t L \widetilde{A}$, which we may write as

$$P^t L P = L, \quad P = n^{-\frac{1}{2}} \widetilde{A}. \tag{3.13}$$

Thus Laguerre's relation states that the linear transformation given by $P = n^{-\frac{1}{2}} \widetilde{A}$ takes the positive definite real quadratic form defined by L into itself. This makes $\widetilde{A}$ quite special, as Frobenius realized. For example, when $L = I$ (3.13) becomes $P^t P = I$, which says that $P = n^{-\frac{1}{2}} \widetilde{A}$ is real and orthogonal. More generally, since L is positive definite, the principal axes theorem implies that an orthogonal transformation $\mathbf{x} \to \mathbf{y}$ exists so that $\mathbf{x}^t L \mathbf{x} = \lambda_1 y_1^2 + \cdots + \lambda_{2g} y_{2g}^2$, and since all $\lambda_i > 0$, the further transformation $y_i = \lambda_i^{-\frac{1}{2}} z_i$ gives $\mathbf{x}^t L \mathbf{x} = z_1^2 + \cdots + z_{2g}^2$. Expressed in terms of matrix algebra, this says that K exists such that $K^t L K = I$ or, equivalently, $L = Q^t Q$, where $Q = K^{-1}$. Substituting this expression for L in (3.13) we obtain $P^t Q^t Q P = Q^t Q$ and if this equation is multiplied on the left by $(Q^t)^{-1} = (Q^{-1})^t$ and on the right by Q^{-1}, the result may be expressed in the form

$$(QPQ^{-1})^t (QPQ^{-1}) = I, \tag{3.14}$$

[24] Of course this means that $\Psi_0 = (\det \Psi)\Psi^{-1}$, but, surprisingly, Laguerre introduced no notation for an inverse. In the theory of determinants attention was focused on the adjoined system, and Laguerre apparently adhered to custom.

which implies that $S = QPQ^{-1}$ is a real orthogonal transformation. Thus Frobenius could see from his Theorem 2.3 on orthogonal matrices that since $P = n^{-\frac{1}{2}}\widetilde{A}$ is similar to an orthogonal matrix S it inherits the two properties of orthogonal matrices given in that theorem: every characteristic root has absolute value 1 and all the elementary divisors are linear. Since $\widetilde{A} = \sqrt{n}\,P$ it then follows that the characteristic roots of $\widetilde{A}$ all have absolute value $\sqrt{n}$ and that its elementary divisors must also be linear.

According to Frobenius, once it is realized that these two properties of a principal abelian matrix are necessary, it is "easy to show" that they are also sufficient [11, p. 111]. Of course this was easy for Frobenius because he was master of the algebraic aspects of the theory of abelian and theta functions. The point to be made here is that it was through a fertile combination of matrix algebra and Weierstrass' elementary divisor theory that he was led, with some invaluable assistance from Hermite and Laguerre, to the discovery of the following theorem, which solves Kronecker's problem as originally formulated.

Theorem 3.5 (Frobenius, 1883) *If $\widetilde{A}$ is an abelian matrix of order n, then $\widetilde{A}$ is principal if and only if* (1) *all characteristic roots of $\widetilde{A}$ have absolute value $\sqrt{n}$ and* (2) *the elementary divisors of $\widetilde{A}$ are all linear ($\widetilde{A}$ can be diagonalized).*

What had made Kronecker's problem seem so intractable, was the lack of realization of property (1), which is not generally true for abelian matrices; in fact all the equivalence class representatives given by Hermite for $g = 2$ and by Weber for $g = 3$ fail to have property (1). As for property (2), it indicated the extent to which the ad hoc generic assumptions of Kronecker and Weber were justified: as part (iv) of Frobenius' Theorem 2.2 implies, it is the minimal polynomial of $\widetilde{A}$, rather than the characteristic polynomial, that must have distinct roots.

The above line of reasoning leading from Laguerre's equation (3.12) via (3.13) to properties (1) and (2) is my reconstruction of how Frobenius discovered them based on his remarks [11, p. 98]. Once having discovered them, however, he proceeded in characteristic fashion to develop the theory ab initio along the lines he deemed most suitable for publication. The new approach led him to introduce the notion of a unitary matrix—apparently for the first time in the history of mathematics—and to establish the main properties of such matrices. Since this was all done by means of matrix algebra, a brief digression is in order.[25]

In the reasoning leading to Theorem 3.4, Weber had shown that when $\widetilde{A}$ is principal and all $2g$ of its characteristic roots are distinct, then exactly half of them,

[25] In the 1850s and 1860s when hermitian symmetric matrices and forms were introduced and studied as analogs of real symmetric matrices and quadratic forms, the focus was exclusively on the reality of the characteristic roots. No interest was shown in generalizing the principal axes theorem, which would have led naturally to the notion of a unitary matrix as the "hermitian" analog of a real orthogonal matrix. That Frobenius was the first to have found a use for unitary matrices in his 1883 paper [11] on principal transformations is suggested by a remark by Hurwitz in 1897. In his seminal paper on invariant integrals on Lie groups, Hurwitz, who knew Frobenius' paper [11] (see below), had occasion to introduce unitary transformations (and the special unitary group) in order to perform what Weyl later called "the unitarian trick" (see [24, p. 393]). Apparently because unitary transformations (or substitutions) were still sufficiently novel in 1897, he pointed out to his readers that "these substitutions also come into consideration in other investigations" [29, p. 556, n. 1] and referred them to Frobenius' paper [11].

say $\alpha_1, \ldots, \alpha_g$, are characteristic roots of the $g \times g$ matrix M defined by (3.2), viz $M = \mathrm{A} + T\Gamma$, and the other half are all of the form $\alpha'_1, \ldots, \alpha'_g$ where $\alpha_j \alpha'_j = n$ [57, pp. 141–142]. In view of property (1) of Theorem 3.5, namely $\alpha\bar{\alpha} = n$, Frobenius could see that $\alpha'_j = \bar{\alpha}_j$ when $\widetilde{A}$ is principal. Thus, at least in the generic case, $\widetilde{A}$ and $\widetilde{M} = \begin{pmatrix} M & 0 \\ 0 & \overline{M} \end{pmatrix}$ have the same characteristic roots, viz., $\alpha_1, \ldots, \alpha_g, \bar{\alpha}_1, \ldots, \bar{\alpha}_g$. Since the generic case considered by Weber is not far removed from the case of linear elementary divisors guaranteed by property (2), it was perhaps natural for Frobenius to consider showing that $\widetilde{A}$ and $\widetilde{M}$ are similar when $\widetilde{A}$ is principal. This he did using matrix algebra, including the expansion of $\widetilde{A}^t J \widetilde{A} = nJ$ by the block multiplication of partitioned matrices, to show that $\widetilde{A} = \widetilde{P}^{-1} \widetilde{M} \widetilde{P}$, where $\widetilde{P} = \begin{pmatrix} I & T \\ I & \overline{T} \end{pmatrix}$ and T is a singular parameter system for $\widetilde{A}$ [11, p. 105]. This meant that in order to establish the necessity of conditions (1) and (2), it sufficed to establish them for $\widetilde{M}$ and, given how $\widetilde{M}$ is related to M, it sufficed to show (1) and (2) hold for M.

Further and more complicated matrix algebra enabled Frobenius to establish a result analogous to Laguerre's (3.12) but for M rather than $\widetilde{A}$: If $\widetilde{A}$ is abelian and $T \to T'$ in the sense (3.3) then

$$M \Psi' \overline{M}^t = n\Psi, \tag{3.15}$$

where Ψ, Ψ' are the imaginary parts of T, T', respectively [11, p. 105]. Frobenius was well acquainted with the theory of hermitian forms, where conjugate transpositions such as $\overline{M}^t$ replace ordinary transposition so as to define hermitian symmetry. Although he did not do so, I will use the notation

$$M^h \stackrel{\text{def}}{=} \overline{M}^t$$

to denote the hermitian transpose of a matrix M. Thus (3.15) is a hermitian transpose analog of Laguerre's equation (3.12), and when $T' = T$ so that $\Psi' = \Psi$ it yields an analog of (3.13), namely

$$S^h \Psi S = \Psi, \quad S = n^{-\frac{1}{2}} M^h. \tag{3.16}$$

Reasoning completely analogous to that leading from (3.13) to (3.14)—but with hermitian transposes replacing ordinary transposition—then shows that

$$(QSQ^{-1})^h (QSQ^{-1}) = I \tag{3.17}$$

and so implies that S is similar to $R = QSQ^{-1}$ where R satisfies $R^h R = I$.

Frobenius realized that this is the hermitian conjugate analog of the defining equation for a real orthogonal transformation, and he observed that the proof of his Theorem 2.3 on real orthogonal matrices "can be carried over to the more general systems R considered here without the least change" [11, p. 100]. Over two decades later Frobenius and his student I. Schur named the more general systems R "unitary" (*unitär*) [16, p. 356]—presumably because of the important role they had come to play

in the representation theory of finite groups that Frobenius had begun creating in 1896. Thus Frobenius had established that if R is unitary then all its characteristic roots have absolute value one and its elementary divisors are linear so that R is similar to a diagonal matrix. Thus M and therefore also $\widetilde{M}$ and $\widetilde{A}$ have properties (1) and (2). I should point out that Frobenius deduced these conclusions by applying the reasoning leading from (3.16) to (3.17) to prove a still more general theorem, which he characterized as the "fundamental theorem" underlying his paper: If H is any positive definite hermitian symmetric matrix and S is such that $S^h H S = H$, then all the characteristic roots of S have absolute value one and the elementary divisors of S are all linear [11, p. 100].

We saw that Kronecker's remarks about his Theorem 3.1 tacitly suggested a further problem: characterize the principal $\widetilde{A}$ for which the associated singular parameter system T is unique. Frobenius completely solved this problem as well. Again, ideas in Weber's proof of Theorem 3.4 formed the starting point. Let α denote a characteristic root of the principal abelian matrix $\widetilde{A}$. Then it is a characteristic root of $\widetilde{A}^t$ as well and so $\mathbf{x} \neq \mathbf{0}$ exists such that $\widetilde{A}^t\mathbf{x} = \alpha\mathbf{x}$. Since $\widetilde{A}J\widetilde{A}^t = nJ$ follows from (3.5), we have $nJ\mathbf{x} = (\widetilde{A}J)\widetilde{A}^t\mathbf{x} = \alpha(\widetilde{A}J)\mathbf{x}$, or $\widetilde{A}(J\mathbf{x}) = (n/\alpha)\mathbf{x}$. So much was in effect implicit in Weber's proof. By virtue of property (1) of Frobenius' Theorem 3.5, however, $n/\alpha = \bar{\alpha}$, and so $\widetilde{A}(J\mathbf{x}) = \bar{\alpha}(J\mathbf{x})$. Taking complex conjugates then gives $\widetilde{A}(J\bar{\mathbf{x}}) = \alpha(J\bar{\mathbf{x}})$, since $\widetilde{A}$ and J are real. Summing up: if $\mathbf{x} = \begin{pmatrix}\mathbf{y} & \mathbf{z}\end{pmatrix}^t$ is a characteristic vector for α with respect to $\widetilde{A}^t$, then $\mathbf{x}^* = J\bar{\mathbf{x}} = \begin{pmatrix}\bar{\mathbf{z}} & -\bar{\mathbf{y}}\end{pmatrix}^t$ is a characteristic vector for α with respect to $\widetilde{A}$. Now suppose α is a characteristic root of multiplicity m for $\widetilde{A}$ (and so for $\widetilde{A}^t$ as well). Then by virtue of property (2) of Theorem 3.5 there are m linearly independent characteristic vectors for α as a characteristic root of $\widetilde{A}^t$. If these are denoted by $\mathbf{x}_j = \begin{pmatrix}\mathbf{y}_j & \mathbf{z}_j\end{pmatrix}^t$, $j = 1, \ldots, m$, then $\mathbf{x}_j^* = \begin{pmatrix}\bar{\mathbf{z}_j} & -\bar{\mathbf{y}_j}\end{pmatrix}^t$ are linearly independent characteristic vectors for α with respect to $\widetilde{A}$. Frobenius introduced the matrix $Z_\alpha = (c_{jk})$ defined by the dot product[26] $c_{jk} = (1/2\mathrm{i})(\mathbf{x}_j \cdot \mathbf{x}_k^*) = (1/2\mathrm{i})(\mathbf{y}_j \cdot \bar{\mathbf{z}}_k - \mathbf{z}_j \cdot \bar{\mathbf{y}}_k)$ [11, Sect. 4]. Then Z_α is an $m \times m$ hermitian symmetric matrix of full rank m, and Frobenius discovered that the properties of the Z_α were the key to determining when the singular parameter system T of $\widetilde{A}$ is unique: T is unique precisely when for every characteristic root α of $\widetilde{A}$, Z_α is positive or negative definite [11, p. 114, Satz V].

The reasoning leading to this result did not utilize matrix algebra, but in order to express his uniqueness condition directly in terms of constructs coming from the coefficients of $\widetilde{A}$, Frobenius turned to matrix algebra [11, Sect. 8]. He showed Z_α could be replaced by a $2g \times 2g$ hermitian symmetric matrix $\widetilde{Z}_\alpha$ of rank m (the multiplicity of α) and related to Z_α by $\widetilde{Z}_\alpha = R^h \begin{pmatrix} Z_\alpha & 0 \\ 0 & 0 \end{pmatrix} R$ where the unitary matrix R is chosen so that $\widetilde{Z}_\alpha$ has the following simple symbolical form. Let $\psi(\lambda)$ denote the minimal polynomial of $\widetilde{A}$. Then since (2) of Theorem 3.5 states that the elementary divisors of $\widetilde{A}$ are all linear, Frobenius' Theorem 2.2 on minimal polynomials implies that $\psi(\lambda) = \prod_{j=1}^{d}(\lambda - \alpha_j)$ where $\alpha_1, \ldots, \alpha_d$ denote the distinct characteristic roots of $\widetilde{A}$. Set $\psi_k(\lambda) = \prod_{j \neq k}(\lambda - \alpha_j)$. Then

$$\widetilde{Z}_{\alpha_k} = \mathrm{i}\psi'(\bar{\alpha}_k)\psi_k(\widetilde{A})J, \quad k = 1, \ldots, d. \tag{3.18}$$

[26] Frobenius did not utilize the notion of a dot product. I have used it for succinctness of expression.

The second part of Frobenius' solution to Kronecker's problem may now be stated in the following form.

Theorem 3.6 (Frobenius, 1883) *Let $\widetilde{A}$ be a principal abelian matrix with distinct characteristic roots $\alpha_1, \ldots, \alpha_d$ and let $\widetilde{Z}_{\alpha_k}$ be as in (3.18). Then $\widetilde{Z}_{\alpha_k}$ is hermitian symmetric of rank equal to the multiplicity of α_k, and there is a unique singular parameter system T associated to $\widetilde{A}$ if and only if for all $k = 1, \ldots, d$, $\widetilde{Z}_{\alpha_k}$ is either nonnegative definite or nonpositive definite.*

Part (2) of Weber's Theorem 3.4 is an immediate consequence of the above theorem, since when all the roots of $\widetilde{A}$ are distinct, as Weber assumed, each Z_{α_k} has rank 1 and so has exactly one nonzero root. Thus Z_{α_k} is nonnegative or nonpositive definite, depending on the sign of the nonzero root, and Theorem 3.6 implies that T is unique. Frobenius also gave examples illustrating, respectively, that when a principal $\widetilde{A}$ has multiple roots: (1) the singular parameter T system can still be unique; (2) there can be infinitely many singular parameter systems; (3) all the characteristic roots of $\widetilde{A}$ can be real—and so all equal to $\pm\sqrt{n}$ [11, Sect. 9]. Thus (1) of Weber's Theorem 3.4, which asserts that a principal $\widetilde{A}$ cannot have real characteristic roots when the latter are distinct, is not indicative of the more general situation.

Taken together, Frobenius' Theorems 3.5–3.6 constituted an elegant and definitive solution to Kronecker's problem in its elaborated form. A few years after its publication, Theorem 3.5 was applied by Hurwitz in a paper on Riemann surfaces [28, Sect. 5], a subject closely connected with the theory of theta functions; and in 1903 the theorem was given a detailed exposition in Adolf Krazer's treatise on abelian and theta functions [33, pp. 214–234]. Because Frobenius' results on matrix algebra could not be taken for granted as common knowledge even in 1903, Krazer included an exposition of the basics of matrix algebra and related theorems, such as Frobenius' Theorem 2.3 on orthogonal matrices. In this manner matrix algebra, Frobenius-style, and its mathematical advantages were called to the attention of the many mathematicians of the period with an interest in abelian and theta functions.[27]

[27] Frobenius' Theorems 3.5 and 3.6, however, seem to have little to do mathematically and nothing to do historically with the modern theory of abelian varieties with complex multiplication. That theory was initiated in 1955 in three papers by Weil, Shimura, and Taniyama and drew upon arithmetical developments having their roots in the elliptic case $g = 1$ as developed by Deuring and Hasse. (See Shimura's account [53, pp. ix–x].) These developments involved a stronger formulation of complex multiplication than that suggested by Kronecker's problem. To indicate the difference, it is necessary to regard abelian functions with period matrix $P = \begin{pmatrix} I & T \end{pmatrix}$ as analytic functions on the abelian manifold $\mathcal{T} = \mathbb{C}^g/\Lambda$, where Λ is the lattice in $\mathbb{C}^g$ generated by the columns of the period matrix P. The fact that $\widetilde{A} = mI_{2g}$, $m \in \mathbb{Z}$, is a principal abelian matrix corresponds to the nineteenth-century fact that any abelian function admits ordinary multiplication by m. That is, the mapping $\mathbf{z} \to m\mathbf{z}$ takes the lattice Λ into itself and so induces an endomorphism on $\mathcal{T}$. Thus the integers $\mathbb{Z}$ always embed in the ring of endomorphisms of $\mathcal{T}$. If $\widetilde{A} \neq mI_{2g}$ is principal with singular parameter system T, then $\mathbf{z} \to M\mathbf{z}$, $M = \mathrm{A} + T\Gamma$, maps Λ into itself and so induces on $\mathcal{T}$ a further endomorphism. In the modern theory, however, much more is required of the ring of endomorphisms of $\mathcal{T}$ for it to be an abelian manifold with complex multiplication: the algebraic integers in a "CM field" $\mathbb{K}$ of degree $2g$ over $\mathbb{Q}$ must embed in the endomorphism ring of $\mathcal{T}$ [46, p. 76]. $\mathbb{K}$ is a CM field if it is a totally imaginary quadratic extension of a totally real field. An example of a $g = 1$ dimensional $\mathcal{T}$ with complex multiplication is given by the example corresponding to the principal abelian matrix $\widetilde{A}_1$ discussed at the end of Sect. 3.2. The construction of abelian manifolds with complex multiplication in the modern sense, however, makes no use of principal abelian matrices [46, p. 14ff.].

4 The congruent transformation problem

The congruent transformation problem (defined further on) was first posed and resolved by Frobenius in a paper published in 1896 [14], but it has its origins in the work of Weierstrass and Kronecker on the transformation of pencils of quadratic and bilinear forms. In order to appreciate the problem and its solution via matrix algebra by Frobenius, it is necessary to consider first the work that motivated it.

4.1 Origins of the problem

In his paper of 1868 on the theory of elementary divisors [60] Weierstrass was primarily concerned with pencils of bilinear forms and the problem of determining a complete set of invariants associated to a such pencil $\beta(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t(\lambda A - B)\mathbf{y}$ when it is nonsingular, i.e., when $\det A \neq 0$. Theorem 2.1 summarizes his main result: the elementary divisors of the pencil form a complete set of invariants. Thus if two such pencils $\beta(\mathbf{x}, \mathbf{y})$ and $\gamma(\mathbf{X}, \mathbf{Y})$ have the same elementary divisors then nonsingular transformations $\mathbf{x} = P\mathbf{X}$ and $\mathbf{y} = Q\mathbf{Y}$ exist so that $\beta(\mathbf{x}, \mathbf{y}) = \gamma(\mathbf{X}, \mathbf{Y})$, i.e., so that β can be transformed into γ and vice versa. In this case β and γ are said to be *equivalent*. Expressed in terms of matrix algebra, equivalence means that $P^t(\lambda A - B)Q = \lambda C - D$. With this in mind, in his 1877 paper on matrix algebra [8, p. 361] Frobenius defined any two $n \times n$ matrices A and B to be equivalent if

$$RAQ = B \quad \text{where } R, Q \text{ are nonsingular.}^{28} \tag{4.1}$$

In what follows I will use the notation

$$A \sim B \tag{4.2}$$

to indicate that A and B are equivalent in the sense of (4.1). Thus for nonsingular pencils, Weierstrass' theory says that $\lambda A - B \sim \lambda C - D$ if and only if the two pencils have the same elementary divisors.

Weierstrass also wanted to apply his theory to nonsingular pencils of quadratic forms, so as to generalize his Theorem 3.3 of 1858. If $\lambda A - B$ and $\lambda C - D$ are two such pencils, so that A, B, C, D are symmetric, and if they have the same elementary divisors then of course by Weierstrass' Theorem 2.1 $\lambda A - B \sim \lambda C - D$, but this does not imply the one pencil can be transformed into the other in the usual sense of that term as it applies to quadratic forms. That is, these forms depend on a single variable set $\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^t$, so $\beta(\mathbf{x}) = \mathbf{x}^t(\lambda A - B)\mathbf{x}$ and $\gamma(\mathbf{X}) = \mathbf{X}^t(\lambda C - D)\mathbf{X}$ and the question is whether a nonsingular transformation $\mathbf{x} = R\mathbf{X}$ exists that transforms β into γ in the sense that $\beta(\mathbf{x}) = \gamma(\mathbf{X})$ for $\mathbf{x} = R\mathbf{X}$. This means that $R^t(\lambda A - B)R = \lambda C - D$. In this case Kronecker spoke of the congruent transformation of β into γ [40, p. 424]. With this in mind, in his paper of 1877 [8, p. 364] Frobenius defined two $n \times n$ matrices

[28] Here R is playing the role of P^t above.

A and B to be congruent when

$$R^t AR = B \quad \text{for some nonsingular } R. \tag{4.3}$$

In what follows I will use the notation

$$A \cong B \tag{4.4}$$

to indicate that A and B are congruent. Thus the problem facing Weierstrass in applying his theory to nonsingular pencils of quadratic forms $\mathbf{x}^t(\lambda A - B)\mathbf{x}$ and $\mathbf{x}^t(\lambda C - D)\mathbf{x}$ was to show that $\lambda A - B \cong \lambda C - D$ when they have the same elementary divisors.

The transformations $\mathbf{x} = P\mathbf{X}$, $\mathbf{y} = Q\mathbf{Y}$ that Weierstrass used to transform $\mathbf{x}^t(\lambda A - B)\mathbf{x}$ into its canonical form were defined using determinant-theoretic constructs, and Weierstrass showed from the form of these transformations and the manner in which they depended upon the coefficients of A and B that when A and B were symmetric then $P = Q$, so that $P^t(\lambda A - B)Q = \lambda I - J$ becomes $P^t(\lambda A - B)P = \lambda I - J$. Thus $\lambda A - B \cong \lambda I - J$ and since $\lambda B - D \cong \lambda I - J$ also, it followed that $\lambda A - B \cong \lambda C - D$. Implicit in these deliberations was the corollary that for nonsingular families of quadratic forms equivalence implies congruence. It turned out, however, that Weierstrass' proof was flawed. It had depended on certain results from the previous section of his paper [60, Sect. 4] that turned out to be far more difficult to justify than Weierstrass had realized when he published it.[29] Various determinant-theoretic ways around the problem were proposed by Frobenius' friend and frequent collaborator, Ludwig Stickelberger, by Kronecker, and by Frobenius himself.[30] Apparently Weierstrass was not pleased with any of these approaches, since with the preparation of a version of his 1868 paper for inclusion in his collected works in mind, he asked Frobenius to look into the matter anew. Frobenius consented and in a paper of 1894 [13] resolved the problem in a way that pleased Weierstrass enough that, in the his collected works, he omitted the problematic section 4 of his paper [60] and referred the reader instead to Frobenius' paper [13].[31] Thus for nonsingular pencils of quadratic forms the theorem that equivalence implies congruence had been rigorously established but only by means of what Frobenius himself described as "subtle deliberations" involving determinants [14, p. 697].

As indicated at the end of Sect. 3.3, it was Kronecker who proposed to himself the problem of extending Weierstrass' theory to all pencils of forms. When a pencil $\mathbf{x}^t(\lambda A - B)\mathbf{y}$ is singular (det $A = 0$) it can happen that $\varphi(\lambda) = \det(\lambda A - B)$ is identically zero, so that Weierstrass' definition of elementary divisors does not directly apply and the problem arises as to how to define a complete set of invariants to characterize the equivalence classes of such pencils. After some preliminary attacks on the problem in 1868, Kronecker discovered a viable approach for carrying out his agenda in 1874. To any pencil of quadratic or bilinear forms $\lambda A - B$ he showed how to associate a set

[29] This according to Frobenius [13, p. 577].

[30] See [13, pp. 577–578] for a detailed summary of these efforts.

[31] See Weierstrass's *Werke* 2 (1895), p. 31.

$\mathcal{K}_{(\lambda A-B)}$ of invariants.[32] He convinced himself that his invariants formed a complete set, i.e., that two pencils of bilinear forms $\lambda A - B$ and $\lambda C - D$ are equivalent if and only if $\mathcal{K}_{(\lambda A-B)} = \mathcal{K}_{(\lambda C-D)}$ and that two pencils of quadratic forms are congruent if and only if $\mathcal{K}_{(\lambda A-B)} = \mathcal{K}_{(\lambda C-D)}$. Bits and pieces of Kronecker's theory were indicated in several papers in the proceedings of the Berlin Academy in 1874 [37–40] but it was only towards the end of 1874 that he wrote up a coherent development of his theory for pencils of bilinear forms. He finally published it in 1890 [42], and then during 1890–1891 he published his theory for pencils of quadratic forms [43,44]. The quadratic theory was developed ab initio; Kronecker did not attempt to derive it from the bilinear theory as had been Weierstrass' strategy in 1868. From Kronecker's papers of 1890–1891 it now followed that if any two pencils of quadratic forms, singular or not, are equivalent, then they are congruent; for if $\lambda A - B \sim \lambda C - D$ then $\mathcal{K}_{(\lambda A-B)} = \mathcal{K}_{(\lambda C-D)}$ by virtue of the bilinear theory developed in [42], whereas $\mathcal{K}_{(\lambda A-B)} = \mathcal{K}_{(\lambda C-D)}$ implied $\lambda A - B \cong \lambda C - D$ by virtue of the quadratic theory developed in [43,44].

It should be noted that if it could be proved independently of Kronecker's theories that, for quadratic pencils, equivalence implies congruence, then Kronecker's entire quadratic theory could be dispensed with in the sense that it would now be an immediate consequence of the bilinear theory. That is, since congruence obviously implies equivalence we would have for quadratic pencils $\mathcal{K}_{(\lambda A-B)} = \mathcal{K}_{(\lambda C-D)}$ if and only if $\lambda A - B \sim \lambda C - D$ by Kronecker's bilinear theory. Then if it were proved that equivalence implies congruence for quadratic pencils, the main theorem of the quadratic theory follows: $\mathcal{K}_{(\lambda A-B)} = \mathcal{K}_{(\lambda C-D)}$ if and only if $\lambda A - B \cong \lambda C - D$.

One theory that Kronecker did fully develop and publish in 1874 had to do with the congruence of the special pencils of bilinear forms with coefficient matrices $\lambda A - A^t$—the type that had arisen in connection with the complex multiplication problem of Sect. 3.3. In 1866 when Kronecker considered that problem he had been content with a generic result, namely Theorem 3.2. In 1874 he was in a position to deal with it on a nongeneric level and without assuming the pencil nonsingular or the number of variables to be even, as in Theorem 3.2. Thus in the course of sixty pages he showed that for any A and B, $\lambda A - A^t \cong \lambda B - B^t$ if and only if $\mathcal{K}_{(\lambda A-A^t)} = \mathcal{K}_{(\lambda B-B^t)}$ [40]. Of course, in view of the bilinear theory Kronecker published in 1890 [42], the above result would follow as an immediate consequence of the main theorem of [42] *if* it could be shown directly that for pencils of bilinear forms of the special type $\mathbf{x}^t(\lambda A - A^t)\mathbf{y}$, equivalence implies congruence. That is, by virtue of Kronecker's general bilinear theory as published in 1890, $\mathcal{K}_{(\lambda A-A^t)} = \mathcal{K}_{(\lambda B-B^t)}$ if and only if $\lambda A - A^t \sim \lambda B - B^t$ if and only if $\lambda A - A^t \cong \lambda B - B^t$, *provided* it could be proved directly that equivalence implies congruence for this type of pencil.

This then is the background to the *congruent transformation problem* that Frobenius posed to himself: Show directly that equivalence implies congruence for (a) pencils

[32] Kronecker actually considered two different sets. The first one he considered is easiest to understand and is the one adopted by subsequent mathematicians interested in singular pencils. It consists of n homogeneous polynomials $D_i(\lambda)$, $i = 0, \ldots, n-1$, where $D_i(\lambda)$ is the polynomial greatest common divisor of the $n - i$ by $n - i$ minors of $D_0(\lambda) = \det(\lambda A - B)$, together with two series of integers associated to the null spaces of $\lambda A - B$ and its transpose. For a definition of the integer sequences see [17, v. 2, p. 37ff.], where they are called the minimal indices of $\lambda A - B$.

of quadratic forms and (b) pencils of bilinear forms of the type $\mathbf{x}^t(\lambda A - A^t)\mathbf{y}$. As we shall see, he solved it in a matter of a few pages by means of matrix algebra in a paper of 1896. As Frobenius observed in the introduction [14, p. 697]:

> The extremely simple argument presented here provides a complete replacement for the lengthy analysis that Kronecker employed ... [in the above-mentioned papers [40,43,44]] ..., and also with its help the subtle deliberations in the work of Weierstrass that are required for a precise treatment of ... quadratic ... forms can be avoided.

4.2 Frobenius' solution

In order to suggest the line of reasoning that surely led Frobenius to realize how the congruent transformation problem could be readily solved, consider the problem for pencils of quadratic forms, the case of primary importance. Clearly it will be solved provided it can be shown that for any two symmetric matrices A and B if $A \sim B$ then $A \cong B$. That is, A and B can each represent a pencil of quadratic forms so the problem is solved in case (a). So suppose that A and B are symmetric and $A \sim B$. Then nonsingular P and Q exist such that

$$PAQ = B. \tag{4.5}$$

Taking transposes of both sides and using the symmetry of A and B we get

$$Q^t A P^t = B. \tag{4.6}$$

Frobenius discovered that equations (4.5) and (4.6) together provide the key to the solution of the congruent transformation problem in both cases (a) and (b).

Eliminating B from these two equations we have $PAQ = Q^t A P^t$, which may be rewritten as

$$(Q^t)^{-1} PA = AP^t Q^{-1}. \tag{4.7}$$

Thus if we set $U = (Q^t)^{-1} P$, then $U^t = P^t Q^{-1}$ and so (4.7) becomes

$$UA = AU^t, \quad U = (Q^t)^{-1} P. \tag{4.8}$$

It then follows from (4.8) that $U^k A = A(U^t)^k$ for any positive integer k and therefore that for any polynomial $\chi(t) \in \mathbb{C}[t]$

$$\chi(U)A = A\chi(U^t). \tag{4.9}$$

If $\chi(t)$ is such that $\det[\chi(U)] \neq 0$, then $\chi(U^t) = [\chi(U)]^t$ is invertible and (4.9) can be written as

$$\chi(U)A[\chi(U^t)]^{-1} = A, \quad U = (Q^t)^{-1} P. \tag{4.10}$$

If we now go back to (4.6) and express A by the left-hand side of (4.10) we get

$$B = Q^t \chi(U) A [\chi(U^t)^{-1}] P^t = RAS, \tag{4.11}$$

with R and S defined by

$$R = Q^t \chi(U) \quad \text{and} \quad S = [\chi(U^t)]^{-1} P^t, \tag{4.12}$$

where $\chi(t)$ is any polynomial with the property that $\det[\chi(U)] \neq 0$ for $U = (Q^t)^{-1} P$.

In effect, (4.11)–(4.12) gives an infinite number of equivalence transformations R, S taking A into B. The question is whether it is possible to choose $\chi(t)$ so that $R = S^t$, for then (4.11) asserts that $A \cong B$. From the expressions in (4.12) for R and S, it follows that the condition that $R = S^t$ can be expressed in the form $Q^t \chi(U) = P[\chi(U)]^{-1}$ or as

$$[\chi(U)]^2 = U, \quad U = (Q^t)^{-1} P. \tag{4.13}$$

It was by means of the above sort of matrix algebraic reasoning (given here essentially as Frobenius' presented it in his paper [14, Sect. 2] that he surely first realized that part (a) of the congruent transformation problem would be solved if he could prove the following square root theorem:

Theorem 4.1 (Frobenius, 1896) *If U is any square matrix with* $\det U \neq 0$ *and if m is the degree of its minimal polynomial, then a polynomial $\chi(z)$ of degree $m-1$ exists such that* $[\chi(U)]^2 = U$.

The relation $[\chi(U)]^2 = U$ of course implies that $\det[\chi(U)] \neq 0$ and so application of this theorem to $U = (Q^t)^{-1} P$ then implies by the above reasoning that S as given in (4.12) satisfies $S^t AS = B$. Thus for A, B symmetric $A \sim B$ does indeed imply $A \cong B$.

Frobenius' proof of Theorem 4.1 will be discussed below in Section 4.4. First, however, it should be observed, as Frobenius did, that the reasoning leading to (4.13) derived entirely from (4.5) and (4.6), so that for any A and B satisfying these two equations, symmetric or not, the same reasoning as given above would lead to (4.13), so that Theorem 4.1 implies S as defined in (4.12) satisfies $S^t AS = B$. With that in mind consider part (b) of the congruent transformation problem: show that $\lambda A - A^t \sim \lambda B - B^t$ implies $\lambda A - A^t \cong \lambda B - B^t$. Suppose $\lambda A - A^t \sim \lambda B - B^t$, i.e., that $P(\lambda A - A^t)Q = \lambda B - B^t$. Evidently the only way this can hold for all λ is if $PAQ = B$ and $PA^tQ = B^t$. The former equality is (4.5) and the latter, after transposition, is (4.6). Thus for this A and B we have (4.13) and so by Theorem 4.1 we may conclude that a nonsingular S exists for which $S^t AS = B$. Transposition of this equality yields $S^t A^t S = B^t$ from which $S^t(\lambda A - A^t)S = \lambda B - B^t$ follows immediately. Thus $\lambda A - A^t \sim \lambda B - B^t$ does indeed imply $\lambda A - A^t \cong \lambda B - B^t$ and part (b) is also established.

4.3 Cayley, Sylvester, and matrix square roots

Prior to Frobenius' work on the congruence problem, both Cayley and Sylvester had considered the matter of matrix square roots. Although it is uncertain whether Frobenius was familiar with what they had to say, their remarks provide some historical and mathematical perspective on his Theorem 4.1 and its proof.

Cayley had already considered the idea of a square root of a matrix in his paper of 1858 on matrix algebra [2, p. 483ff.]. This occurred in his discussion of the implications of the Hamilton–Cayley Theorem that $\varphi(A) = 0$ where $\varphi(A) = \det(\lambda I - A)$. It follows immediately from this theorem, Cayley observed, that if $f(t)$ is any polynomial or rational function and M is an $n \times n$ matrix, then $L = f(M)$ is expressible as a polynomial of degree at most $n - 1$.[33] "But it is important to consider," Cayley continued, "how far or in what sense the like theorem is true with respect to irrational functions of a matrix" [2, p. 383]. By "irrational functions" of a matrix M Cayley had in mind expressions such as $L = \sqrt{M}$, which he considered by way of example. In this case, if M is $n \times n$ then $L = \sqrt{M}$ exists precisely when the system of n^2 quadratic equations in the n^2 unknown coefficients of L that corresponds to $L^2 = M$ has a solution. Cayley focused on how to determine L when it exists, presumably with an eye towards determining "how far or in what sense" L is expressible as a polynomial in M. To this end he showed how the Hamilton–Cayley Theorem could be used to facilitate finding L.

Cayley's method applies to matrices of any size, but presumably to avoid complicated notation he illustrated it in the case $n = 2$. For the purposes at hand, I will illustrate it for 3×3 matrices. If M is 3×3 and $L^2 = M$, then the Hamilton–Cayley Theorem implies that constants a, b, c exist so that $L^3 + aL^2 + bL + cI = 0$. Since $L^2 = M$, $L^3 = ML = LM$ and so the equation for L becomes $LM + aM + bL + cI = 0$ or $L(M + bI) = -(aM + cI)$. Squaring both sides of this last equation and substituting M for L^2 we obtain $M(M + bI)^2 = (aM + cI)^2$. This matrix equation corresponds to $n^2 = 9$ quadratic equations in the $n = 3$ unknowns a, b, c rather than the $n^2 = 9$ unknown coefficients of L in the $n^2 = 9$ equations implied by $L^2 = M$. This reduction of unknowns is the point of Cayley's method.

Cayley's method, however, was ineffective in dealing with the question he had posed, namely "how far and to what extent" is it the case that irrational functions such as $\sqrt{M}$ can be expressed as polynomials in M. For example if

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then since $M^2 = 0$, the equation $M(M + bI)^2 = (aM + cI)^2$ derived above reduces to $c^2 I = (b^2 - 2ac)M$ and implies $b = c = 0$ with a left undetermined. Thus $L = (aM + cI)/(M + bI) = aM/M$, but since M is not invertible the method loses its meaning and might seem to suggest that L does not exist, although this turns out

[33] This is indeed correct, as Frobenius showed in his 1877 paper [8, p. 355], assuming that when $f(t) = p(t)/q(t)$, $\det q(A) \neq 0$ so that $[q(A)]^{-1}$ exists.

to be incorrect. M does have square roots, there are solutions to the system of n^2 equations in n^2 unknowns symbolized by $L^2 = M$, and they are given by

$$L(\alpha, \beta) = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & \beta^{-1} & 0 \end{pmatrix},$$

where α and β are arbitrary parameters with $\beta \neq 0$ [17, vol. 1, p. 239]. Since the minimal polynomial of M is $\psi(t) = t^2$, and so of degree 2, it follows that if $L(\alpha, \beta)$ were expressible as a polynomial in M, then it would be expressible as a polynomial of degree 1. However, it is easily seen that $L(\alpha, \beta) = pM + qI$ is impossible since it implies $\beta = 0$. This shows that $L = \sqrt{M}$ cannot always be expressed as a polynomial in M. The example $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ shows that $L = \sqrt{M}$ need not exist,[34] In the two above examples of exceptional M, $\det M = 0$. Frobenius' Theorem 4.1 shows that when $\det M \neq 0$ M always has a square root and it is expressible as a polynomial in M.

During the period 1882–1884, Cayley's friend J. J. Sylvester became interested in matrix algebra.[35] Among other things, he considered Cayley's question concerning the determination of irrational functions of a matrix M such as $\sqrt{M}$. Sylvester's general conclusion was that if $f(z)$ is any single- or multiple-valued function of z and if M is $n \times n$ with distinct characteristic roots $\lambda_1, \ldots, \lambda_n$, then $f(M)$ is given by

$$f(M) = \sum_{i=1}^{n} \prod_{j \neq i} \frac{(M - \lambda_j I)}{(\lambda_i - \lambda_j)} f(\lambda_i). \tag{4.14}$$

Formula (4.14) is based on the idea behind the Lagrange interpolation formula and is sometimes called Sylvester's interpolation formula. As Sylvester realized, it applies to the case in which the characteristic roots λ_i are all distinct. When some λ_i are equal "the formula must be replaced by another obtained from it by the usual method of infinitesimal variation" [54, p. 111]. Perhaps Sylvester could have done this in a specific case, e.g., when M has one double root, but no less an algebraist than Lagrange had been led to false conclusions using "the usual method of infinitesimal variation" because he continued to reason generically.[36] Indeed, when $f(z) = \sqrt{z}$ the examples given above indicate that $f(M)$ need not exist or may exist but not be derivable by infinitesimal considerations applied to (4.14), which would lead to an expression for $f(M)$ as a polynomial in M.

[34] If $L^2 = M$, then $L^4 = M^2 = 0$. This means the characteristic roots of L must all be 0 and so $\varphi(t) = t^2$ is the characteristic polynomial of L. The Hamilton–Cayley Theorem then implies $L^2 = 0 \neq M$, and so $\sqrt{M}$ does not exist.

[35] For the personal and institutional background to Sylvester's brief flurry of interest in matrix algebra see [50, pp. 135–138]. A fairly detailed mathematical discussion of Sylvester's work on matrix algebra is given in [22, Sect. 6].

[36] I refer to Lagrange's attempt to extend his elegant generic solution to $\ddot{\mathbf{y}} + A\mathbf{y} = \mathbf{0}, \mathbf{y}(0) = \underline{y}_0$, A $n \times n$, to the case in which $f(\rho) = \det(\rho^2 + A)$ has one root of multiplicity two. See [21] for details and references.

4.4 Frobenius' proof of his square root theorem

Frobenius, by contrast with Sylvester (who is not mentioned by Frobenius) provided a proof of Theorem 4.1 that is nongeneric, yet simple, and completely rigorous by present-day standards. For pedagogical reasons it would probably not be used today because it invokes results from complex analysis, but for that reason it provides another good example of how in forging the link between formal matrix algebra and Weierstrass' theory of elementary divisors, Frobenius followed the example of his mentor and employed considerations drawn from Weierstrass-style complex analysis. What follows is an exposition of the resulting proof.

The goal of Theorem 4.1 is to determine a polynomial $\chi(z)$ such that $[\chi(U)]^2 = U$, where U is a given matrix with $\det U \neq 0$. Frobenius turned to the minimal polynomial $\psi(z)$ of U. Since $\psi(U) = 0$ by definition, it suffices to define $\chi(z)$ in such a way that $\psi(z)$ divides $[\chi(z)]^2 - z$, for then $[\chi(U)]^2 - U = 0$, i.e., $V^2 = U$ for $V = \chi(U)$. The idea is to utilize the identity

$$[\chi(z)]^2 - z = \left[\chi(z) - \sqrt{z}\right]\left[\chi(z) + \sqrt{z}\right], \tag{4.15}$$

which is valid for any determination of $\sqrt{z}$, $z \neq 0$, to show, in a sense to be explained below, that $\psi(z)$ divides the first factor and hence divides the left-hand side.

Let

$$\psi(z) = \prod_{j=1}^{d} (z - \lambda_i)^{e_j} \tag{4.16}$$

denote the factorization of the minimal polynomial, so that $\lambda_1, \ldots, \lambda_d$ are the distinct characteristic roots of U. Since $\det U \neq 0$, each root $\lambda_j \neq 0$, and so a branch $f_j(z)$ of $\sqrt{z}$ exists in a sufficiently small neighborhood of $z = \lambda_j$. Thus $f_j(z)$ is analytic in this neighborhood and $f_j(\lambda_j) = \sqrt{\lambda_j} \neq 0$ since $\lambda_j \neq 0$, which means that $f_j(z)/\psi(z)$ has a pole of order e_j at λ_j, and so we may write

$$\frac{f_j(z)}{\psi(z)} = \frac{a_{-e_j}}{(z - \lambda_j)^{e_j}} + \cdots + \frac{a_{-1}}{z - \lambda_j} + \mathfrak{P}_j(z - \lambda_j), \tag{4.17}$$

where $\mathfrak{P}_j(z - \lambda_j)$ is my notation for a power series in $z - \lambda_j$. The singular part of the Laurent expansion (4.17), can be expressed as a simple fraction

$$\frac{a_{-e_j}}{(z - \lambda_j)^{e_j}} + \cdots + \frac{a_{-1}}{z - \lambda_j} = \frac{A_j(z)}{(z - \lambda_j)^{e_j}}, \tag{4.18}$$

where $A_j(z)$ is a polynomial in z.

With these preliminaries in place, define $\chi(z)$ by

$$\chi(z) = \sum_{j=1}^{d} \frac{A_j(z)\psi(z)}{(z - \lambda_j)^{e_j}}. \tag{4.19}$$

Thus $\chi(z)$ is $\psi(z)$ times the sum of the singular parts of $f_j(z)/\psi(z)$ at each pole $z = \lambda_j$, $j = 1, \ldots, d$. It is easy to see that $\chi(z)$ is a polynomial since in the jth term of (4.19) the factor $(z - \lambda_j)^{e_j}$ divides $\psi(z)$.[37]

Now consider for any fixed k the difference $\chi(z) - \sqrt{z} = \chi(z) - f_k(z)$. From (4.17)–(4.19) it follows that

$$\chi(z) - f_k(z) = \sum_{j \neq k} \frac{A_j(z)\psi(z)}{(z - \lambda_j)^{e_j}} - \psi(z)\mathfrak{P}_k(z - \lambda_k). \tag{4.20}$$

The presence of $\psi(z)$ as a factor in every term above means that $(z - \lambda_k)^{e_k}$ can be factored from every term, so that (4.20) may be written as

$$\chi(z) - f_k(z) = (z - \lambda_k)^{e_k}\mathfrak{P}_k^*(z - \lambda_k).$$

Since it is clear that $\chi(z) + \sqrt{z} = \chi(z) + f_k(z) = \mathfrak{P}_k^{**}(z - \lambda_k)$, the identity (4.15) becomes

$$[\chi(z)]^2 - z = (z - \lambda_k)^{e_k}\mathfrak{P}_k^*(z - \lambda_k)\mathfrak{P}_k^{**}(z - \lambda_k) = (z - \lambda_k)^{e_k}\mathfrak{P}_k^{***}(z - \lambda_k),$$

which shows that, for any fixed k, $(z - \lambda_k)^{e_k}$ divides $[\chi(z)]^2 - z$; and so $\psi(z)$ divides $[\chi(z)]^2 - z$. This complete's Frobenius' proof of Theorem 4.1.[38]

5 The spread of Frobenius' ideas

By Frobenius' ideas, I mean not simply the idea of a symbolical algebra of matrices but also, and more importantly, the idea that such an algebra can be developed with profit in conjunction with the theory of canonical matrix forms of Weierstrass, Jordan, and Kronecker. The idea of matrix algebra had been considered independently by several mathematicians besides Frobenius—Cayley, Laguerre, and Sylvester—but he alone pursued the second idea as exemplified by his solution to Kronecker's complex multiplication problem and the congruence problem. By 1896, when he published his solution to the congruence problem, however, his work on matrix algebra was still not widely known or appreciated.

For example in 1887 Lipschitz published a paper [47] that was prompted by a passing remark that Camille Jordan had made in a lengthy paper on linear differential equations [32, p. 112, no. 36], which appeared in Crelle's *Journal* 26 pages after Frobenius' 1877 paper on matrix algebra [8]. Jordan observed without proof that if $x_i' = \sum_{i,j=1}^{n} a_{ij}x_j$ is a linear substitution S that belongs to a group of finite order

[37] It also follows that at $z = \lambda_j$ χ and its derivatives up to the $(e_j - 1)$st agree with those of f_j. That is because $\chi - f_j = (z - \lambda_j)^{e_j}\mathfrak{P}(z - \lambda_j)$ near λ_j. It also follows that deg $\chi = m - 1$ because in (4.18) $a_{-1} = f^{(j-1)}(\lambda_j)/(j-1)! \neq 0$ and so deg $A_j = e_j - 1$. Thus each term of χ in (4.19) has degree $m - 1$.

[38] The reason that χ has degree $m-1$ is indicated in the previous footnote. Frobenius' proof [14, p. 697ff.] is expressed somewhat more generally than expounded here so as to allow a brief discussion of the problematic case det $U = 0$ as well as other functions of a matrix.

then S has a diagonal canonical form with roots of unity along the diagonal. Lipschitz realized that Jordan's remark implied that if any linear substitution S composed with itself k times gives the identical substitution, then it has a diagonal form with k roots of unity along the diagonal. Being well-versed in Weierstrass' theory of elementary divisors, Lipschitz devoted his paper to a proof of the elementary divisor analog: if S has the above property then all its characteristic roots are kth roots of unity and all its elementary divisors are linear. He failed to realize that his proposition was a special case of a more general theorem already proved by Frobenius a few pages earlier in the same issue of Crelle's *Journal* that contained Jordan's paper.[39] Lipschitz' proposition is also an easy consequence of Frobenius' Theorem 2.2 on the minimal polynomial $\psi(t)$ of S (Sect. 2.1) since $S^k = I$ implies by that theorem that $\psi(t)$ divides $f(t) = t^k - 1$. Thus all the roots of $\psi(t)$ are (1) distinct and are (2) kth roots of unity. By part (iv) of Theorem 2.2 (1) implies the elementary divisors of A are all linear and by part (iii) of Theorem 2.2 (2) implies all the characteristic roots of A are kth roots of unity. Lipschitz had clearly overlooked Frobenius' paper! Worse yet, Kronecker responded to Lipschitz' paper by suggesting in a paper of 1890 [41] how Lipschitz' theorem could be deduced (nontrivially) by means of considerations similar to some he had published earlier that year for orthogonal systems. As Frobenius said in 1896 regarding his minimal polynomial theorem, "So far little attention has been paid to this consequential theorem" [15, p. 711]. Not only, he continued, were Lipschitz and Kronecker unfamiliar with the theorem and with his 1877 paper containing it but also almost all "the English and American algebraists, who have concerned themselves considerably with the theory of matrices" [15, p. 712]. Ironically, the activity in the realm of matrix algebra to which Frobenius referred seems to have been prompted by the need to develop Cayley's and Sylvester's ideas more generally and rigorously, whereas this had already been done by Frobenius.[40]

After 1896, however, Frobenius' accomplishments in the realm of matrix algebra and its applications became better known. As already mentioned, his solution to Kronecker's complex multiplication problem, together with the attendant results on matrix algebra, were highlighted by Adolf Krazer in his treatise on the theory of theta functions, which appeared in 1903 [33, Chap. 6]. Even earlier, in 1899, Peter Muth (1860–1909), who had been one of Moritz Pasch's doctoral students at the University of Giessen, published the first book devoted to a systematic exposition of the theory of quadratic and bilinear forms that had been created by Weierstrass, Kronecker and Frobenius.[41] Entitled *Theory and Application of Elementary Divisors* [49], Muth's book made matrix algebra, which is developed along the lines set out in Frobenius' 1877 paper [8], central to the theory.[42] The basics were expounded in the second

[39] Frobenius proved that if S is a matrix with the property that the sequence S^j, $j = 0, 1, 2, \ldots$ has only a finite number of distinct terms, then all characteristic roots of S are either 0 or roots of unity and the elementary divisors corresponding to the roots of unity are all linear [8, Satz VI, p. 357].

[40] See in this connection [22, p. 107n.15].

[41] Regarding Muth's life and work see [51].

[42] In the preface Muth wrote that he had been encouraged by the fact that "From the outset my undertaking was of special interest to several outstanding experts in the theory of elementary divisors," namely Frobenius, S. Gundelfinger, and K. Hensel (Kronecker's former student) [49, p. iv].

chapter, which closely followed the development as given in Frobenius' 1877 paper, and was then used throughout the book. In particular, Muth stressed the importance of Frobenius' "simple, elegant" solution to the congruence problem [49, pp. xii–xiii, p. 125] and used it to derive Kronecker's theory of singular pencils of quadratic forms as an easy consequence of his theory for pencils of bilinear forms [49, pp. 125–128] as well as Kronecker's congruence theory of the special pencils $\mathbf{x}^t(\lambda A + A^t)\mathbf{y}$ [49, pp. 142–143]. Muth's book became a standard reference, and Frobenius' matrix algebraic, square root technique became the standard one for deducing the theory of the congruence of pencils of quadratic forms from the equivalence theory of pencils of bilinear forms.[43]

Also in 1896 Frobenius began creating the theory of characters and representations of finite groups, which attracted the attention of many mathematicians.[44] One of the principal tools in developing representation theory was linear algebra, and so at the hands of Frobenius and his brilliant student Issai Schur, Frobenius' special brand of linear algebra became a familiar aspect of the theory.

Thanks largely to Frobenius' work, matrix algebra became an additional tool for dealing with problems of a linear algebraic nature. In the case of the congruence problem, matrix algebra rather than the lengthy determinant-based considerations of the papers of Weierstrass and Kronecker had, as Frobenius said, revealed the "proper reason" (*eigentlicher Grund*) as to why equivalence implies congruence for the pencils they had considered. This was one of the ways in which Frobenius contributed to the decline of the theory of determinants as a principal tool of linear algebra. Another way that was more far-reaching resulted from his "rational" development of elementary divisor theory in 1879 [9], which was based upon arithmetical considerations rather than determinants; but that is another story altogether.

References

1. M. Bôcher. *Introduction to Higher Algebra*. Macmillan, New York, 1907. Republished by Dover Publications, New York, 1964. German translation by H. Beck, Leipzig, 1910.
2. A. Cayley. A memoir on the theory of matrices. *Phil. Trans. R. Soc. London*, 148:17–37, 1858. Reprinted in *Papers* 2, 475–496.
3. E. B. Christoffel. Verallgemeinerung einiger Theoreme des Herrn Weierstrass. *Jl. für die reine u. angew. Math.*, 63:255–272, 1864. Reprinted in *Abhandlungen* 1, pp. 129–145.
4. A. Clebsch. Theorie der circularpolarisirenden Medien. *Jl. für die reine u. angew. Math.*, 57:319–358, 1860.
5. A. Clebsch. Über eine Classe von Gleichungen, welche nur reelle Wurzeln besitzen. *Jl. für die reine u. angew. Math.*, 62:232–245, 1863.
6. C. W. Curtis. *Pioneers of Representation Theory: Frobenius, Burnside, Schur and Brauer*. American Mathematical Society, 1999.
7. L. E. Dickson. *Modern Algebraic Theories*. Sandborn, Chicago, 1926.
8. G. Frobenius. Über lineare Substitutionen und bilineare Formen. *Jl. für die reine u. angew. Math.*, 84:1–63, 1877. Reprinted in *Abhandlungen* 1, 343-405.
9. G. Frobenius. Theorie der linearen Formen mit ganzen Coefficienten. *Jl. für die reine u. angew. Math.*, 86:146–208, 1879. Reprinted in *Abhandlungen* 1, 482–544.

[43] See, e.g., [1, pp. 297–301], [7, pp. 120–125], [55, pp. 130–131], [48, pp. 60–61], [17, v. 2, 41–42].

[44] On this part of Frobenius' mathematics and its background, see [19,20,6].

10. G. Frobenius. Zur Theorie der Transformation der Thetafunctionen. *Jl. für die reine u. angew. Math.*, 89:40–46, 1880. Reprinted in *Abhandlungen* 2, 1–7.
11. G. Frobenius. Über die principale Transformation der Thetafunctionen mehrerer Variableln. *Jl. für die reine u. angew. Math.*, 95:264–296, 1883. Reprinted in *Abhandlungen* 2, 97–129.
12. G. Frobenius. Gedächtnisrede auf Leopold Kronecker. *Abhandlungen d. Akad. der Wiss. zu Berlin*, pages 3–22, 1893. Reprinted in *Abhandlungen* 3, 707–724.
13. G. Frobenius. Über die Elementartheiler der Determinanten. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 7–20, 1894. Reprinted in *Abhandlungen* 2, 577–590.
14. G. Frobenius. Über die cogredienten Transformationen der bilinearen Formen. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 7–16, 1896. Reprinted in *Abhandlungen* 2, 695–704.
15. G. Frobenius. Über vertauschbare Matrizen. *S'ber. Akad. der Wiss. Berlin, 1896*, 2:601–614, 1896. Reprinted in *Abhandlungen* 2, 705-718.
16. G. Frobenius and I. Schur. Über die reellen Darstellungen der endlichen Gruppen. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 186–208, 1906. Reprinted in Frobenius, *Abhandlungen* 3, 355–377.
17. F. Gantmacher. *Matrix Theory*. AMS Chelsea Publishing, 2000. This work, published in two volumes, is an English translation of Gantmacher's *Teoriya Matrits* (Moscow, 1953). It first appeared in 1959.
18. W. R. Hamilton. *Lectures on Quaternions*. Hodges and Smith, Dublin, 1853.
19. T. Hawkins. The origins of the theory of group characters. *Arch. Hist. Exact. Sci.*, 7:142–170, 1971.
20. T. Hawkins. New light on Frobenius' creation of the theory of group characters. *Arch. Hist. Exact Sci.*, 12:217–243, 1974.
21. T. Hawkins. Cauchy and the spectral theory of matrices. *Historia Math.*, 2:1–29, 1975.
22. T. Hawkins. Another look at Cayley and the theory of matrices. *Archives internationales d'histoire des sciences*, 26:82–112, 1977.
23. T. Hawkins. Weierstrass and the theory of matrices. *Archive for History of Exact Sciences*, 17:119–163, 1977.
24. T. Hawkins. *Emergence of the Theory of Lie Groups. An Essay on the History of Mathematics 1869–1926*. Springer, New York, 2000.
25. T. Hawkins. Frobenius, Cartan, and the Problem of Pfaff. *Archive for History of Exact Sciences*, 59:381–436, 2005.
26. C. Hermite. Remarque sur un théorème de M. Cauchy. *Comptes rendus Acad. Sci. Paris*, 41:459–481, 1855. Oeuvres **1** (Paris, 1905).
27. C. Hermite. Sur la théorie de la transformation des fonctions abéliennes. *Comptes Rendus, Acad. Sci. Paris*, 40, 1855. Reprinted in *Oeuvres* 1, 444–477.
28. A. Hurwitz. Ueber diejenigen algebraische Gebilde, welche eindeutige Transformationen in sich zulassen. *Math. Ann.*, 32:290–308, 1888. Reprinted in *Werke* 1, 241–259.
29. A. Hurwitz. Über die Erzeugung der Invarianten durch Integration. *Göttinger Nachrichten*, pages 71–90, 1897. Reprinted in *Werke* 2, 546–564.
30. C. G. J. Jacobi. De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis, quae solis quadratis variabilium constant; .... *Jl. für die reine u. angew. Math.*, 12:1–69, 1834. Reprinted in *Werke* 3, 191–268.
31. C. G. J. Jacobi. Über eine elementare Transformation eines in Bezug auf jedes von zwei Variablen-Systemen linearen und homogenen Ausdrucks. *Jl. für die reine u. angew. Math.*, 53:265–270, 1857. Reprinted in *Werke* 3, 583–590.
32. C. Jordan. Mémoire sur les équations différentielles linéaires à intégrale algébrique. *Jl. für die reine u. angew. Math.*, 84:89–215, 1878. Reprinted in *Oeuvres* 2, 13–139.
33. A. Krazer. *Lehrbuch der Thetafunktionen*. Teubner, Leipzig, 1903. Reprinted by Chelsea Publishing Company (New York, 1970).
34. L. Kronecker. Über die elliptischen Functionen, für welche complexe Multiplication stattfindet. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 455–460, 1857. Reprinted in *Werke* 3, 177–183.
35. L. Kronecker. Über bilineare Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, 1:145–162, 1866. Reprinted in *Jl. für die reine u. angew. Math.*, 68:273–285 and in *Werke* 1, 145–162.
36. L. Kronecker. Über Schaaren quadratischer Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 339–346, 1868. Reprinted in *Werke* 1, 163–174. The above title for the work was added by the editor (K. Hensel). See *Werke* 1, 163n. 1.
37. L. Kronecker. Über Schaaren von quadratischen und bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 59–76, 1874. Presented Jan 19, 1874. Reprinted in *Werke* 1, 349–372.

38. L. Kronecker. Über Schaaren von quadratischen und bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 149–156, 1874. Presented Feb 16, 1874. Reprinted in *Werke* 1, 373–381.
39. L. Kronecker. Über Schaaren von quadratischen und bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 206–232, 1874. Presented March 16, 1874. Reprinted in *Werke* 1, 382–413.
40. L. Kronecker. Über die congruenten Transformation der bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 397–447, 1874. Presented April 23, 1874. Reprinted in *Werke* 1, 423–483.
41. L. Kronecker. Über die Composition der Systeme von n^2 Grössen mit sich selbst. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 1081–1088, 1890. Reprinted in *Werke* 3_1, 463–473.
42. L. Kronecker. Algebraische Reduction der Schaaren bilinearer Formen. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 1225–1237, 1890. Reprinted in *Werke* 3_2, 141–155.
43. L. Kronecker. Algebraische Reduction der Schaaren quadratischer Formen. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 1375–1388, 1890. Reprinted in *Werke* 3_2, 159–174.
44. L. Kronecker. Algebraische Reduction der Schaaren quadratischer Formen. *Sitzungsberichte der Akademie der Wiss. zu Berlin*, pages 9–17, 34–44, 1891. Reprinted in *Werke* 3_2, 175–198.
45. E. Laguerre. Sur le calcul des systèmes linéaires. *Journal École Polytechnique*, 62, 1867. Reprinted in *Oeuvres* 1, 221-267.
46. S. Lang. *Complex Multiplication*. Springer-Verlag, New York, 1983.
47. R. Lipschitz. Beweis eines Satzes aus der Theorie der Substitutionen. *Acta Mathematica*, 10:137–144, 1887.
48. C. C. MacDuffee. *The Theory of Matrices*. Springer, Berlin, 1933.
49. P. Muth. *Theorie und Anwendung der Elementartheiler*. Teubner, Leipzig, 1899.
50. K. H. Parshall and D. Rowe. *The Emergence of the American Mathematical Research Community, 1876–1900: J. J. Sylvester, Felix Klein, and E. H. Moore*. History of Mathematics, Vol. 8. American Mathematical Society, 1994.
51. M. Pasch. Peter Muth. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 18:454–456, 1909.
52. M. Rosen. Abelian varieties over $\mathbb{C}$. In G. Cornell and J. Silverman, editors, *Arithmetic Geometry*, pages 79–101. Springer-Verlag, New York, 1986.
53. G. Shimura. *Abelian Varieties with Complex Multiplication and Modular Functions*. Princeton University Press, 1998.
54. J. J. Sylvester. On the equation to the secular inequalities in the planetary theory. *Phil. Mag.*, 16:110–11, 1883. Reprinted in *Papers*, v. 4, 110—111.
55. H. W. Turnbull and A. C. Aitken. *An Introduction to the Theory of Canonical Matrices*. Blackie and Son, London & Glasgow, 1932.
56. S. G. Vlăduţ. *Kronecker's Jugendtraum and Modular Functions*. Gordon and Breach, 1991.
57. H. Weber. Ueber die Transformationstheorie der Theta-Functionen, ins Besondere derer von drei Veränderlichen. *Annali di matematica*, (2) 9:126–166, 1878.
58. H. Weber. *Elliptische Functionen und algebraische Zahlen*. Vieweg, Braunschweig, 1891. A second edition was published in 1908 under the same title but as the third volume of the second edition of Weber's *Lehrbuch der Algebra*.
59. K. Weierstrass. Über ein die homogenen Functionen zweiten Grades betreffendes Theorem. *Monatsberichte der Akademie der Wiss. zu Berlin*, 1858. Reprinted in *Werke* 1, 233–246.
60. K. Weierstrass. Zur Theorie des quadratischen und bilinearen Formen. *Monatsberichte der Akademie der Wiss. zu Berlin*, pages 311–338, 1868. Reprinted with modifications in *Werke* 2, 19–44.