CHAPTER XXII

THE EQUIVALENCE AND CLASSIFICATION OF PAIRS OF QUADRATIC FORMS

101. Two Theorems in the Theory of Matrices. In order to justify the applications we wish to make of the theory of elementary divisors to the subject of quadratic forms, it will be necessary for us to turn back for a moment to the general theory of matrices.

DEFINITION. *If $\phi(x)$ is a polynomial:*

$$\phi(x) \equiv a_0 x^m + a_1 x^{m-1} + \cdots + a_{m-1}x + a_m,$$

then
$$a_0 \mathbf{x}^m + a_1 \mathbf{x}^{m-1} + \cdots + a_{m-1}\mathbf{x} + a_m \mathbf{I}$$

is called a polynomial in the matrix $\mathbf{x}$ *and is denoted by* $\phi(\mathbf{x})$.*

We come now to one of the most fundamental theorems in the whole theory of matrices:

THEOREM 1. *If* $\mathbf{a}$ *is a matrix, and* $\phi(\lambda)$ *its characteristic function, then*

$$\phi(\mathbf{a}) = 0.$$

This equation is called *the Hamilton-Cayley equation.*

Let $\mathbf{c}$ be the characteristic matrix of $\mathbf{a}$:

$$\mathbf{c} = \mathbf{a} - \lambda \mathbf{I}.$$

This being a λ-matrix of the first degree, its adjoint $\mathbf{C}$ will be a λ-matrix of degree not higher than $n-1$, if n is the order of the matrix $\mathbf{a}$:

(1)
$$\mathbf{C} \equiv \mathbf{C}_{n-1}\lambda^{n-1} + \mathbf{C}_{n-2}\lambda^{n-2} + \cdots + \mathbf{C}_0.$$

We may also write

(2)
$$\phi(\lambda) \equiv k_n\lambda^n + k_{n-1}\lambda^{n-1} + \cdots + k_0.$$

Now referring to formula (5), § 25, we see that

$$\mathbf{aC} - \lambda\mathbf{C} \equiv \phi(\lambda)\mathbf{I}.$$

* It should be noticed that, according to this definition, the coefficients of a polynomial in $\mathbf{x}$ are scalars. Contrast this with a λ-matrix, in which the coefficients are matrices and the variable a scalar. Both of these conceptions would be included in expressions of the form:

$$\mathbf{a}_0\mathbf{x}^m\mathbf{b}_0 + \mathbf{a}_1\mathbf{x}^{m-1}\mathbf{b}_1 + \cdots + \mathbf{a}_{m-1}\mathbf{x}\mathbf{b}_{m-1} + \mathbf{a}_m.$$

Substituting here from (1) and (2), we have, on equating corresponding powers of λ,

$$\begin{aligned}
\mathbf{aC}_0 \qquad\quad &= k_0\mathbf{I},\\
\mathbf{aC}_1 - \mathbf{C}_0 &= k_1\mathbf{I},\\
\mathbf{aC}_2 - \mathbf{C}_1 &= k_2\mathbf{I},\\
\cdots\cdots &\cdots\\
\cdots\cdots &\cdots\\
\mathbf{aC}_{n-1} - \mathbf{C}_{n-2} &= k_{n-1}\mathbf{I},\\
- \mathbf{C}_{n-1} &= k_n\mathbf{I}.
\end{aligned}$$

If we multiply these equations in succession by $\mathbf{I}$, $\mathbf{a}$, $\mathbf{a}^2$, $\cdots$ $\mathbf{a}^n$, and add, the first members cancel out, and we get

$$k_0\mathbf{I} + k_1\mathbf{a} + k_2\mathbf{a}^2 + \cdots + k_n\mathbf{a}^n = 0.$$

This is precisely the equation

$$\phi(\mathbf{a}) = 0$$

which we wished to establish.

As a means of deducing our second theorem, we next establish a lemma which relates merely to scalar quantities.

LEMMA. *If $\psi(x)$ is a polynomial of the nth degree $(n > 0)$ whose constant term is not zero, there exists a polynomial $\chi(x)$ of degree less than n such that*

$$(\chi(x))^2 - x$$

is divisible by $\psi(x)$.

Let $x-a$, $x-b$, $x-c$, $\cdots$ be the distinct linear factors of $\psi(x)$, so that we may write

$$\psi(x) \equiv k(x-a)^\alpha(x-b)^\beta(x-c)^\gamma \cdots \qquad (\alpha+\beta+\gamma+\cdots=n).$$

None of the constants a, b, c, $\cdots$ are zero, since, by hypothesis, the constant term of ψ is not zero. Let us, further, denote by $\psi_1(x)$ the polynomial obtained from ψ by omitting the factor $(x-a)^\alpha$, by $\psi_2(x)$ the polynomial obtained from ψ by omitting the factor $(x-b)^\beta$, etc., and finally let us form, with undetermined coefficients, the polynomials

$$\begin{aligned}
A(x) &\equiv A_0 + A_1(x-a) + A_2(x-a)^2 + \cdots + A_{\alpha-1}(x-a)^{\alpha-1},\\
B(x) &\equiv B_0 + B_1(x-b) + B_2(x-b)^2 + \cdots + B_{\beta-1}(x-b)^{\beta-1},\\
C(x) &\equiv C_0 + C_1(x-c) + C_2(x-c)^2 + \cdots + C_{\gamma-1}(x-c)^{\gamma-1}.\\
&\cdots\cdots\cdots\cdots\\
&\cdots\cdots\cdots\cdots
\end{aligned}$$

From these polynomials we now form the polynomial

$$\chi(x) \equiv A(x)\psi_1(x) + B(x)\psi_2(x) + C(x)\psi_3(x) + \cdots$$

whose degree can obviously not exceed $n-1$. We wish to show that the coefficients $A_i, B_i, \cdots$ can be so determined that this polynomial $X(x)$ satisfies the conditions of our lemma.

Since $\psi_2, \psi_3, \cdots$ are all divisible by $(x-a)^\alpha$, a necessary and sufficient condition that $(\chi(x))^2 - x$ be divisible by this factor is that the polynomial

$$\phi(x) \equiv (A(x))^2(\psi_1(x))^2 - x$$

be divisible by $(x-a)^\alpha$. We have

$$\phi(a) = A_0^2 k^2 (a-b)^{2\beta}(a-c)^{2\gamma} \cdots - a.$$

In order that $\phi(x)$ be divisible by $x - a$ it is therefore necessary and sufficient that

$$(3) \qquad A_0^2 = \frac{a}{k^2(a-b)^{2\beta}(a-c)^{2\gamma}\cdots}.$$

Neither numerator nor denominator here being zero, we thus obtain two distinct values for A_0, both different from zero. If we give to A_0 one of these values, $\phi(x)$ is divisible by $x-a$. A necessary and sufficient condition that it be also divisible by $(x-a)^2$ is that $\phi'(a) = 0$, accents here, and in what follows, denoting differentiation. We shall see in a moment that this condition can be imposed in one, and only one, way by a suitable choice of A_1. The condition that $\phi(x)$ be divisible by $(x-a)^3$ is then simply $\phi''(a) = 0$. We wish to show that this process can be continued until we have finally imposed the condition that $\phi(x)$ be divisible by $(x-a)^\alpha$. For this purpose we use the method of mathematical induction, and assume that $A_0, \cdots A_{s-1}$ have been so determined that $\phi(a) = \phi'(a) = \cdots = \phi^{[s-1]}(a) = 0$. It remains then merely to show that A_s can be so determined that $\phi^{[s]}(a) = 0$. For this purpose we notice that

$$(4) \qquad \phi^{[s]}(x) \equiv 2A^{[s]}(x)A(x)(\psi_1(x))^2 + R_s(x)$$

where $R_s(x)$ is an integral rational function with numerical coefficients of $\psi_1, \psi_1', \cdots \psi_1^{[s]}, A, A', \cdots A^{[s-1]}$. Since

$$A(a) = A_0,\ A'(a) = A_1,\ A''(a) = 2!A_2, \cdots A^{[s-1]}(a) = (s-1)!A_{s-1},$$

it follows that $R(a)$ is a known constant, that is, that it does not depend on any of the still undetermined constants $A_s, A_{s+1}, \cdots A_{\alpha-1}$, nor on the B's, C's, etc. Consequently we see from (4) that a necessary and sufficient condition that $\phi^{[s]}(a) = 0$ is that A_s have the value

$$(5) \qquad A_s = \frac{-R_s(a)}{2\,s!A_0(\psi_1(a))^2}.$$

Determining the coefficients $A_1, A_2, \cdots A_{\alpha-1}$ in succession by means of this formula, we finally determine the polynomial $A(x)$ in such a way that $\phi(x)$ is divisible by $(x-a)^\alpha$. For this determination, $(\chi(x))^2 - x$ will, as we saw above, be divisible by $(x-a)^\alpha$.

In the same way we can now determine the coefficients of $B(x)$ so that $(\chi(x))^2 - x$ is divisible by $(x-b)^\beta$; then we determine the coefficients of $C(x)$ so that $(\chi(x))^2 - x$ is divisible by $(x-c)^\gamma$; etc. When all the polynomials $A, B, C, \cdots$ are thus determined, $(X(x))^2 - x$ is divisible by $\psi(x)$, and our lemma is proved.

THEOREM 2. *If* $\mathfrak{a}$ *is a non-singular matrix of order n, there exist matrices* $\mathfrak{b}$ *of order n (necessarily non-singular) with the following properties:*

$$\mathfrak{b}^2 = \mathfrak{a},$$

$\mathfrak{b}$ *is a polynomial in* $\mathfrak{a}$ *of degree less than n.*

Since $\mathfrak{a}$ is non-singular, its characteristic function $\phi(\lambda)$ is a polynomial of the nth degree whose constant term is not zero. Hence, by the preceding lemma, a polynomial $\chi(\lambda)$ of degree less than n can be determined such that

$$(\chi(\lambda))^2 - \lambda \equiv \phi(\lambda)f(\lambda)$$

where $f(\lambda)$ is also a polynomial. From this identity it follows that

$$(\chi(\mathfrak{a}))^2 - \mathfrak{a} = \phi(\mathfrak{a})f(\mathfrak{a}).$$

Since, by Theorem 1, $\phi(\mathfrak{a}) = 0$, the last equation may be written

$$(\chi(\mathfrak{a}))^2 = \mathfrak{a},$$

so that $\mathfrak{b} = \chi(\mathfrak{a})$ is a matrix satisfying the conditions of our theorem, which is thus proved.

102. Symmetric Matrices. The application of the theory of elementary divisors to the subject of quadratic forms rests on the following proposition:

THEOREM 1. *If* $\mathfrak{a}_1$ *and* $\mathfrak{a}_2$ *are symmetric matrices and if there exist two non-singular matrices* $\mathfrak{p}$ *and* $\mathfrak{q}$ *such that*

$$(1) \qquad \mathfrak{a}_2 = \mathfrak{p}\mathfrak{a}_1\mathfrak{q},$$

then there also exists a non-singular matrix P *such that*

(2) $$a_2 = P'a_1P$$

where P' *is the conjugate of* P.*

Let us denote by p' and q' the conjugates of p and q respectively. Taking the conjugates of both sides of (1), and remembering that a_1 and a_2, being symmetric, are their own conjugates, we get, by Theorem 6, § 22,

(3) $$a_2 = q'a_1p'.$$

By equating the values of a_2 in (1) and (3), we readily deduce the further relation

(4) $$(q')^{-1}pa_1 = a_1p'q^{-1}.$$

For brevity we will let

(5) $$U = (q')^{-1}p, \qquad U' = p'q^{-1},$$

and note that U' is the conjugate of U; cf. Exercise 6, § 25. Equation (4) may then be written

(6) $$Ua_1 = a_1U'.$$

From this equation we infer at once the following further ones:

(7) $$\begin{cases} U^2a_1 = Ua_1U' = a_1U'^2, \\ U^3a_1 = Ua_1U'^2 = a_1U'^3, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ U^ka_1 = Ua_1U'^{k-1} = a_1U'^k. \end{cases}$$

Let us now multiply the equations (6) and (7) and also the equation $a_1 = a_1$ by any set of scalar constants and add them together. We see in this way that if $\chi(U)$ is any polynomial in U,

(8) $$\chi(U)a_1 = a_1\chi(U').$$

* A proof of this theorem much simpler than that given in the text is the following: From (1) we infer at once that a_1 and a_2 have the same rank. Hence the quadratic forms of which a_1 and a_2 are the matrices are equivalent to each other by Theorem 4, § 46. If we denote by P the matrix of the linear transformation which carries over the quadratic form a_1 into the form a_2, we see, from Theorem 1, § 43, that equation (2) holds.

This proof would not enable us to infer that P can be expressed in terms of p and q alone, and this is essential for our purposes.

We will choose the polynomial

$$V = \chi(U)$$

so that V is non-singular and

$$V^2 = U,$$

as is seen to be possible by Theorem 2, § 101. Denoting by V' the conjugate of V, we evidently have

$$V' = \chi(U'),$$

so that we may write (8) in the form

$$Va_1 = a_1V',$$

or $$a_1 = V^{-1}a_1V'.$$

We now substitute this value in (1) and get

(9) $$a_2 = pV^{-1}a_1V'q.$$

From the first equation (5) we infer the formula

$$pV^{-1} = q'V.$$

Consequently pV^{-1} is the conjugate of $V'q$, so that if we let

$$P = V'q,$$

equation (9) may be written

$$a_2 = P'a_1P,$$

and our theorem is proved.

The proof just given enables us to add the

COROLLARY. *As the matrix* P *of the foregoing theorem may be taken the matrix* $V'q$ *where* V' *is the conjugate of any one of the square roots, determined by Theorem* 2, § 101, *of* $(q')^{-1}p$.

In particular it will be seen that P depends on p and q but *not* on a_1 or a_2. Hence if a_1, a_2, b_1, b_2 are symmetric matrices, and there exist two non-singular matrices p and q such that

$$a_2 = pa_1q, \qquad b_2 = pb_1q,$$

then there exists a non-singular matrix P such that

$$a_2 = P'a_1P, \qquad b_2 = P'b_1P.$$

From this and Theorem 2, § 96, we infer

THEOREM 2. *If* $\mathsf{a}_1, \mathsf{a}_2, \mathsf{b}_1, \mathsf{b}_2$, *are symmetric matrices of which* $\mathsf{b}_1, \mathsf{b}_2$ *are non singular, a necessary and sufficient condition that a non-singular matrix* P *exist such that*

$$(10) \qquad \mathsf{a}_2 = \mathsf{P}'\mathsf{a}_1\mathsf{P}, \qquad \mathsf{b}_2 = \mathsf{P}'\mathsf{b}_1\mathsf{P},$$

where P' *is the conjugate of* P, *is that the matrices*

$$\mathsf{a}_1 - \lambda\mathsf{b}_1, \qquad \mathsf{a}_2 - \lambda\mathsf{b}_2$$

have the same invariant factors, — or, if we prefer, the same elementary divisors.

If, in particular, $\mathsf{b}_1 = \mathsf{b}_2 = \mathsf{I}$, where I is the unit matrix, we have, from the second equation (10), the formula

$$\mathsf{I} = \mathsf{P}'\mathsf{P}.$$

Such a matrix P we call an *orthogonal matrix* according to the definition, which will readily be seen to be equivalent to the one given in the first footnote on page 154:

DEFINITION. *By an orthogonal matrix we understand a non-singular matrix whose inverse is equal to its conjugate.*

In the special case just referred to, Theorem 2 may be stated in the following form:

THEOREM 3. *If* a_1 *and* a_2 *are two symmetric matrices, a necessary and sufficient condition that an orthogonal matrix* P *exist such that*

$$\mathsf{a}_2 = \mathsf{P}'\mathsf{a}_1\mathsf{P}$$

is that the characteristic matrices of a_1 *and* a_2 *have the same invariant factors, — or, if we prefer, the same elementary divisors.*

If this theorem is compared with Theorem 3, § 96, it will be seen that it differs from it only in two respects, first that a_1 and a_2 are assumed to be symmetric, and secondly that P is required to be orthogonal.

*** 103. The Equivalence of Pairs of Quadratic Forms.** Let us consider the two pairs of quadratic forms

$$\phi_1 \equiv \sum_1^n a'_{ij}x_ix_j, \qquad \psi_1 \equiv \sum_1^n b'_{ij}x_ix_j,$$

and

$$\phi_2 \equiv \sum_1^n a''_{ij}x_ix_j, \qquad \psi_2 \equiv \sum_1^n b''_{ij}x_ix_j,$$

of which the two forms ψ_1 and ψ_2 are assumed to be non-singular. We will inquire under what conditions these two pairs of forms are equivalent; that is, under what conditions a linear transformation

$$\mathsf{c} \begin{cases} x_1 = c_{11}x'_1 + \cdots + c_{1n}x'_n \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_n = c_{n1}x'_1 + \cdots + c_{nn}x'_n \end{cases}$$

exists which carries over ϕ_1 into ϕ_2 and, at the same time, ψ_1 into ψ_2.

If we denote the conjugate of the matrix c by c', and the matrices of the forms $\phi_1, \psi_1, \phi_2, \psi_2$ by $\mathsf{a}_1, \mathsf{b}_1, \mathsf{a}_2, \mathsf{b}_2$ respectively, we know, by Theorem 1, § 43, that the transformation c carries over ϕ_1 and ψ_1 into forms with the matrices

$$\mathsf{c}'\mathsf{a}_1\mathsf{c}, \qquad \mathsf{c}'\mathsf{b}_1\mathsf{c}$$

respectively; so that, if these are the forms ϕ_2 and ψ_2, we have

$$(1) \qquad a_2 = \mathsf{c}'\mathsf{a}_1\mathsf{c}, \qquad \mathsf{b}_2 = \mathsf{c}'\mathsf{b}_1\mathsf{c}.$$

Consequently, by Theorem 2, § 102, the two λ-matrices

$$\mathsf{a}_1 - \lambda\mathsf{b}_1, \qquad \mathsf{a}_2 - \lambda\mathsf{b}_2$$

have the same invariant factors and elementary divisors.

Conversely, by the same theorem, if these two λ-matrices have the same invariant factors (or elementary divisors), a matrix c, independent of λ, exists which satisfies both equations (1); and hence the two pairs of quadratic forms are equivalent. Thus we have proved

THEOREM 1. *If* ϕ_1, ψ_1 *and* ϕ_2, ψ_2 *are two pairs of quadratic forms in n variables, in which* ψ_1 *and* ψ_2 *are non-singular, a necessary and sufficient condition that these two pairs of forms be equivalent is that the matrices of the two pencils*

$$\phi_1 - \lambda\psi_1, \qquad \phi_2 - \lambda\psi_2$$

*have the same invariant factors, — or, if we prefer, the same elementary divisors.**

A special case of this theorem which is of considerable importance is that in which both of the forms ψ_1 and ψ_2 reduce to

$$x_1^2 + x_2^2 + \cdots + x_n^2.$$

*For brevity, we shall speak of these invariant factors and elementary divisors as the invariant factors and elementary divisors of the *pairs of forms* ϕ_1, ψ_1 and ϕ_2, ψ_2 respectively.

In this case we have to deal with orthogonal transformations (cf. the Definition in Exercise 1, § 52), and our theorem may be stated in the form *

THEOREM 2. *If* a_1 *and* a_2 *are the matrices of two quadratic forms, a necessary and sufficient condition that there exist an orthogonal transformation which carries over one of these forms into the other is that the characteristic matrices of* a_1 *and* a_2 *have the same invariant factors, — or, if we prefer, the same elementary divisors.*

To illustrate the meaning of the theorems of this section, let us consider again briefly the problem of the simultaneous reduction of two quadratic forms to sums of squares. In Chapter XIII we became acquainted with two cases in which this reduction is possible; cf. Theorem 2, § 58, and Theorem 2, § 59. We are in a position now to state a necessary and sufficient condition for the possibility of this reduction, provided that one of the two forms is non-singular.

For this purpose, consider the two quadratic forms

$$\phi \equiv k_1x_1^2 + k_2x_2^2 + \cdots + k_nx_n^2,$$
$$\psi \equiv c_1x_1^2 + c_2x_2^2 + \cdots + c_nx_n^2,$$

where we assume, in order that the second form may be non-singular, that none of the c's vanish. The matrix of the pencil $\phi - \lambda\psi$ is

$$\left\| \begin{matrix} k_1 - c_1\lambda & 0 & 0 & \cdots & 0 \\ 0 & k_2 - c_2\lambda & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & k_n - c_n\lambda \end{matrix} \right\|,$$

and the elementary divisors of this matrix are

$$\lambda - \frac{k_1}{c_1}, \qquad \lambda - \frac{k_2}{c_2}, \qquad \cdots\cdots \qquad \lambda - \frac{k_n}{c_n},$$

all of the first degree. Consequently, any pair of quadratic forms equivalent to the pair just considered must have a λ-matrix whose elementary divisors are all of the first degree.

Conversely, if we have a pair of quadratic forms, of which the first is non-singular, whose λ-matrix has elementary divisors all of

* This theorem is, of course, essentially equivalent to Theorem 5, § 102, of which it may be regarded as an immediate consequence.

the first degree, we can obviously choose the constants k and c in such a way that the λ-matrix of the forms ϕ and ψ just considered has these same elementary divisors, and therefore the given forms are equivalent to these special forms ϕ and ψ. Thus we have proved the theorem:

THEOREM 3. *If* ϕ *and* ψ *are quadratic forms and* ψ *is non-singular, a necessary and sufficient condition that it be possible to reduce* ϕ *and* ψ *simultaneously by a non-singular linear transformation to forms into which only the square terms enter is that all the elementary divisors of the pair of forms be of the first degree.*

This theorem obviously includes as a special case Theorem 2 of § 58, since the elementary divisors are necessarily of the first degree when the λ-equation has no multiple roots.

Comparing the theorem just proved with Theorem 2, § 59, we see that under the conditions of that theorem the elementary divisors must be of the first degree. Hence

THEOREM 4. *If* ψ *is a non-singular, definite, quadratic form, and* ϕ *is a real quadratic form, all the elementary divisors of this pair of forms are necessarily of the first degree.*

104. Classification of Pairs of Quadratic Forms. We consider the pair of quadratic forms

$$(1) \qquad \phi \equiv \sum_1^n a_{ij}x_ix_j, \qquad \psi \equiv \sum_1^n b_{ij}x_ix_j,$$

and assume, as before, that ψ is non-singular. We denote the elementary divisors of these forms, as in § 99, by

$$(\lambda - \lambda_1)^{e_1}, \quad (\lambda - \lambda_2)^{e_2}, \cdots (\lambda - \lambda_k)^{e_k} \qquad (e_1 + e_2 + \cdots + e_k = n).$$

The symbol $[e_1\ e_2 \cdots e_k]$ we call the *characteristic* of the pair of quadratic forms; and all pairs of quadratic forms which have the same characteristic we speak of as forming a *category*.*

We have here, precisely as in the case of bilinear forms, the theorem:

THEOREM. *If* $\lambda_1, \lambda_2, \cdots \lambda_k$ *are any constants, equal or unequal, and* $e_1, e_2, \cdots e_k$ *are any positive integers whose sum is* n, *there exist pairs*

* Thus, for instance, all pairs of forms of which the second is non-singular and which admit of simultaneous reduction to sums of squares, form a category whose characteristic is $[1\ 1 \cdots 1]$. Cf. Theorem 3, § 103.

of quadratic forms in n variables, the second form in each pair being non-singular, which have the elementary divisors

$$(2) \qquad (\lambda-\lambda_1)^{e_1}, \quad (\lambda-\lambda_2)^{e_2}, \quad \cdots\cdots(\lambda-\lambda_k)^{e_k}.$$

The proof of this theorem consists in considering the following pair of quadratic forms, analogous to the normal form (3′) of § 99:

$$(3)\left\{\begin{aligned}\phi &\equiv \left(\sum_1^{e_1}\lambda_1 c_1 x_i x_{e_1-i+1} + \sum_1^{e_1-1} d_1 x_i x_{e_1-i}\right) + \left(\sum_{e_1+1}^{e_1+e_2}\lambda_2 c_2 x_i x_{2e_1+e_2-i+1} + \sum_{e_1+1}^{e_1+e_2-1} d_2 x_i x_{2e_1+e_2-i}\right) \\ &\qquad + \cdots + \left(\sum_{n-e_k+1}^{n-1}\lambda_k c_k x_i x_{2n-e_k-i+1} + \sum_{n-e_k+1}^{n-1} d_k x_i x_{2n-e_k-i}\right), \\ \psi &\equiv \sum_1^{e_1} c_1 x_i x_{e_1-i+1} + \sum_{e_1+1}^{e_1+e_2} c_2 x_i x_{2e_1+e_2-i+1} + \sum_{e_1+e_2+1}^{e_1+e_2+e_3} c_3 x_i x_{2e_1+2e_2+e_3-i+1} \\ &\qquad + \cdots + \sum_{n-e_k+1}^{n} c_k x_i x_{2n-e_k-i+1},\end{aligned}\right.$$

where $c_1, \cdots c_k, d_1, \cdots d_k$ are constants which may be chosen at pleasure, provided none of them are zero.

The λ-matrix of this pair of forms is the same as the λ-matrix of the pair of bilinear forms (3′) of § 99, and therefore has the desired elementary divisors.

A reference to Theorem 1, § 103, shows that formula (3) yields a normal form to which every pair of quadratic forms, of which the second is non-singular and whose elementary divisors are given by (2), can be reduced.

The categories, of which we have so far spoken, may be divided into classes by the same methods we used in § 99 in the case of bilinear forms. This may be done, as before, either by simply noting which of the λ_i's are equal to each other, or by further distinguishing between the cases where some of the λ_i's are zero.

We are now in a position to see exactly in what way our elementary divisors give us a more powerful instrument than we had in the invariants Θ_i of § 57. These invariants Θ_i, being the coefficients of the λ-equation of our pair of forms, determine the constants λ_i, which are the roots of this equation, as well as the multiplicities of these roots. They do *not* determine the degrees e_i of the elementary divisors, and the use of the Θ_i's alone does not, in all cases, enable us to determine whether two pairs of forms are equivalent or not. Thus, for instance, we may have two pairs of forms with exactly the

same invariants Θ_i but with characteristics $[(1\,1)\,1\,1 \cdots 1]$ and $[2\,1\,1 \cdots 1]$ respectively.* It will be seen, therefore, that the Θ_i's form in only a very technical sense a complete system of invariants.

EXERCISES

1. Form a numerical example in the case $n = 3$ to illustrate the statement made in the next to the last sentence of this section.

2. Prove that if two equivalent pairs of quadratic forms have two elementary divisors of the first degree which correspond to the same linear factor, there exist an infinite number of linear transformations which carry over one pair of forms into the other.

3. Prove the general theorem, of which Exercise 2 is a special case, namely, that if two equivalent pairs of quadratic forms have a characteristic in which one or more parentheses appear, there exist an infinite number of linear transformations which carry over one pair of forms into the other.

4. Prove that if two equivalent pairs of quadratic forms have a characteristic in which no parentheses appear, only a finite number of linear transformations exist which carry over one pair of forms into the other.†

How are these transformations related to each other?

105. Pairs of Quadratic Equations, and Pencils of Forms or Equations.‡ In dealing with quadratic forms, the questions of equivalence and classification do not always present themselves to us in precisely the form in which we have considered them in the last two sections. We frequently have to deal not with the quadratic forms themselves but with the *equations* obtained by setting the forms equal to zero. Two such pairs of equations we shall regard as equivalent, not merely if the forms in them are equivalent, but also if one pair of forms can be obtained from the other by multiplication by constants different from zero.

Let us consider two quadratic forms ϕ, ψ, of which we assume, as before, that the second is non-singular, and inquire what the effect on the elementary divisors

$$(1) \qquad (\lambda-\lambda_1)^{e_1}, \quad (\lambda-\lambda_2)^{e_2}, \cdots\cdots (\lambda-\lambda_k)^{e_k}$$

* We may, in the case $n = 3$, put the same thing geometrically (cf. the next section) by saying that it is impossible to distinguish between the case of two conics having double contact and that of two conics having simple contact at a single point by the use of the invariants Θ_i alone, whereas these two cases are at once distinguished by the use of elementary divisors.

† The exercise in § 58 is practically a special case of this.

‡ Questions similar to those treated in this section might have been taken up in the last chapter for the case of bilinear forms.

of these forms will be if the forms are multiplied respectively by the constants p, q which are both assumed to be different from zero. Let us write

$$\phi_1 \equiv p\phi, \qquad \psi_1 \equiv q\psi.$$

Then

(2) $$\phi_1 - \lambda\psi_1 \equiv p(\phi - \lambda'\psi)$$

where $$\lambda' = \frac{q}{p}\lambda.$$

Let $\lambda - \alpha$ be any one of the linear factors of the matrix of $\phi - \lambda\psi$, so that α is any one of the constants $\lambda_1, \lambda_2, \cdots \lambda_k$; and let us denote, as in the footnote to Definition 3, § 92, by l_i the exponent of the highest power of $\lambda - \alpha$ which is a factor of all the i-rowed determinants of this matrix. Then it is clear, from (2), that l_i is the exponent of the highest power of $\lambda' - \alpha$ which is a factor of all the i-rowed determinants of the matrix of $\phi_1 - \lambda\psi_1$. In other words,

$$\left(\lambda - \frac{p\alpha}{q}\right)^{l_i}$$

is the highest power of the linear factor $\lambda - p\alpha/q$ which is a factor of all the i-rowed determinants of the matrix of $\phi_1 - \lambda\psi_1$. Turning now to the definition of elementary divisors as given in the footnote to Definition 3, § 92, we see that the elementary divisors of the matrix of $\phi_1 - \lambda\psi_1$ differ from those of the matrix of $\phi - \lambda\psi$ only in having the constants λ_i replaced by the constants $p\lambda_i/q$. We thus have the result:

THEOREM 1. *If the pair of quadratic forms ϕ, ψ, of which the second is assumed to be non-singular, has the elementary divisors*

$$(\lambda - \lambda_1)^{e_1}, \quad (\lambda - \lambda_2)^{e_2}, \quad \cdots\cdots (\lambda - \lambda_k)^{e_k},$$

and if p, q are constants different from zero, then the pair of quadratic forms $p\phi$, $q\psi$ has the elementary divisors

$$(\lambda - \lambda_1')^{e_1}, \quad (\lambda - \lambda_2')^{e_2}, \quad \cdots\cdots (\lambda - \lambda_k')^{e_k}$$

where $$\lambda_i' = \frac{p}{q}\lambda_i.$$

In particular, it will be seen that these two pairs of forms have the same characteristic, even when the conception of the characteristic is refined not merely by inserting parentheses but also by the use of the small zeros.

The theorem just proved shows that pairs of homogeneous quadratic equations, of which the second equation in each pair is non-singular, may be classified by the use of their characteristics precisely as was done in the last section for pairs of quadratic forms. We proceed to illustrate this in the case $n = 3$, where we may consider that we have to deal with the classification of pairs of conics in a plane, one of the conics being non-singular.

We have here three categories represented by the following normal forms:*

I. [1 1 1] $$\begin{cases} \phi \equiv \lambda_1 x_1^2 + \lambda_2 x_2^2 - \lambda_3 x_3^2 \\ \psi \equiv x_1^2 + x_2^2 - x_3^2. \end{cases}$$

II. [2 1] $$\begin{cases} \phi \equiv 2\lambda_1 x_1 x_2 + x_1^2 + \lambda_2 x_3^2 \\ \psi \equiv 2x_1x_2 + x_3^2. \end{cases}$$

III. [3] $$\begin{cases} \phi \equiv 2\lambda_1 x_1 x_3 + \lambda_1 x_2^2 + 2x_1x_2 \\ \psi \equiv 2x_1x_3 + x_2^2. \end{cases}$$

We next subdivide these categories into classes, and, by an examination of the normal form in each case, we are enabled at once to characterize each class by certain projective properties which it has, and which are shared by no other class.† Since the conic ψ is non-singular in all cases, this fact need not be explicitly stated.

[1 1 1] ϕ and ψ intersect in four distinct points.

[(1 1) 1] ϕ and ψ have double contact.

[(1 1 1)] ϕ and ψ coincide.

[2 1] ϕ and ψ meet in three distinct points at one of which they touch.

[(2 1)] ϕ and ψ have contact of the third order.

[3] ϕ and ψ have contact of the second order.

In all of the above cases ϕ, as well as ψ, is non-singular.

In the next five cases, ϕ consists of a pair of distinct straight lines.

* We assign to the constants c_i and k_i, in formula (3) of the last section, values so chosen that the loci $\phi = 0$, $\psi = 0$ are real when the constants λ_i are real. This is, of course, not essential, since we are not concerned with questions of reality.

† In order to verify the statements made below, the reader should have some knowledge of the theory of the contact of conics; cf. for instance Salmon's *Conic Sections*, Chapter XIV., pages 232–238.

$[1\,1\,\overset{0}{1}]$ ϕ and ψ intersect in four distinct points.

$[(1\,1)\,\overset{0}{1}]$ Both of the lines of which ϕ consists touch ψ.

$[2\,\overset{0}{1}]$ One of the lines of which ϕ consists touches ψ, while the other cuts it in two points distinct from the point of contact of the first.

$[\overset{0}{2}\,1]$ The two lines of which ϕ consists intersect on ψ, and neither of them touches ψ.

$[\overset{0}{3}]$ The two lines of which ϕ consists intersect on ψ, and one of them touches ψ.

In the next two cases, ϕ consists of a single line.

$[(\overset{0}{1}\,\overset{0}{1})\,1]$ The line ϕ meets ψ in two distinct points.

$[(\overset{0}{2}\,\overset{0}{1})]$ The line ϕ touches ψ.

Finally we have the case:

$[(\overset{0}{1}\,\overset{0}{1}\,\overset{0}{1}])$ Here $\phi \equiv 0$, and we have no conic other than ψ.

Suppose finally that we wish to classify not *pairs* of quadratic forms or equations but *pencils* of quadratic forms or equations. Consider the pencil of quadratic forms

$$\phi - \lambda\psi$$

where ϕ and ψ are quadratic forms, and ψ is non-singular, and suppose that the elementary divisors of the pair of forms ϕ, ψ are given by formula (1) above. The question presents itself whether, if, in place of the forms ϕ, ψ, we take any other two forms of the pencil

$$\phi_1 \equiv \phi - \mu\psi, \qquad \psi_1 \equiv \phi - \nu\psi,$$

the constants μ, ν being so chosen that $\mu \neq \nu$ and that ψ_1 is non-singular, the pair of forms ϕ_1, ψ_1, will have these same elementary divisors (1). If this were the case, we could properly speak of (1) as the elementary divisors of the pencil. This, however, is *not* the case, and *the pencil of quadratic forms cannot properly be said to have elementary divisors.**

* We here regard the pencil as merely an aggregate of an infinite number of quadratic forms, namely, all the forms which can be obtained from the expression $\phi - \lambda\psi$ by giving to λ different values. In this sense we cannot speak of the elementary divisors of the pencil. If, however, we wish to regard the polynomial in the x's and λ, $\phi - \lambda\psi$, as the pencil, we may speak of its elementary divisors, meaning thereby simply what we have called the elementary divisors of the pair of forms ϕ, ψ.

There is, however, a simple relation between the elementary divisors of two pairs of forms taken from the same pencil. In order to show this, let us determine the elementary divisors of the pair of forms ϕ_1, ψ_1, above. For this purpose consider the expression $\phi_1 - \lambda\psi_1$, which, when $\lambda \neq 1$, may be written

$$\phi_1 - \lambda\psi_1 \equiv (1-\lambda)[\phi - \lambda'\psi] \tag{3}$$

where
$$\lambda' = \frac{\mu - \nu\lambda}{1 - \lambda}.$$

Now suppose, as above, that $\lambda - \alpha$ is any one of the linear factors of the matrix of $\phi - \lambda\psi$, and that l_i is the exponent of the highest power of $\lambda - \alpha$ which is a factor of all the i-rowed determinants of this matrix. Then any one of the i-rowed determinants of the matrix of $\phi - \lambda'\psi$ may, when $\lambda \neq 1$, be written in the form

$$(\lambda' - \alpha)^{l_i} f(\lambda')$$

where f is a polynomial in λ' of degree not greater than $i - l_i$. Accordingly, by (3), the corresponding i-rowed determinant of the matrix of $\phi_1 - \lambda\psi_1$ may be written

$$[\mu - \nu\lambda - \alpha(1-\lambda)]^{l_i} f_1(\lambda)$$

where f_1 is a polynomial in λ. Thus we see that

$$\left[\lambda - \frac{\alpha - \mu}{\alpha - \nu}\right]^{l_i}$$

is a factor of every i-rowed determinant of the matrix of $\phi_1 - \lambda\psi_1$. Similar reasoning, carried through in the reverse order, shows that this is the highest power of
$$\lambda - \frac{\alpha - \mu}{\alpha - \nu}$$
which is a factor of all these i-rowed determinants. Hence

THEOREM 2. *If the pair of quadratic forms ϕ, ψ, of which the second is non-singular, have the elementary divisors*

$$(\lambda - \lambda_1)^{e_1}, \quad (\lambda - \lambda_2)^{e_2}, \quad \cdots\cdots \quad (\lambda - \lambda_k)^{e_k},$$

and if μ, ν are any two constants distinct from each other and such that ν is distinct from all the constants $\lambda_1, \lambda_2, \cdots \lambda_k$, then the two forms

$$\phi_1 \equiv \phi - \mu\psi, \qquad \psi_1 \equiv \phi - \nu\psi,$$

of which the second will then be non-singular, will have the elementary divisors
$$(\lambda - \lambda_1')^{e_1}, \quad (\lambda - \lambda_2')^{e_2}, \quad \cdots\cdots \quad (\lambda - \lambda_k')^{e_k}$$

where
$$\lambda_i' = \frac{\lambda_i - \mu}{\lambda_i - \nu} \qquad (i = 1, 2, \cdots k).$$

In particular, it will be seen that the two pairs of forms ϕ, ψ and ϕ_1, ψ_1 have the same characteristic $[e_1\, e_2 \cdots e_k]$ even if we put in parentheses to indicate which of the e's correspond to equal λ_i's. The characteristics will not, however, necessarily be the same if we put in small zeros to indicate which of the e's correspond to vanishing λ_i's, since λ_i and λ_i' do not usually vanish together. Accordingly, in classifying pencils of quadratic forms, we may use the characteristic of any pair of distinct forms of the pencil, the second of which is non-singular, *but we must not introduce the small zeros into these characteristics.* This classification, of course, applies only to what may be called *non-singular pencils*, that is, pencils whose forms are not all singular.

It will readily be seen that what has just been said applies without essential change to the case of pencils of homogeneous quadratic equations. We may therefore illustrate it by the classification of non-singular pencils of conics.* We have here six classes of pencils which we characterize as follows:

[1 1 1] The conics all pass through four distinct points.

[(1 1) 1] The conics all pass through two points at which they have double contact with each other.

[(1 1 1)] The conics all coincide.

[2 1] The conics all pass through three points at one of which they touch one another.

[(2 1)] The conics all pass through one point at which they have contact of the third order.

[3] The conics all pass through two points, at one of which they have contact of the second order.

EXERCISES

1. Determine, by the use of elementary divisors, the nature of each of the following pairs of conics:

$$(a)\quad \begin{cases} 3x_1^2 + 7x_2^2 \qquad\quad + 8x_1x_2 - 10x_2x_3 + 4x_1x_3 = 0 \\ 2x_1^2 + 3x_2^2 - x_3^2 + 4x_1x_2 - 6x_2x_3 + 6x_1x_3 = 0. \end{cases}$$

$$(b)\quad \begin{cases} 3x_1^2 - x_2^2 - 3x_3^2 - 3x_1x_2 + 3x_2x_3 + x_1x_3 = 0 \\ 2x_1^2 + x_2^2 - x_3^2 - 2x_1x_2 - 2x_2x_3 + 2x_1x_3 = 0. \end{cases}$$

2. Give a classification of pairs of binary quadratic equations, the second equation of each pair being non-singular, and interpret the work geometrically.

* For a similar classification of pencils of quadrics we refer to p. 46 of Bromwich's book: *Quadratic Forms and their Classification by Means of Invariant Factors.*

106. Conclusion. We wish, in this section, to point out some of the important questions connected with the subject of elementary divisors, which, in order to keep our treatment within proper limits, we have been obliged to leave out of consideration.

If ϕ_1, ψ_1 and ϕ_2, ψ_2 are two pairs of bilinear or quadratic forms of which ψ_1, ψ_2 are non-singular, we have found a method of determining whether these two pairs of forms are equivalent or not. If we use the invariant factors instead of the elementary divisors, our method involves only the use of the rational operations (addition, subtraction, multiplication, and division), and can, therefore, be actually carried through in any concrete case. In fact we have explained in § 93 some really practical methods of determining the invariant factors of a λ-matrix, so that the problem of determining whether or not two pairs of bilinear or quadratic forms, the second form in each pair being non-singular, are equivalent, may be regarded as solved, not merely from the theoretical, but also from the practical point of view.

There is, however, another question here, which we have not treated, namely, if the two pairs of forms turn out to be equivalent, to find a linear transformation which carries over one into the other. This problem, too, we may consider that we have solved from a theoretical point of view; for the proof we have given that if two pairs of forms have the same elementary divisors there exists a linear transformation which carries over one pair of forms into the other, consisted, as will be seen on examination, in actually giving a method whereby such a linear transformation could be determined. In fact, in the case of bilinear forms, the processes involved are, here again, merely the rational processes; so that, given two equivalent pairs of bilinear forms, the second form of each pair being non-singular, we are in a position to find, in any concrete case, linear transformations of the x's and y's which carry over one pair of forms into the other. Even here the arrangement of the work in a practical manner might require further consideration.

In the case of quadratic forms the problem becomes a much more difficult one, inasmuch as the processes involved in the determination of the required linear transformation are no longer rational; cf. the Lemma of § 101. That this is not merely a defect of the method we have used, but is inherent in the problem itself, will be seen by a consideration of simple numerical examples. Let, for instance,

$$\phi_1 \equiv 2x_1^2 + 3x_2^2, \qquad \phi_2 \equiv 2x_1^2 - 3x_2^2,$$
$$\psi_1 \equiv x_1^2 + x_2^2, \qquad \psi_2 \equiv x_1^2 - x_2^2.$$

Here the pairs of forms ϕ_1, ψ_1 and ϕ_2, ψ_2 both have the elementary divisors

$$\lambda - 2, \qquad \lambda - 3,$$

and are therefore equivalent. The linear transformation which carries over one pair of forms into the other cannot, however, be real (and therefore its coefficients cannot be determined rationally from the coefficients of the given forms) since ϕ_1 and ψ_1 are definite, ϕ_2 and ψ_2 indefinite.

We have, therefore, here the problem of devising a practical method of determining a linear transformation which carries over a first pair of quadratic forms into a second given equivalent pair. A method of this sort, which is a practical one when once the elementary divisors have been determined, will be found in Bromwich's book on quadratic forms referred to in the footnote on p. 312.

Another point at which our treatment is incomplete is in the restriction we have always made in assuming that, in the pair of bilinear or quadratic forms ϕ, ψ, the form ψ is non-singular. Although this is the case in many of the most important problems to which one wishes to apply the method of elementary divisors, it is still a restriction which it is desirable to remove. This may be done in part by making use not, as we have done, of the pencil $\phi - \lambda\psi$, but of the more general pencil $\mu\phi - \lambda\psi$, μ and λ being variable parameters. The determinants of the matrix of this pencil are binary forms in (μ, λ), and the whole subject of elementary divisors admits an easy extension to this case, the elementary divisors being now integral powers of linear binary forms. The only case which cannot be treated in this way is that in which not only ϕ and ψ are both singular, but every form of the pencil $\mu\phi - \lambda\psi$ is singular. This *singular case*, which was explicitly excluded by Weierstrass in his original paper, requires a special treatment which has been given by Kronecker. Cf., for the case of quadratic forms, the book of Bromwich already referred to.

Still another question is the application of the method of elementary divisors to the case in which the two forms ϕ, ψ are real, and only real linear transformations are admitted. In the case of bilinear forms, this question presents no serious difficulty; cf. the exercises of §§ 97, 99. In the case of quadratic forms, however, the irrational processes involved in the proof of the Lemma of § 101 introduce an essential difficulty, since they are capable of introducing imaginary quantities. Moreover, this difficulty does not lie merely in the method of treatment. The theorems themselves which we have established do not remain true, as is seen by a reference to the numerical example given earlier in this section for another purpose, where we have two pairs of real quadratic forms which, although they have the same elementary divisors, are not equivalent with regard to real linear transformations.

We must content ourselves with merely mentioning this important subject, and referring, for one of the fundamental theorems, to p. 69 of the book of Bromwich.

For further information concerning the subject of elementary divisors the reader is referred to Muth's *Theorie und Anwendung der Elementartheiler*, Leipzig, Teubner, 1899. In English, the book of Bromwich already referred to and some sections in Mathews' revision of Scott's *Determinants* will be found useful.