SYMMETRIC AND ALTERNATE MATRICES IN AN ARBITRARY FIELD, I*

BY
A. ADRIAN ALBERT

INTRODUCTION

The elementary theorems of the classical treatment of symmetric and alternate matrices may be shown, without change in the proofs, to hold for matrices whose elements are in any field of characteristic not two. The proofs fail in the characteristic two case and the results cannot hold since here the concepts of symmetric and alternate matrices coincide. But it is possible to obtain a unified treatment. We shall provide this here by adding a condition to the definition of alternate matrices which is redundant except for fields of characteristic two. The proofs of the classical results will then be completed by the addition of two necessary new arguments.

The theorems on the definiteness of real symmetric matrices have had no analogues for general fields. They have been based on the property that the sum of any two non-negative real numbers is non-negative. This is equivalent to the property that for every real a and b we have $a^2+b^2=c^2$ for a real c. But $a^2+b^2=(a+b)^2$ in any field of characteristic two and we shall use this fact to obtain complete analogues for arbitrary fields of characteristic two of the usual theorems on the definiteness of real symmetric matrices.

Quadratics forms may be associated with symmetric matrices and the problem of their equivalence is equivalent to the problem of the congruence of the corresponding matrices. This is true except when the field of reference has characteristic two where no matric treatment has been given. We shall associate quadratic forms in this case with a certain type of non-symmetric matrix and shall use our results on the congruence of alternate matrices to obtain a matrix treatment of the quadratic form problem.

The classical theorems† on pairs of symmetric or alternate matrices with complex elements will be shown here to be true for matrices with elements in any algebraically closed field whose characteristic is not two. This will be seen to imply that any two symmetric (or alternate) matrices are orthogonally equivalent if and only if they are similar. But the proof fails for fields of characteristic two.

* Presented to the Society, April 10, 1937; received by the editors April 26, 1937.

† Cf. L. E. Dickson, *Modern Algebraic Theories*, chap. 6. See also the report of C. C. MacDuffee, Ergebnisse der Mathematik, vol. 21 (1933), part 5, for this material as well as the classical results referred to above. The theory will also be found in J. H. M. Wedderburn's *Lectures on Matrices*, American Mathematical Society Colloquium Publications, vol. 17, 1934.

We shall prove the existence of two similar symmetric matrices with elements in a field $\mathfrak{F}$ of characteristic two which are not orthogonally equivalent in the algebraically closed extension of $\mathfrak{F}$. Our treatment of the theory of the orthogonal equivalence in $\mathfrak{F}$ of characteristic two will be rational, that is, no algebraic closure properties of $\mathfrak{F}$ will be assumed. Our formulation will involve a recasting of the theory of similarity of square matrices and then a corresponding parallel treatment of the theory of orthogonal equivalence. In particular we shall obtain a complete determination of the invariant factors of any symmetric matrix in $\mathfrak{F}$ of characteristic two.

The generalized transposition concept called an involution* J of the set of all n-rowed square matrices arises naturally in any rational treatment of orthogonal equivalence. The consequent study of the J-orthogonal equivalence of J-symmetric and J-alternate matrices will be introduced here and various important special types treated in subsequent papers.

I. Congruence theory

1. **Elementary concepts.** Let $\mathfrak{F}$ be an arbitrary field, and let $A=(a_{ij})$ $(i,j=1,\cdots,n)$ be an n-rowed square matrix with elements a_{ij} in $\mathfrak{F}$. We use the customary notation A' for the transpose of A and call A symmetric if $A'=A$. We shall modify the usual definition of alternate matrices however, and make the classically consistent definition:

Definition. *A matrix A is called alternate (or skew-symmetric) if $A'=-A$, and the diagonal elements a_{ii} of A are all zero.*

Notice that when the characteristic of $\mathfrak{F}$ is not two the final part of our definition is redundant. But when the characteristic is two every symmetric matrix has the property $A=-A'$ and the condition will be shown to be essential. We shall also call a matrix A *non-alternate symmetric* if $A=A'$ and A is not alternate according to our definition above. This last condition is redundant except for fields of characteristic two, in which case we are simply assuming that $A=A'$ has a non-zero diagonal element.

Two square matrices A and B with elements in a field $\mathfrak{F}$ are said to be congruent in $\mathfrak{F}$ if there exists a non-singular matrix P with elements in $\mathfrak{F}$ such that $B=PAP'$. It is easy to prove the following lemmas:†

Lemma 1. *Let B be obtained from A by any permutation of its rows followed by the same permutation on the columns. Then B and A are congruent in $\mathfrak{F}$.*

* See the author's paper, *Involutorial simple algebras and real Riemann matrices*, Annals of Mathematics, vol. 36 (1935), pp. 886–964. p. 894 for the definition and some elementary properties of involutions.

† The proofs of these lemmas may be found in the author's *Modern Higher Algebra*, chap. 5, University of Chicago Press, 1937.

LEMMA 2. *Replace the ith row of A by the sum of this row and any linear combination of the remaining rows. Follow this with the corresponding column replacement. Then the resulting matrix B is congruent to A.*

If G and H are square matrices, the matrix

$$A = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} \tag{1}$$

is called the direct sum of G and H. This notion has an immediate generalization to the direct sum

$$\begin{pmatrix} G_1 & & \\ & \ddots & \\ & & G_s \end{pmatrix} \tag{2}$$

of square matrices G_i* called the components of A. It is clear that A is symmetric if and only if its components are symmetric. Also A is alternate if and only if the components of A are all alternate. But in a field of characteristic two a matrix A may be non-alternate symmetric and yet may have some alternate components but has at least one non-alternate component. We shall show later that *such matrices are always congruent to diagonal matrices*, that is, *to direct sums of one-rowed square matrices*. Our proofs will depend partly on the almost trivial consequence of Lemmas 1 and 2.

LEMMA 3. *Let A have the form* (1). *Then A is congruent to*

$$\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix},$$

for any G_0 congruent to G and H_0 congruent to H.

A principal sub-matrix of A is a sub-matrix whose main diagonal is a part of the main diagonal of A. The determinants of principal sub-matrices are called principal minors of A. Then it may easily be shown† that Lemmas 1 and 2 yield the following theorem:

LEMMA 4. *Let G be a non-singular principal sub-matrix of a symmetric or alternate matrix A. Then A is congruent in $\mathfrak{F}$ to*

$$B = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}$$

whose principal minors having $|G|$ as sub-determinant have the same values as those of A.

* We shall henceforth use the notation diag$[G_1, \cdots, G_s]$ for (2) to simplify printing.

† See the author's *Modern Higher Algebra*, chap. 5.

The one-rowed principal minors of A are the elements a_{ii}. When they are all zero and $A' = \pm A$ the two-rowed principal minors are

$$\begin{vmatrix} 0 & a_{ij} \\ a_{ji} & 0 \end{vmatrix} = \pm (a_{ij})^2 = 0 \qquad (i, j = 1, \cdots, n), \tag{3}$$

if and only if $A = 0$. Thus we have the following lemma:

LEMMA 5. *Every symmetric or alternate matrix $A \neq 0$ has a non-zero one- or two-rowed principal minor.*

We shall close our discussion of the tools of our theory by proving the lemma:

LEMMA 6. *Let A and B be r-rowed non-singular matrices, and let*

$$A_0 = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \qquad B_0 = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \tag{4}$$

be n-rowed square matrices. Then $B_0 = PA_0P'$ for a non-singular P if and only if

$$P = \begin{pmatrix} Q & S \\ 0 & R \end{pmatrix}, \qquad QAQ' = B, \tag{5}$$

where Q and R are non-singular.

It is clear that the lemma implies that if $QAQ' = B$ for Q non-singular, we may choose any R and S such that R is non-singular and obtain $PA_0P' = B_0$. These results are an immediate consequence of the computation

$$PA_0P' = \begin{pmatrix} Q & S \\ K & R \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Q' & K' \\ S' & R' \end{pmatrix} = \begin{pmatrix} QAQ' & QAK' \\ KAQ' & KAK' \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \tag{6}$$

if and only if $QAQ' = B$, $QAK' = KAQ' = KAK' = 0$. But B is non-singular and so is A. Hence Q is non-singular, so therefore is AQ', and $K = 0$ is our only condition.

2. **Congruence of alternate matrices.** We shall prove the following theorem:

THEOREM 1. *Every matrix congruent to an alternate matrix is an alternate matrix.*

For let $x_1, \cdots, x_n$ be independent indeterminates over $\mathfrak{F}$

$$x = (x_1, \cdots, x_n), \qquad A = (a_{ij}) \qquad (i, j = 1, \cdots, n). \tag{7}$$

If $a_{ji} = -a_{ij}$, the quadratic form

$$xAx' = \sum_{i,j} x_i a_{ij} x_j = \sum_{i=1}^{n} a_{ii} x_i^2 . \tag{8}$$

When also A is alternate the a_{ii} are all zero and

$$xAx' \equiv 0.$$

We let $B=PAP'$, $y=(y_1, \cdots, y_n)$, and have $B'=-B$. But

$$yBy' = \sum b_{ii}y_i^2 = xAx', \qquad x = yP,$$

so that $yBy'=xAx'\equiv 0$ in the y_i and the $b_{ii}=0$. Hence B is alternate.

The proof is of course unnecessary for fields $\mathfrak{F}$ of characteristic not two. Notice that it implies the theorem:

THEOREM 2. *Every matrix congruent to a non-alternate symmetric matrix is non-alternate symmetric.*

We may also easily prove the following theorem:

THEOREM 3. *Two alternate matrices are congruent in $\mathfrak{F}$ if and only if they have the same rank $2t$.*

For if $A\neq 0$ is alternate, its diagonal elements are all zero. By Lemma 5 A has a two-rowed principal minor

$$E_0 = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \tag{9}$$

with $a\neq 0$. By Lemma 4 the matrix A is congruent to

$$\begin{pmatrix} E_0 & 0 \\ 0 & A_0 \end{pmatrix}.$$

But

$$E = \begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}\begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ a & 0 \end{pmatrix}\begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{10}$$

so that by Lemma 3, A is congruent to

$$\begin{pmatrix} E & 0 \\ 0 & A_0 \end{pmatrix}.$$

We apply Theorem 1 to see that A_0 is alternate. A repetition of this process by the use of Lemma 3 shows that A is congruent to the direct sum

$$\text{diag}\,[E_1, \cdots, E_t, 0], \qquad E_i = E, \tag{11}$$

where $2t$ is the rank of A. Any other alternate matrix of rank $2t$ is congruent to (11) and hence to A, and we have Theorem 3.

The proof given above is valid for general fields only because we have proved Theorem 1. Notice that we have the following consequence:

THEOREM 4. *Every alternate matrix of rank 2t is congruent in* $\mathfrak{F}$ *to*

$$\begin{pmatrix} 0 & -I_t & 0 \\ I_t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{12}$$

where I_t *is the t-rowed identity matrix.*

This new form of Theorem 3 is a consequence of (11) and Lemma 1.

3. **Congruence of non-alternate symmetric matrices.** The Lemmas 1, 3, 4, 5 may be applied to an arbitrary symmetric matrix. They show that every symmetric matrix is congruent to a direct sum of matrices of the forms

$$(a), \qquad \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}, \qquad (a \neq 0 \text{ in } \mathfrak{F}).$$

But as in (10) we have

$$P = \begin{pmatrix} a^{-1} & 0 \\ 0 & 1 \end{pmatrix}, \qquad P\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} P^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{13}$$

By Lemma 1 we have the preliminary reduction given by the following lemma:

LEMMA 7. *Every symmetric matrix is congruent in* $\mathfrak{F}$ *to a matrix*

$$\text{diag } [D, G, 0], \tag{14}$$

where D is a diagonal matrix with elements in $\mathfrak{F}$ *and G is a direct sum of two-rowed matrices*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{15}$$

The reduction of symmetric matrices with elements in a field $\mathfrak{F}$ of characteristic not two is evidently completed by the fact that

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \tag{16}$$

and then that the matrix G is congruent to a diagonal matrix whose diagonal elements are all 2 or -2. But the corresponding transformation in $\mathfrak{F}$ of characteristic two is clearly singular. We complete our reduction in this case by the computation in the following theorem:

THEOREM 5. *Let $a \neq 0$ be in $\mathfrak{F}$ of characteristic two and*

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & a \\ 1 & 1 & a \end{pmatrix}. \tag{17}$$

Then

$$PAP' = aI_3,$$

where I_3 is the three-rowed identity matrix.

The result above seems quite remarkable, as one might expect that the matrix A which has an alternate sub-matrix would not be congruent to a multiple of the identity matrix. We verify the computation in

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & a \\ 1 & 1 & a \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} P' = \begin{pmatrix} a & 0 & 1 \\ a & a & 0 \\ a & a & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & a & a \end{pmatrix}$$

$$= \begin{pmatrix} a & a+a & a+a \\ a+a & a & a+a \\ a+a & a+a & a+a+a \end{pmatrix} = aI_3, \tag{18}$$

since $a+a=0$.

As an immediate consequence of Theorem 5 and Lemma 7 we have the following theorem:

THEOREM 6. *Every non-alternate symmetric matrix is congruent to a diagonal matrix.*

The problem of finding when two diagonal matrices are congruent is not solvable in a general field, as the structural properties of the field are involved in this question. It is usual then to assume that $\mathfrak{F}$ is a field such that for every a of $\mathfrak{F}$ there exists a b in $\mathfrak{F}$ such that $b^2 = a$. Then for this case *two non-alternate symmetric matrices are congruent if and only if they have the same rank.* Finally, as in the classical theory, we may obtain a so-called Kronecker reduction of non-alternate symmetric matrices to diagonal form. The results are nearly the same as in the classical theory; they depend essentially on Lemma 6, and we shall not give the proofs. The only difference is that when in the Kronecker reduction we obtain a matrix

$$\begin{pmatrix} A_{r-2} & 0 \\ 0 & E \end{pmatrix},$$

with E the matrix of (15), we use (16) if $\mathfrak{F}$ does not have characteristic two

and obtain a corresponding pair of diagonal elements 2, -2. But when $\mathfrak{F}$ has characteristic two the Kronecker reduction is completed by the use of Theorem 5. Thus we replace E by

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},$$

where a is any diagonal element obtained at any stage of the reduction.

4. **Definite symmetric matrices.** The field $\mathfrak{R}'$ of all real numbers has the characteristic property that if a and b are in $\mathfrak{R}'$, there exists a c in $\mathfrak{R}'$ such that $c^2 = a^2 + b^2$. This result has the analogue $a^2 + b^2 = (a+b)^2$ in fields of characteristic two, and we shall use these results to obtain an analogue of the concept of definiteness of real symmetric matrices.

DEFINITION. *A symmetric matrix A with elements in a field $\mathfrak{F}$ will be called semi-definite if A is congruent in $\mathfrak{F}$ to*

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

where r is the rank of A, and I_r is the r-rowed identity matrix. A non-singular semi-definite matrix will be called definite.

In the remainder of the section we assume that $\mathfrak{F}$ has characteristic two.* Clearly Theorem 1 implies that alternate matrices are never semi-definite. However by Theorems 4 and 5 the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \tag{19}$$

is semi-definite for every alternate matrix A. We use this result in the proof of the following theorem:

THEOREM 7. *A non-alternate symmetric matrix with elements in $\mathfrak{F}$ of characteristic two is semi-definite if and only if its diagonal elements are the squares of elements of $\mathfrak{F}$.*

If $A \neq 0$ is one-rowed, it has the form (a^2), is congruent to I_1, and is definite. Hence our theorem is true for one-rowed matrices, and we make an induction on the order of A. Let $A = (a_{ij})$ be n-rowed,

$$a_{ij} = a_{ji}, \qquad a_{ii} = a_i^2 \qquad (a_i, a_{ij} \text{ in } \mathfrak{F}).$$

At least one $a_i \neq 0$, and there is no loss of generality if we assume that $a_1 \neq 0$.

* As the field of reference in what follows will sometimes be general and sometimes of characteristic two we shall henceforth designate that it has characteristic two by writing $\mathfrak{F}^{(2)}$ except when the condition is explicitly stated.

Multiply the first row and column of A by a_1^{-1} and replace a_{11} by 1, A by a congruent matrix $B=(b_{ij})$, $b_{ii}=b_i^2$, $b_{ij}=b_{ji}$. We add b_{i1} times the first row of B to its ith row and replace b_{i1} by 0, b_{ii} by $b_{ii}+b_{i1}^2=(b_i+b_{i1})^2=c_i^2$. The corresponding column transformation then replaces b_{1i} by 0, and leaves c_i^2 unaltered. Thus A is congruent to the direct sum

$$\begin{pmatrix} 1 & 0 \\ 0 & C \end{pmatrix}, \tag{20}$$

where the diagonal elements of C are c_i^2, c_i in $\mathfrak{F}^{(2)}$. If C is alternate we have seen that the matrix (20) is semi-definite. Otherwise C is semi-definite by our induction and so is A.

Conversely let A be semi-definite so that

$$A=\begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} P_1' & P_3' \\ P_2' & P_4' \end{pmatrix}=\begin{pmatrix} P_1P_1' & P_1P_3' \\ P_3P_1' & P_3P_3' \end{pmatrix}.$$

Then if $P_1=(d_{ij})$, the diagonal elements of P_1P_1' have the form $\sum_{j=1}^r d_{ij}^2=(\sum_{j=1}^r d_{ij})^2$. Similarly the diagonal elements of P_3P_3' are squares.

COROLLARY. *The principal minors of a semi-definite symmetric matrix are the squares of elements of* $\mathfrak{F}^{(2)}$.

For every principal sub-matrix B of a semi-definite matrix is semi-definite by our theorem. If $|B|=0$, our result is true. Let $|B|\neq 0$, so that B is definite and $B=PP'$, $|B|=|P|^2$ as desired.

The classical result on real symmetric matrices states that A is positive semi-definite if and only if every principal minor of A is non-negative, that is, the square of a real number. Theorem 7 has a weaker hypothesis than the theorem about the real field but, in view of our corollary, the same conclusion. Thus our result is a true analogue of the corresponding real theorem.

In a later section we shall require the theorem:

THEOREM 8. *Let A be a semi-definite matrix with elements in an infinite field $\mathfrak{F}$ of characteristic two. Then there exist quantities a in $\mathfrak{F}$ such that*

$$a^4I+A \tag{21}$$

is definite.

For the diagonal elements of (21) are squares in $\mathfrak{F}^{(2)}$ when this is true of A. The determinant $d(a)$ of (21) is a polynomial in a with leading coefficient unity, and thus there exist infinitely many elements a in $\mathfrak{F}^{(2)}$ such that $d(a)\neq 0$, (21) is definite.

5. **Hermitian matrices.** The classical theory of the conjunctivity of Hermitian matrices with elements in a field $\mathfrak{R}$ holds for arbitrary fields. To verify

this note that the theory has already been shown to be valid for fields of characteristic not two.* Let now $\mathfrak{F}^{(2)}$ be a field of characteristic two, and let $\mathfrak{R}$ be a separable quadratic field over $\mathfrak{F}^{(2)}$. This is the only case that need be considered. Then

$$\mathfrak{R} = \mathfrak{F}^{(2)}(\theta), \qquad \theta^2 = \theta + c, \tag{22}$$

so that every k of $\mathfrak{R}$ has the form

$$k = k_1 + k_2\theta \qquad (k_1, k_2 \text{ in } \mathfrak{F}^{(2)}). \tag{23}$$

The correspondence

$$k \longleftrightarrow \bar{k} = k_1 + k_2(\theta + 1) \tag{24}$$

is an automorphism of $\mathfrak{R}$ over $\mathfrak{F}^{(2)}$ with the property $\bar{\bar{k}} = k$. The theory of Hermitian matrices is then a theory of matrices A with elements in $\mathfrak{R}$. Write

$$A = (a_{ij}), \qquad \overline{A} = (\bar{a}_{ij}). \tag{25}$$

Then $(\overline{A'}) = \overline{A}'$ is called the conjugate transpose of A; and we call a matrix A Hermitian if $A = \overline{A}'$.

Two Hermitian matrices A and B with elements in $\mathfrak{R}$ are said to be conjunctive in $\mathfrak{R}$ if

$$B = DA\overline{D}' \tag{26}$$

for a non-singular D with elements in $\mathfrak{R}$. It is clear that all of the results leading up to our reduction theorem to diagonal form hold.

Our reduction theory is now an immediate consequence of

$$\begin{pmatrix} 1 & \theta \\ 1 & \bar{\theta} \end{pmatrix} E \begin{pmatrix} 1 & 1 \\ \bar{\theta} & \theta \end{pmatrix} = \begin{pmatrix} \theta & 1 \\ \bar{\theta} & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \bar{\theta} & \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{27}$$

since $\bar{\theta} = \theta + 1$, $\theta + \bar{\theta} = 1$. We combine this result with the Hermitian analogue of Lemma 7 and have proved the following theorem:

THEOREM 9. *Every Hermitian matrix is conjunctive in $\mathfrak{R}$ to a diagonal matrix with elements in $\mathfrak{F}^{(2)}$.*

When $\mathfrak{F}$ is a perfect field† we have the usual result:

THEOREM 10. *Any two Hermitian matrices with elements in $\mathfrak{R}$ over a perfect $\mathfrak{F}$ of characteristic two are conjunctive in $\mathfrak{R}$ if and only if they have the same rank.*

* Cf. the author's *Modern Higher Algebra*. The theory is almost exactly the same as in L. E. Dickson's *Modern Algebraic Theories*.

† A perfect field $\mathfrak{F}$ of characteristic two has the property that every a of $\mathfrak{F}$ is equal to b^2, b in $\mathfrak{F}$. Such fields with an over-field $\mathfrak{R} = \mathfrak{F}(\theta)$ exist. For the definition see van der Waerden's *Moderne Algebra*, vol. 1, as well as the author's own *Modern Higher Algebra*.

The proof of the above result is trivial and will be omitted.

II. Quadratic forms

1. The matrices of a quadratic form. Let

$$x = (x_1, \cdots, x_n), \qquad y = (y_1, \cdots, y_n)$$

with independent indeterminates $x_1, \cdots, y_n$ over any field $\mathfrak{F}$. If A is a symmetric matrix with elements in $\mathfrak{F}$ the form xAy is called a symmetric bilinear form in the variables x_i, y_j. A trivial computation then shows that two such bilinear forms are equivalent if and only if their matrices are congruent. Thus the symmetric matrix and the symmetric bilinear form theories are equivalent. Analogous results evidently hold when A is Hermitian, $xA\bar{y}$ is an Hermitian bilinear form. Moreover the theory of Hermitian quadratic forms $xA\bar{x}$ is also equivalent to the theory of Hermitian matrices since we may always choose $x_1, \cdots, x_n, \bar{x}_1, \cdots, \bar{x}_n$ to be independent indeterminates over $\mathfrak{R}$. In the theory of fields of characteristic not two the theory of quadratic forms is equivalent to that of symmetric matrices. This is not true for fields $\mathfrak{F}$ of characteristic two, and we shall develop this theory here. It was developed for the case of a finite field by L. E. Dickson (American Journal of Mathematics, vol. 21 (1899), p. 194), but the results obtained there do not hold for an arbitrary field $\mathfrak{F}$ of characteristic two. We introduce the theory as follows:

Every quadratic form in n independent indeterminates has the form

$$f = f(x_1, \cdots, x_n) = \sum_{i,j}^{1,\cdots,n} x_i a_{ij} x_j. \tag{28}$$

This expression is clearly not unique, but if also

$$f = \sum_{i,j} x_i a_{0ij} x_j, \tag{29}$$

then, by equating the coefficients of x_i^2 and $x_i x_j$, we have

$$a_{ii} = a_{0ii}, \qquad a_{ij} + a_{ji} = a_{0ij} + a_{0ji}. \tag{30}$$

Write $x=(x_1, \cdots, x_n)$, $A=(a_{ij})$, $A_0=(a_{0ij})$, so that

$$f = xAx' = xA_0x'. \tag{31}$$

Then (30) states that $A+A'=A_0+A_0'$, A and A_0 have the same diagonal elements. The matrix A_0-A has zero diagonal elements and $(A_0-A)'=A_0'-A'=A-A_0=-(A_0-A)$. Conversely let $f=xAx'$ and $A_0=A+N$, where $N=-N'$ and the diagonal elements of N are all zero. Then $A_0+A_0'=A+A'$, $f=xA_0x'$. We have proved the following theorem:

THEOREM 11. *Let* $f=xAx'$, $g=xA_0x'$, $A_0=A+N$. *Then* $f=g$ *if and only if* N *is an alternate matrix.*

In the study of quadratic forms with coefficients in a field $\mathfrak{F}$ of characteristic not two it is customary to choose a unique matrix

$$A = (a_{ij}), \qquad a_{ij} = \tfrac{1}{2}(a_{0ij} + a_{0ji}) \tag{32}$$

so that A is symmetric. Then the theory of quadratic forms is equivalent to that of symmetric matrices. This is impossible in $\mathfrak{F}^{(2)}$ of characteristic two both because (32) is impossible and because if $A=(a_{ij})=A'$, then $f=xAx'=\sum_{i=1}^{n}a_{ii}x_i^2$. It is now natural to make the following definition:

DEFINITION. *A quadratic form* f *with coefficients in a field* $\mathfrak{F}$ *of characteristic two is called diagonal or non-diagonal according as* f *does or does not have the expression*

$$f = a_1x_1^2 + \cdots + a_nx_n^2 \qquad (a_i \text{ in } \mathfrak{F}). \tag{33}$$

We shall now assume that the characteristic of $\mathfrak{F}^{(2)}$ is two. By a non-singular linear transformation

$$x_i = \sum_{i=1}^{n} y_id_{ij} \cdot \tag{34}$$

with matrix $D=(d_{ij})$ we carry a quadratic form f into what is called an equivalent form g. If A is a matrix, $f=xAx'$, and $y=(y_1, \cdots, y_n)$, then $x=yD$,

$$g = yDA(yD)' = yDAD'y'. \tag{35}$$

Hence DAD' is one possible matrix of g.

Consider in particular the case where

$$f = x_1^2 + x_2^2, \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{36}$$

and write

$$x_1 = y_1 + y_2, \qquad x_2 = y_2, \qquad x = (y_1, y_2)\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = yD. \tag{37}$$

Then in a field $\mathfrak{F}^{(2)}$ we have

$$y_1 = x_1 + x_2, \qquad f = g = y_1^2, \tag{38}$$

since $(x_1+x_2)^2=x_1^2+x_2^2$. A matrix of g is

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{39}$$

which is not merely incongruent to the two-rowed identity matrix A but does not even have the same rank. However

$$DAD' = DD' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = B + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{40}$$

We now consider the general theory of quadratic forms in the light of the above example. Write

$$f = \sum_{i \leqq j}^{i,j=1,\cdots,n} x_i a_{ij} x_j. \tag{41}$$

Then the matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2\,n-1} & a_{2n} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & \cdots & 0 & a_{n\,n} \end{pmatrix} \tag{42}$$

is uniquely determined by f. We make a non-singular transformation $x = yD$ and carry f into

$$g = yBy', \tag{43}$$

where B has the form $B = (b_{ij})$, $b_{ij} = 0$ for $i > j$. Then

$$B - DAD' = N \tag{44}$$

is alternate by Theorem 11.

THEOREM 12. *Let f be a quadratic form with unique matrix A of* (42). *Then a non-singular transformation with matrix D carries f into an equivalent form with matrix $B = DAD' + N$ where N is the unique alternate matrix chosen so that the elements below the diagonal in B are all zero.*

2. **Diagonal quadratic forms.** If $A = A'$, then $B = DAD' + N$ is symmetric. Hence $g = yBy'$ is a diagonal quadratic form. This gives the following theorem:

THEOREM 13. *Every quadratic form equivalent to a diagonal quadratic form is a diagonal quadratic form in $\mathfrak{F}^{(2)}$.*

A simpler proof is given as follows. We let $f = a_1x_1^2 + \cdots + a_nx_n^2$ and use (34). Then

$$f = g(y_1, \cdots, y_n) = \sum_i a_i \Big(\sum_j d_{ij} y_j \Big)^2 = \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} a_i d_{ij}^2 \Big) y_j^2 = \sum_{j=1}^{n} b_j y_j^2,$$

since $(a+b)^2 = a^2 + b^2$ in $\mathfrak{F}^{(2)}$. However we have now proved the theorem:

THEOREM 14. *Two diagonal quadratic forms* $f=\sum_{i=1}^{n}a_ix_i^2$ *and* $g=\sum_{i=1}^{n}b_iy_i^2$ *are equivalent in* $\mathfrak{F}^{(2)}$ *if and only if the coefficients* b_j *are representable as values* $f(d_{1j}, \cdots, d_{nj})$ *of* f *such that the corresponding determinant* $|d_{ij}| \neq 0$.

We cannot go into questions of representation in a general field $\mathfrak{F}$. However we do have the following theorem:

THEOREM 15. *Let* $\mathfrak{F}^{(2)}$ *be perfect. Then any two non-zero diagonal quadratic forms in n indeterminates are equivalent.*

For

$$f = \sum_{i=1}^{n} a_ix_i^2 = \sum_{i=1}^{n} (\alpha_ix_i)^2 = \left(\sum_{i=1}^{n} \alpha_ix_i\right)^2 = z_1^2,$$

where $\alpha_i^2 = a_i$, α_i in $\mathfrak{F}$. Similarly

$$g = \sum b_iy_i^2 = z_1^2.$$

Then both f and g are equivalent to the same quadratic form and hence to each other.

3. **Non-diagonal quadratic forms.** Non-diagonal quadratic forms f are not equivalent to diagonal quadratic forms by Theorem 13. Then $f=xAx'$, $A+A'\neq 0$. By Theorem 12 we have $B=DAD'+N$, $N+N'=0$, and $B+B'=D(A+A')D'$. The matrix $A+A'$ is alternate and has rank $2r$ by Theorem 3. Moreover we may always carry f into a form g whose matrix B has the property

$$D(A + A')D' = B + B' = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad W_r = \begin{pmatrix} 0 & I_r \\ I_r & 0 \end{pmatrix}, \tag{45}$$

with I_r the r-rowed identity matrix. Write

$$B = \begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix},$$

where B_1 has $2r$ rows and columns. Since B has elements below the diagonal all zero this is true of B_1 and B_3. But

$$B + B' = \begin{pmatrix} B_1 + B_1' & B_2 \\ B_2' & B_3 + B_3' \end{pmatrix} = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix},$$

so that $B_2=0$, $B_1+B_1'=W_r$. Then

$$B = \begin{pmatrix} G_1+\Gamma_r & 0 \\ 0 & G_2 \end{pmatrix}, \qquad \Gamma_r = \begin{pmatrix} 0 & I_r \\ 0 & 0 \end{pmatrix}, \tag{46}$$

and G_1 and $G_2 = B_3$ are diagonal matrices. It is clear that r is an invariant of f, and we have proved the following theorem:

THEOREM 16. *Every non-diagonal quadratic form is equivalent to a form*

$$f = \sum_{i=1}^{n} a_i x_i^2 + (x_1 x_{r+1} + \cdots + x_r x_{2r}). \tag{47}$$

Moreover two forms f and g are equivalent only if they have the same rank invariant r.

Let f of the form (47) go into g of the form (47) under a transformation of matrix D. Then D leaves $A + A'$ invariant. We suppose that

$$A = \begin{pmatrix} G_1 + \Gamma_r & 0 \\ 0 & G_2 \end{pmatrix}, \qquad A + A' = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix}, \tag{48}$$

and by Lemma 6 see that

$$D(A + A')D' = D\begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix} D' = \begin{pmatrix} W_r & 0 \\ 0 & 0 \end{pmatrix} \tag{49}$$

if and only if R is a non-singular matrix, where

$$D = \begin{pmatrix} H & K \\ 0 & R \end{pmatrix}, \qquad HW_rH' = W_r. \tag{50}$$

Then

$$B + N = DAD' = \begin{pmatrix} H(G_1 + \Gamma_r)H' + KG_2K' & KG_2R' \\ RG_2K' & RG_2R' \end{pmatrix} \tag{51}$$

with N alternate, and

$$B = \begin{pmatrix} G_{10} + \Gamma_r & 0 \\ 0 & G_{20} \end{pmatrix}, \tag{52}$$

where G_{20} is a non-alternate symmetric matrix whose diagonal elements coincide with those of RG_2R'. This gives the theorem:

THEOREM 17. *Let*

$$f = \left(\sum_{i=2}^{2r} a_i x_i^2 + x_1 x_{r+1} + \cdots + x_r x_{2r} \right) + \sum_{i=2r+1}^{n} a_i x_i^2$$

and

$$g = \left(\sum_{i=1}^{2r} b_i y_i^2 + y_1 y_{r+1} + \cdots + y_r y_{2r} \right) + \sum_{i=2r+1}^{n} b_i y_i^2 .$$

Then f and g are equivalent in $\mathfrak{F}^{(2)}$ only if the forms

$$\sum_{i=2r+1}^{n} a_i x_i^2, \qquad \sum_{i=2r+1}^{n} b_i y_i^2 \tag{53}$$

are equivalent.

We next write

$$H = \begin{pmatrix} L & U \\ V & M \end{pmatrix}, \tag{54}$$

where L, M, U, V are r-rowed square matrices. Then

$$\begin{aligned} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} &= H \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} H' = \begin{pmatrix} U & L \\ M & V \end{pmatrix} \begin{pmatrix} L' & V' \\ U' & M' \end{pmatrix} \\ &= \begin{pmatrix} UL' + LU' & UV' + LM' \\ VU' + ML' & MV' + VM' \end{pmatrix}, \end{aligned} \tag{55}$$

if and only if

$$UL' = (UL')', \qquad MV' = (MV')', \qquad UV' + LM' = I, \tag{56}$$

where I is the r-rowed identity matrix. Then if

$$G_1 = \begin{pmatrix} C & 0 \\ 0 & J \end{pmatrix}, \tag{57}$$

where C and J are r-rowed diagonal matrices, we have

$$\begin{aligned} HG_1H' &= \begin{pmatrix} L & U \\ V & M \end{pmatrix} \begin{pmatrix} C & 0 \\ 0 & J \end{pmatrix} H' = \begin{pmatrix} LC & UJ \\ VC & MJ \end{pmatrix} \begin{pmatrix} L' & V' \\ U' & M' \end{pmatrix} \\ &= \begin{pmatrix} LCL' + UJU' & 0 \\ 0 & VCV' + MJM' \end{pmatrix} + N_0, \end{aligned} \tag{58}$$

where N_0 is alternate. Also

$$\begin{aligned} H\Gamma_r H' &= \begin{pmatrix} L & U \\ V & M \end{pmatrix} \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} H' = \begin{pmatrix} 0 & L \\ 0 & V \end{pmatrix} \begin{pmatrix} L' & V' \\ U' & M' \end{pmatrix} = \begin{pmatrix} LU' & LM' \\ VU' & VM' \end{pmatrix} \\ &= \begin{pmatrix} LU' & 0 \\ 0 & VM' \end{pmatrix} + \Gamma_r + \begin{pmatrix} 0 & UV' \\ VU' & 0 \end{pmatrix}, \end{aligned} \tag{59}$$

since $I + LM' = UV'$. The last matrix in (59) is alternate; hence the diagonal matrix G_{10} has the same diagonal elements as

$$KG_2K' + \begin{pmatrix} LCL' + UJU' + LU' & 0 \\ 0 & VCV' + MJM' + VM' \end{pmatrix}. \tag{60}$$

The quadratic form with matrix (60) is a quadratic form $\sum_{i=1}^{2r} b_i y_i^2$. The part corresponding to the arbitrary matrix K is clearly

$$\sum_{j=1}^{2r}\Big(\sum_{i=2r+1}^{n} a_i k_{ij}^2\Big) y_j^2, \tag{61}$$

for arbitrary k_{ij} and a_i given as in Theorem 17. We now let f and g be two arbitrary forms satisfying the necessary conditions of Theorems 16, 17 so that we may write

$$f = \sum_{i=2r+1}^{n} a_i x_i^2 + \sum_{i=1}^{r} (a_i x_i^2 + x_i x_{i+r} + a_{i+r} x_{i+r}^2), \tag{62}$$

$$g = \sum_{i=2r+1}^{n} a_i y_i^2 + \sum_{i=1}^{r} (b_i y_i^2 + y_i y_{i+r} + b_{i+r} y_{i+r}^2). \tag{63}$$

Also write

$$z = (y_1, \cdots, y_r), \qquad w = (y_{r+1}, \cdots, y_{2r}),$$

$$C = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_r \end{pmatrix}, \qquad J = \begin{pmatrix} a_{r+1} & & \\ & \ddots & \\ & & a_{2r} \end{pmatrix}. \tag{64}$$

Then we have proved the following theorem:

THEOREM 18. *The forms f and g are equivalent in $\mathfrak{F}^{(2)}$ if and only if there exist r-rowed square matrices L, M, U, V such that LU' and VM' are symmetric, $UV'+LM'=I_r$, and the quadratic form*

$$\sum_{i=1}^{2r} b_i y_i^2 - z(LCL' + UJU' + LU')z' - w(VCV' + MJM' + VM')w' \tag{65}$$

may be expressed as (61) *for k_{ij} in $\mathfrak{F}$.*

It does not seem possible to materially simplify (65) for an arbitrary field $\mathfrak{F}^{(2)}$. However we may obtain an analogue of the classical complex number case by proving the theorem:

THEOREM 19. *Every non-diagonal quadratic form with elements in an algebraically closed field $\mathfrak{F}^{(2)}$ of characteristic two and rank invariant r is equivalent to one and only one of the forms*

$$x_1 x_{r+1} + \cdots + x_r x_{2r}, \tag{66}$$

$$x_1 x_{r+1} + \cdots + x_r x_{2r} + x_{2r+1}^2. \tag{67}$$

Hence two non-diagonal quadratic forms are equivalent in $\mathfrak{F}^{(2)}$ if and only if they have the same rank invariant r and the same type (66) *or* (67).

For if $\mathfrak{F}^{(2)}$ is algebraically closed, we use Theorem 15 and transform the form (53) into x_{2r+1}^2 if it is not identically zero. By Theorem 16 our above theorem is true if we can show that the form

$$\sum_{i=1}^{r} (a_i x_i^2 + x_i x_{i+r} + a_{i+r} x_{i+r}^2)$$

is equivalent to (66). This is clearly true if it can be proved that the two forms

$$ax^2 + xy + by^2, \qquad XY$$

are equivalent for every a and b of $\mathfrak{F}^{(2)}$. If $a=b=0$ the result is trivial. We may therefore assume $a \neq 0$ without loss of generality. The equation

$$\omega^2 + \omega + ab = 0$$

has two distinct roots $\lambda, \lambda+1$ in $\mathfrak{F}$. Thus

$$(\omega - \lambda)(\omega - \lambda - 1) \equiv \omega^2 + \omega + ab.$$

Put $\omega = axy^{-1}$ and multiply by $a^{-1}y^2$ to obtain

$$a^{-1}y^2(a^2x^2y^{-2} + axy^{-1} + ab) \equiv ax^2 + xy + by^2 \equiv XY,$$

where

$$X = a^{-1}y(\omega - \lambda) = a^{-1}y(axy^{-1} - \lambda) = x - a^{-1}\lambda y$$

and

$$Y = y(\omega - \lambda - 1) = y(axy^{-1} - \lambda - 1) = ax - y(\lambda + 1).$$

The determinant of the transformation is

$$\begin{vmatrix} 1 & a \\ a^{-1}\lambda & \lambda + 1 \end{vmatrix} = 1,$$

and we have proved our theorem.

III. Pairs of symmetric matrices

1. **The problem.** The theory of the congruence of pairs of symmetric matrices has been studied only in the classical case of matrices with complex elements. The results hold however for matrices with elements in any algebraically closed field whose characteristic is not two. It will be our purpose in the present chapter to develop these results and to show precisely where the classical proofs fail.

2. **The nth roots of a matrix.** Let $\mathfrak{J}$ be an integral domain with the property that every two elements of $\mathfrak{J}$ have a greatest common divisor in $\mathfrak{J}$ which

is linearly expressible in terms of them. Then the congruence

$$ax \equiv b \quad (c) \tag{68}$$

has a solution x for every a prime to c.

We consider a prime (or irreducible) element π of $\mathfrak{J}$ and study the congruence

$$f(x) \equiv 0 \quad (\pi^e). \tag{69}$$

Here $f(x)$ is a polynomial in an indeterminate x with coefficients in $\mathfrak{J}$, and e is any positive integer. It is clear that (69) implies that, in particular,

$$f(x) \equiv 0 \quad (\pi) \tag{70}$$

must have a solution. Let then $f(x) \equiv 0\ (\pi^{e-1})$ have a solution x_0, and write $x = x_0 + y\pi^{e-1}$. The Taylor expansion of $f(x)$ then implies that (69) is satisfied if and only if $f(x_0) + yf'(x_0)\pi^{e-1} \equiv 0\ (\pi^e)$, that is,

$$yf'(x_0) + q \equiv 0 \quad (\pi), \tag{71}$$

where $\pi^{e-1}q = f(x_0)$. It is clear that (71) has a solution if $f'(x_0) \not\equiv 0\ (\pi)$.* This gives the lemma:

LEMMA 8. *Let $f(x) \equiv 0\ (\pi)$ have a solution x_0 such that $f'(x_0) \not\equiv 0\ (\pi)$. Then there exists a solution of* (69) *for every e.*

The result of our lemma may now be applied to non-singular matrices A with elements in a field $\mathfrak{F}$. Suppose that

$$g = g(\xi) = [\pi_1(\xi)]^{e_1} \cdots [\pi_t(\xi)]^{e_t}$$

is a factorization of the minimum function of A into powers of distinct irreducible functions $\pi_i(\xi)$, and let $\mathfrak{J}$ be the integral domain of all polynomials in the indeterminate ξ. We consider the congruence $x^n \equiv \xi\ (g)$. This congruence may be easily shown to be solvable modulo g if and only if it is solvable modulo $\pi_i^{e_i}$ for $i = 1, \cdots, t$. Now the derivative of $x^n - \xi$ is $nx^{n-1} \equiv 0\ (\pi_i)$ if and only if $nx \equiv 0\ (\pi_i)$. But $x^n - \xi \equiv 0\ (\pi_i)$ so that either $\xi \equiv 0\ (\pi_i)$ or $n \equiv 0\ (\pi_i)$. Then $\xi \equiv 0\ (\pi_i)$ means that ξ is a factor of $g(\xi)$ which is impossible when A is non-singular. Also $n \not\equiv 0\ (\pi_i)$ for π_i irreducible unless the characteristic of $\mathfrak{F}$ divides n. Hence the conditions of our lemma reduce to $x^n - \xi \equiv 0\ (\pi_i)$. The equation $\pi_i(\xi) = 0$ defines a field $\mathfrak{F}(\xi_i)$ over $\mathfrak{F}$ equivalent to the set of all residue classes modulo π_i. Then $x^n - \xi \equiv 0\ (\pi_i)$ if and only if $\xi_i = x_i^n$ for x_i in

* This is the standard technique for the study of congruences $f(x) \equiv 0 \pmod{p}$ in the theory of numbers (as in L. E. Dickson's *Introduction to the Theory of Numbers*, p. 16, ex. 4). We are using the analogous property of the polynomial domain (cf. Lemma 35.21, p. 60, of MacDuffee's tract on *The Theory of Matrices*).

$\mathfrak{F}(\xi_i)$, and when this condition is satisfied we have $[x(\xi)]^n-\xi\equiv 0\ (g(\xi))$, $[x(A)]^n=A$. We have proved the following theorem:

THEOREM 20. *Let n be an integer not divisible by the characteristic of $\mathfrak{F}$, A be a non-singular matrix whose minimum function $g(\xi)$ has the distinct irreducible factors $f_i(\xi)$, $\mathfrak{F}_i=\mathfrak{F}(\xi_i)$ be the corresponding algebraic fields over $\mathfrak{F}$. Then there exists a polynomial $P(A)$ whose nth power is A if and only if the equations*

$$\xi_i = x^n \tag{72}$$

have solutions x_i in $\mathfrak{F}(\xi_i)$.

We next consider singular matrices A. Then $g(\xi)=\xi^r g_0(\xi)$, where $g_0(\xi)\not\equiv 0$ (ξ) and $r\geqq 1$. Thus $x^n\equiv\xi\ (g)$ implies that $x^n-\xi\equiv 0\ (\xi^r)$, $x=\xi x_1$, $\xi^{n-1}x_1^n\equiv 1$ (ξ^{r-1}) which is impossible for $r>1$. This gives the theorem:

THEOREM 21. *Let $n>1$, A be a singular matrix whose minimum function $g(\xi)$ is divisible by ξ^2. Then there exists no polynomial in A whose nth power is A.*

As an immediate corollary of the above argument we have the following theorem:

THEOREM 22. *Let $n>1$, A be a singular matrix whose minimum function has the form $\xi g(\xi)$ where $g(\xi)$ is not divisible by ξ and has irreducible factors with the properties of Theorem 20. Then there exists a polynomial $P(A)$ with coefficients in $\mathfrak{F}$ whose nth power is A.*

Theorem 20 may be applied to prove the following theorem:

THEOREM 23. *Let $\mathfrak{F}$ be an algebraically closed field, n be an integer not divisible by the characteristic of $\mathfrak{F}$, A be a non-singular matrix. Then there exists a polynomial in A with coefficients in $\mathfrak{F}$ whose nth power is A.*

For the fields $\mathfrak{F}_i$ of Theorem 20 are all equal to $\mathfrak{F}$. Moreover $x^n=\xi_i$ has a root x_i in $\mathfrak{F}$; and we have our theorem.

Theorem 23 does not hold for arbitrary matrices A if the characteristic of $\mathfrak{F}$ divides n. For let $\mathfrak{F}$ have characteristic p and B be the p-rowed square matrix all of whose elements are unity. A trivial computation shows that $B^2=0$, $B^p=0$. The matrix $A=I+B$ is non-singular since $A^p=I+B^p=I$. Now any polynomial in A has the form a_0+a_1B, a_0 and a_1 in $\mathfrak{F}$. Then $(a_0+a_1B)^p=a_0^p\neq A$ for any a_0, a_1. If $n=pq$ we have $(a_0+a_1B)^n=a_0^n\neq A$. This proves the following theorem:

THEOREM 24. *Let the characteristic of $\mathfrak{F}$ divide n. Then there exist non-singular matrices A such that no polynomial $\phi(A)$ has the property $[\phi(A)]^n=A$.*

3. **Equivalence of pairs of symmetric and alternate matrices.** The result of Theorem 23 may be applied to the theory of equivalence of pairs of sym-

metric matrices. We assume that $\mathfrak{F}$ is algebraically closed as is usual in the classical case. Suppose now that P and Q are non-singular matrices such that

$$PAQ = B, \tag{73}$$

where A and B are either both symmetric or both alternate. Then $(PAQ)' = Q'A'P' = B'$ so that

$$PAQ = Q'AP', \qquad AG' = GA, \tag{74}$$

where

$$G = P^{-1}Q' \tag{75}$$

is non-singular. We now assume that the characteristic of $\mathfrak{F}$ is not two and use Theorem 23 to obtain a polynomial $\phi(G)$ such that

$$[\phi(G)]^2 = G.$$

Now $AG' = GA$ implies that $AG'^2 = G^2A$, $A(G')^k = G^kA$, $A\phi(G') = \phi(G)A$ for any polynomial in G with coefficients in $\mathfrak{F}$. We use the $\phi(G)$ above and have

$$A = \phi(G)A\phi(G')^{-1}, \qquad G^{-1}A = \phi(G)^{-1}A[\phi(G)^{-1}]' = Q'^{-1}PA.$$

Write $H = Q'[\phi(G)]^{-1}$ and obtain

$$HAH' = Q'[\phi(G)]^{-1}A[\phi(G)^{-1}]'Q = Q'Q'^{-1}PAQ = PAQ.$$

We have proved the following theorem:

THEOREM 25. *Let $\epsilon = \pm 1$, $A' = \epsilon A$, P and Q be non-singular matrices such that $(PAQ)' = \epsilon(PAQ)$. Define*

$$G = P^{-1}Q', \qquad H = Q'[\phi(G)]^{-1}, \tag{76}$$

where $\phi(G)$ is a polynomial in G determined so that its square is G. Then $HAH' = PAQ$.

It is clear that the proof we have given of Theorem 25 is not valid for fields of characteristic two. The result itself does not hold. We shall prove this in an example given later.

DEFINITION. *Let A, B, C, D be n-rowed square matrices with elements in $\mathfrak{F}$. Then we say that the pairs (A, B) and (C, D) are equivalent in $\mathfrak{F}$ if there exist non-singular matrices P and Q with elements in $\mathfrak{F}$ such that*

$$PAQ = C, \qquad PBQ = D.$$

We also call (A, B) and (C, D) congruent in $\mathfrak{F}$ if there exists a non-singular matrix H with elements in $\mathfrak{F}$ such that

$$HAH' = C, \qquad HBH' = D.$$

The notion of congruence of pairs may be applied to either symmetric or alternate matrices. It is clear that if A is symmetric, C must be symmetric, and when A is alternate C must be alternate. We have similar necessary conditions on B and D. Then Theorem 25 implies the following theorem:

THEOREM 26. *Let $\mathfrak{F}$ be an algebraically closed field of characteristic not two. Then two pairs of alternate or symmetric matrices satisfying the trivial necessary conditions above are congruent if and only if they are equivalent.*

Conditions that two pairs of matrices be equivalent are expressed in the literature in terms of the invariant factors of the matrices $Ax+B$, $Cx+B$, where x is an indeterminate over $\mathfrak{F}$. This theory holds for an arbitrary $\mathfrak{F}$, and we shall not discuss these known results.

4. **Elementary applications.** There are two simple consequences of the theory of pairs of matrices which seem interesting and appear never to have been noted in the literature. We first let A, B be two non-singular matrices with elements in an algebraically closed field $\mathfrak{F}$ of characteristic not two, and ask the question as to when they are congruent. The answer is given as the corollary:

COROLLARY I. *Write $A=A_1+A_2$, $B=B_1+B_2$ where $A_1=\frac{1}{2}(A+A')\neq 0$ and $B_1=\frac{1}{2}(B+B')\neq 0$ are symmetric, while $A_2=\frac{1}{2}(A-A')\neq 0$ and $B_2=\frac{1}{2}(B-B')\neq 0$ are alternate matrices. Then A and B are congruent if and only if*

$$Ax+A_1, \qquad Bx+B_1$$

have the same invariant factors.

For the theory of invariant factors states that (A, A_1) and (B, B_1) are equivalent if and only if $Ax+A_1$ and $Bx+B_1$ have the same invariant factors. Then $HAH'=B$ implies that $HA'H'=B'$, $H(A\pm A')H'=B\pm B'$, $HA_1H'=B_1$ so that (A, A_1) and (B, B_1) are equivalent when A and B are congruent. Conversely let (A, A_1) and (B, B_1) be equivalent so that

$$PAQ=B, \qquad PA_1Q=B_1.$$

Then $P(A-A_1)Q=B-B_1=B_2=PA_2Q$ and the pairs (A_1, A_2) and (B_1, B_2) are equivalent. By Theorem 26 they are congruent, $HA_1H'=B_1$, $HA_2H'=B_2$, so that $HAH'=H(A_1+A_2)H'=B_1+B_2=B$ as desired.

We next restrict our attention to the field $\mathfrak{C}$ of complex numbers. Corollary I then states that two Hermitian non-singular matrices

$$A=A_1+A_2i, \qquad B=B_1+B_2i \tag{77}$$

for real A_1, A_2, B_1, B_2 are congruent in $\mathfrak{C}$ if and only if the matrices $Ax+A_1$, $Bx+B_1$ have the same invariant factors. An analogous question arises when

we consider two symmetric matrices (77). Then A_1, A_2, B_1, B_2 are all real symmetric matrices. We then ask for necessary and sufficient conditions that $PA\overline{P}'=B$ for a non-singular P with complex elements. It is clear that then $P\overline{A}'\overline{P}'=\overline{B}'$, $P(A\pm\overline{A}')\overline{P}'=B\pm\overline{B}'$, so that

$$PA_1\overline{P}' = B_1, \qquad PA_2\overline{P}' = B_2.$$

The classical analogue of Theorem 26 states that two pairs of Hermitian complex matrices (A_1, A_2) and (B_1, B_2) are conjunctive if and only if they are equivalent. They are clearly equivalent if and only if (A, A_1) and (B, B_1) are equivalent and we have the statement:

COROLLARY II. *Two non-singular complex symmetric matrices $A=A_1+A_2i$, $B=B_1+B_2i$ for real symmetric A_1, A_2, B_1, B_2 are conjunctive if and only if*

$$Ax + A_1, \qquad Bx + B_1$$

have the same invariant factors.

5. **Orthogonal equivalence.** The theory of the orthogonal equivalence of two symmetric matrices in an algebraically closed field $\mathfrak{F}$ of characteristic not two is equivalent to the theory of the congruence of pairs (A, B), (C, D) of which A and C are non-singular. The concept of orthogonal equivalence is defined as follows. Let D be a square matrix with elements in a field $\mathfrak{F}$. Then we call D orthogonal if $DD'=I$ is the identity matrix. Clearly then $DD'=D'D$. We consider two symmetric or alternate matrices A, B and call them orthogonally equivalent if $DAD'=B$ for an orthgonal D. Since $B=DAD^{-1}$ is similar to A when A and B are orthogonally equivalent, the condition of similarity is a necessary condition. However we may actually prove the following theorem:

THEOREM 27. *Let A and B be both symmetric or both alternate matrices with elements in an algebraically closed field $\mathfrak{F}$ whose characteristic is not two. Then A and B are orthogonally equivalent if and only if they are similar.*

For let $PAP^{-1}=B$. Then $PIP^{-1}=I$ and the pairs (I, A) and (I, B) are equivalent. By Theorem 26 they are congruent, $DID'=I$, $DAD'=B$, D is orthogonal. The converse is trivial.

Theorem 27 is an immediate consequence of Theorem 26. However the connection between the two results is even closer than is indicated by this fact. For assume that the result of Theorem 27 has been proved true independently of Theorem 26. Let now (A, B), (C, G) be any two pairs of matrices such that

$$A' = A, \qquad C' = C, \qquad B' = \epsilon B, \qquad G' = \epsilon G,$$

where $\epsilon = \pm 1$. Assume also that A and C are non-singular. Then (A, B) and (C, G) are congruent if and only if $Ax+B$ and $Cx+G$ have the same invariant factors. This statement of Theorem 26 follows from Theorem 27. We use the existence of a matrix Q such that $QCQ' = I$. Then the invariant factors of $Ax+B$ and $xI+PBP'$ are the same. Similarly those of $Cx+G$ and $xI+QGQ'$ are the same. But when $Ax+B$ and $Cx+G$ have the same invariant factors, so do $xI+QGQ', xI+PBP', QGQ'$ is orthogonally equivalent to PBP'. We let

$$QGQ' = DPBP'D', \qquad DD' = I$$

and have

$$HAH' = Q^{-1}DPAP'D'Q'^{-1} = Q^{-1}Q'^{-1},$$
$$H = Q^{-1}DP, \qquad G = HBH', \qquad C = Q^{-1}Q'^{-1} = HAH'.$$

The above proof cannot, of course, be carried out for fields $\mathfrak{F}$ which are not algebraically closed. For it depends essentially upon the property that $QAQ' = I$ for every non-singular A with elements in $\mathfrak{F}$. However it is clear that in an arbitrary $\mathfrak{F}$ a criterion for the congruence of two pairs always gives a criterion for orthogonal equivalence. For we may take one matrix in each pair to be the identity matrix.

6. **J-orthogonal equivalence** The set $\mathfrak{M}_n$ of all n-rowed square matrices with elements in a field $\mathfrak{F}$ is said to possess an involution J over $\mathfrak{F}$ if there is a one-to-one correspondence

$$J: \qquad A \longleftrightarrow A^J \qquad (A, A^J \text{ in } \mathfrak{M}_n)$$

such that

$$(78) \qquad (A^J)^J = A, \quad (A+B)^J = A^J + B^J, \quad (AB)^J = B^J A^J, \quad (aI_n)^J = aI_n$$

for every A and B of $\mathfrak{M}_n$ and a of $\mathfrak{F}$. Here I_n is the n-rowed identity matrix. I have proved* that every J is determined† by

$$(79) \qquad A^J = E^{-1}A'E,$$

where E is a non-singular matrix $E = \pm E'$. Call A J-symmetric if $A = A^J$, J-alternate if $A = -A^J$ and $\mathfrak{F}$ does not have characteristic two.

Let S be an automorphism of $\mathfrak{M}_n$ over $\mathfrak{F}$. Then there exists a non-singular matrix P such that

$$(80) \qquad A^S = P^{-1}AP$$

for every A. The involution

$$(81) \qquad A \longleftrightarrow A^{S^{-1}JS}$$

* See the author's paper, *Involutorial simple algebras and real Riemann matrices*, loc. cit.

† Note that conversely J determines E only up to a scalar factor.

has been called an involution *cogredient* with J.* In fact

$$A^{S^{-1}} = PAP^{-1}, \qquad (A^{S^{-1}})' = P'^{-1}A'P', \qquad (A^{S^{-1}})^J = E^{-1}P'^{-1}A'P'E,$$

$$A^{S^{-1}JS} = E_0^{-1}A'E_0, \tag{82}$$

where $E_0 = P'EP$ is congruent to E. The set $\mathfrak{M}_n$ may be thought of as the set of linear transformations of a vector space $\mathfrak{R}$, and the replacement of any basis of $\mathfrak{R}$ by any other replaces the matrices of $\mathfrak{M}_n$ by the similar matrices $P^{-1}AP$. They then replace E by $P'EP$, J by $S^{-1}JS$. Hence cogredient involutions are merely different representations of the same abstract involution.

A matrix D is said to be J-orthogonal if

$$D^JD = I_n. \tag{83}$$

Two matrices A and B are called J-orthogonally equivalent if

$$B = D^JAD \tag{84}$$

for a J-orthogonal matrix D. The case where A^J is the transpose of A, that is $E = I_n$, has already been considered in Theorem 27. But we have the following generalization:

THEOREM 28. *Let $\mathfrak{F}$ be an algebraically closed field of characteristic not two, J be an involution of $\mathfrak{M}_n$ over $\mathfrak{F}$, A and B be matrices with elements in $\mathfrak{F}$ such that $A^J = \epsilon A$, $B^J = \epsilon B$, $\epsilon = \pm 1$. Then A and B are J-orthogonally equivalent if and only if they are similar.*

For if A and B are J-orthogonally equivalent they are clearly similar. Conversely let $PAP^{-1} = B$. Now $E^{-1}B'E = \epsilon B$, $E^{-1}A'E = \epsilon A$ so that

$$B_0 = EB, \qquad A_0 = EA, \qquad B_0' = \delta B_0, \qquad A_0' = \delta A_0,$$

where $\delta = \pm 1$ is the product of ϵ and $E'E^{-1} = \pm 1$. Then the pairs (E, A_0) and (E, B_0) have the property $(EPE^{-1})A_0(P^{-1}) = B_0$, $(EPE^{-1})E(P^{-1}) = E$. By Theorem 26 there exists a non-singular matrix D such that

$$D'ED = E, \qquad D'A_0D = B_0.$$

But then

$$D^JD = (E^{-1}D'E)D = I_n,$$

and D is J-orthogonal. Also $B = E^{-1}D'EAD = D^JAD$ is J-orthogonal to A.

---

* N. Jacobson, *A class of normal simple Lie algebras of characteristic zero*, Annals of Mathematics, vol. 38 (1937), pp. 508–517. Note that conversely the property in the preceding footnote implies that E_0 defines an involution cogredient with that defined by E if and only if E_0 is congruent to a scalar multiple of E.

The proof above indicates that the theory of the congruence of pairs of matrices $(E, A), (C, B)$ over any field $\mathfrak{F}$ is equivalent to the theory of J-orthogonal equivalence. We are of course assuming that E and C are non-singular,

$$E' = \epsilon E, \qquad C' = \epsilon C, \qquad A' = \delta A, \qquad B' = \delta B \qquad (\epsilon = \pm 1, \delta = \pm 1).$$

Necessarily C must be congruent to E, and there is no loss of generality if we replace (C, B) by (E, B_1), where B_1 is clearly congruent to B. Then (E, A) and (E, B_1) are congruent if and only if $E^{-1}A$ and $E^{-1}B_1$ are J-orthogonally equivalent. In fact we define $G^J = E^{-1}G'E$ for any matrix G and have

$$A_0 = E^{-1}A, \qquad B_0 = E^{-1}B_1.$$

We then obtain

$$A_0{}^J = E^{-1}A'E'^{-1}E = \epsilon\delta A_0, \qquad B_0{}^J = \epsilon\delta B_0.$$

The proof of our theorem then shows that the J-orthogonal equivalence of A_0, B_0 and the congruence of (E, A) and (E, B_1) are equivalent concepts. Notice that this is true for arbitrary fields $\mathfrak{F}$, and that we are not assuming the algebraic closure of $\mathfrak{F}$ or even that the characteristic of $\mathfrak{F}$ is not two.

IV. Similarity of square matrices

1. **Reduction to primary components.** If two matrices A and B are J-orthogonally equivalent, they are similar. It follows that the properties of J-orthogonal equivalence depend essentially upon the theory of the similarity of square matrices. The usual modern formulation of this theory is valid for an arbitrary field* but is in a form unsuited for the application we shall wish to make. Thus we shall give a new formulation.

Our first assumptions from the classical theory are the following lemmas:

Lemma 9. *Let B be obtained from A by a row permutation followed by the same column permutation. Then B and A are similar.*

Lemma 10. *Let $f(x) = g(x) \cdot h(x)$ be the characteristic function of an n-rowed square matrix A where $g(x)$ is prime to $h(x)$ and has leading coefficient unity. Then A is similar to*

$$\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}, \tag{85}$$

where G, H have respective characteristic functions $g(x), h(x)$.

We are assuming that G, H, A have elements in a field $\mathfrak{F}$ containing the coefficients of $f(x), g(x), h(x)$, and that our similarity is similarity in $\mathfrak{F}$. The

* For an exposition of the validity of this fact and proofs of the results assumed in this chapter see the author's *Modern Higher Algebra*, chap. 4.

importance of the form (85) is principally due to the fact that for any G and H we have

$$\phi(A) = \begin{pmatrix} \phi(G) & 0 \\ 0 & \phi(H) \end{pmatrix}. \tag{86}$$

This means more precisely that if

$$\phi(x) = x^r + a_1 x^{r-1} + \cdots + a_r \qquad (a_i \text{ in } \mathfrak{F}), \tag{87}$$

then G has m rows, H has $n-m$ rows, and

$$\begin{aligned} &\phi(A) = A^r + a_1 A^{r-1} + \cdots + a_r I_n, \qquad \phi(G) = G^r + \cdots + a_r I_m, \\ &\phi(H) = H^r + \cdots + a_r I_{n-m}. \end{aligned} \tag{88}$$

In particular if $\phi(x)$ is the minimum function of G and $a_r \neq 0$, then

$$G_0 = \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix}, \qquad \phi(G_0) = \begin{pmatrix} 0 & 0 \\ 0 & a_r I_{n-m} \end{pmatrix}. \tag{89}$$

But then $x\phi(x)$ is the minimum function of the matrix G_0 formed by bordering G by $n-m$ rows and columns of zeros.

We shall call two square matrices *relatively prime if their characteristic functions are relatively prime.* We also say that a square matrix is *primary* if its characteristic function is a power of an irreducible polynomial. Then Lemma 10 gives the following lemma:

LEMMA 11. *Every square matrix A is similar in $\mathfrak{F}$ to a direct sum of relatively prime primary components the product of whose characteristic functions is that of A.*

Let $g(x)$ and $h(x)$ be relatively prime. Then $a(x)g(x)+b(x)h(x)=1$ for polynomials a and b. Define $\delta(x)=a(x)g(x)$, $\gamma(x)=1-\delta(x)$. Then $\gamma(G_0)=I_m$, $\delta(G_0)=0$, $\gamma(H_0)=0$, and $\delta(H_0)=I_{n-m}$, for any m-rowed square matrix G_0 such that $g(G_0)=0$, and any $(n-m)$-rowed H_0 such that $h(H_0)=0$. We use this result in the proof of the following lemma:

LEMMA 12. *Let P be non-singular and*

$$A_0 = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1} = \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix}, \tag{90}$$

where the characteristic function $g(x)$ of G and G_0 is prime to the characteristic function $h(x)$ of H and H_0. Then

$$P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}, \qquad P_1 G P_1^{-1} = G_0, \qquad P_2 H P_2^{-1} = H_0. \tag{91}$$

Write

$$P = \begin{pmatrix} P_1 & P_3 \\ P_4 & P_2 \end{pmatrix}.$$

It is clearly sufficient to prove P_3 and P_4 zero. Form $\gamma(A_0)$ and obtain

$$P\begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix} P^{-1} = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$$

so that

$$\begin{pmatrix} P_1 & 0 \\ P_4 & 0 \end{pmatrix} = \begin{pmatrix} P_1 & P_3 \\ 0 & 0 \end{pmatrix},$$

and $P_4=0$, $P_3=0$ as desired.

An evident induction now gives the following lemma:

LEMMA 13. *Let*

$$A = \text{diag}\,[G_1, \cdots, G_t], \quad A_0 = \text{diag}\,[G_{01}, \cdots, G_{0t}] = PAP^{-1}, \tag{92}$$

with relatively prime primary components G_i having the same characteristic functions as G_{0i}. Then P is the direct sum of matrices P_i such that $G_{0i}=P_iG_iP_i^{-1}$.

2. **Indecomposable matrices.** A matrix A is called *indecomposable* in $\mathfrak{F}$ if A is not similar in $\mathfrak{F}$ to a direct sum of two matrices. We shall assume the known lemma:*

LEMMA 14. *A matrix A is indecomposable in $\mathfrak{F}$ if and only if its characteristic function is equal to the power*

$$f(x) = x^n - (a_1x^{n-1} + \cdots + a_n) = [d(x)]^e \qquad (a_i \text{ in } \mathfrak{F}) \tag{93}$$

of an irreducible polynomial $d(x)$ and coincides with its minimum function. Every indecomposable matrix is similar in $\mathfrak{F}$ to

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_1 \end{pmatrix}. \tag{94}$$

Conversely the minimum function of (94) *is its characteristic function $f(x)$ and* (94) *is indecomposable if and only if $f(x)=d(x)^e$ for an irreducible $d(x)$.*

Lemmas 11, 13 reduce the problem of the reduction of a matrix to a direct

* See the author's *Modern Higher Algebra*, chap. 4.

sum of indecomposables under similarity transformations to the case where A is primary. When A is primary and has the characteristic function $d(x)^f$, $d(x)$ irreducible, the invariant factors of $xI-A$ are the characteristic functions of its indecomposable components. Then A is similar in $\mathfrak{F}$ to

$$\text{diag } [B_1, \cdots, B_t], \tag{95}$$

where B_i has $d(x)^{e_i}$ as characteristic function and is indecomposable,

$$e_1 + \cdots + e_t = f, \qquad e_1 \geqq e_2 \geqq \cdots \geqq e_t \geqq 1. \tag{96}$$

We shall call the e_i the *indices* of (95). In particular

$$e_1 = e, \tag{97}$$

where $d(x)^e$ is the minimum function of both (95) and B_1.

We now prove the following lemma:

LEMMA 15. *Let $d(x)=c[h(x)]$ be irreducible in $\mathfrak{F}$, A be an n-rowed indecomposable matrix with $c(x)^e$ as minimum function,*

$$h(x) = x^m + b_1x^{m-1} + \cdots + b_m. \tag{98}$$

Then the matrix

$$B = \begin{pmatrix} 0 & I_n & 0 & \cdots 0 \\ 0 & 0 & I_n & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots \cdot \\ 0 & 0 & 0 & \cdots I_n \\ \Gamma_m & \Gamma_{m-1} & \Gamma_{m-2} & \cdots \Gamma_1 \end{pmatrix}, \qquad \Gamma_m = -b_mI_n + A, \quad \Gamma_i = -b_iI_n \quad (i = 1, \cdots, m-1), \tag{99}$$

is indecomposable and has $d(x)^e$ as characteristic function.

For it is clear, from the fact that all elements of B are polynomials in A, that

$$h(B) = B^m + b_1B^{m-1} + \cdots + b_mI = \text{diag } [A, \cdots, A],$$

I the nm-rowed identity matrix. Then $[c(A)]^e=0$ so that $[d(B)]^e=c[h(B)]^e=0$. It follows that the minimum function of B divides $[d(x)]^e$. But $d(x)$ is irreducible; thus it has the form $d(x)^g$, $g \leqq e$. However $c[h(B)]^g=0$ implies that $[c(A)]^g=0$ whence $g \geqq e$. Then $g=e$. The degree of the minimum function $d(x)^e$ of B is the order nm of B, and B is indecomposable.

We next let B be a decomposable primary matrix so that we may assume that B is a direct sum of matrices B_i which are indecomposable. If $d(x)^f$ is the characteristic function of B, and the irreducible polynomial $d(x)=c[h(x)]$ as in Lemma 15, we may assume that each B_i has the form (99). But the b_j

are the same in each B_i, and an evident permutation of the rows and corresponding columns of B carries B into a similar matrix of the form (99), where

$$A = \text{diag } [A_1, \cdots, A_t]$$

has the same indices as B. We state this result as in the following lemma:

LEMMA 16. *Let $d(x) = c[h(x)]$ be irreducible, $h(x)$ be given by* (98), *and B have characteristic function $[d(x)]^f$. Then B is similar in $\mathfrak{F}$ to a matrix* (99), *where A has the characteristic function $c(x)^f$ and the same indices as B.*

3. **Two canonical forms.** A matrix N is called *nilpotent* of index e if $N^e = 0$, $N^{e-1} \neq 0$. The minimum function of N is clearly x^e. Then N is similar in $\mathfrak{F}$ to

$$\text{diag } [N_1, \cdots, N_t],$$

where N_i is nilpotent of index e_i and may be taken to be the e_i-rowed square matrix

$$\begin{pmatrix} 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdot & \cdot & \cdots\cdots \\ 0 & 0 & 0 \cdots 1 \\ 0 & 0 & 0 \cdots 0 \end{pmatrix}. \tag{100}$$

Notice that $N_i = 0$ if $e_i = 1$. Also $e = e_1$, N has order $f = e_1 + \cdots + e_t$. The matrices N_i are indecomposable, and thus a nilpotent matrix is indecomposable if and only if its index is its order. But the indices of N are clearly the respective indices of its nilpotent indecomposable components N_i in the sense in which we defined index above.

We have seen that the terminology of indices which we defined for arbitrary matrices has precise connotations for nilpotent matrices. These connotations are made precise also generally by the following theorem:

THEOREM 29. *Let the characteristic function of B be $d(x)^f$, where $d(x) = x^m + a_1 x^{m-1} + \cdots + a_m$, ($a_i$ in $\mathfrak{F}$), is irreducible. Then B is similar in $\mathfrak{F}$ to*

$$\begin{pmatrix} 0 & I_f & 0 & \cdots 0 \\ 0 & 0 & I_f & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots\cdots \\ 0 & 0 & 0 & \cdots I_f \\ \Delta_m & \Delta_{m-1} & \Delta_{m-2} & \cdots \Delta_1 \end{pmatrix}, \quad \Delta_m = N - a_m I_f, \quad \Delta_i = -a_i I_f, \quad (i = 1, \cdots, m-1), \tag{101}$$

where N is a nilpotent matrix whose indices are the same as those of B.

Our theorem is an immediate application of Lemma 16 to the case where $h(x)$ is irreducible, $c(x)=x$, $d(x)=c[h(x)]=h(x)$. Then the matrix A of Lemma 16 has x^f as characteristic function and is our nilpotent matrix N.

The matrix (101) is a canonical form of a primary matrix. For some purposes other forms may be preferable. One such form is given in the following:

THEOREM 30. *Let $d(x)=c(x^{p^k})$ be irreducible,*

$$(102)\qquad B_j=\begin{pmatrix} 0 & I_\nu & 0\cdots 0 \\ 0 & 0 & I_\nu\cdots 0 \\ \cdot & \cdot & \cdot\cdots\cdot \\ 0 & 0 & 0\cdots I_\nu \\ B_{j-1} & 0 & 0\cdots 0 \end{pmatrix}, \qquad \nu=p^{j-1}m, \qquad (j=1,\cdots,k),$$

where $B_0=A$ is an m-rowed square matrix with $[c(x)]^f$ as characteristic function. Then $B=B_k$ has $[d(x)]^f$ as characteristic function and the same indices as A.

For proof we take $h(x)=x^p$ in Lemma 16 and use an induction on k. This result is of particular use in case $d(x)$ is an inseparable irreducible polynomial over $\mathfrak{F}$ of characteristic p and is equal to $c(x^{p^k})$, $c(x)$ separable.

4. **The algebra $\mathfrak{F}[B]$.** The algebra $\mathfrak{F}[B]$ of all polynomials in a matrix B with characteristic function $d(x)^f$, $d(x)$ irreducible, contains certain sub-fields and certain nilpotent matrices. We first prove the following theorem:

THEOREM 31. *Let the minimum function of B have the form $d(x)^e$ with $d(x)$ irreducible. Write*

$$(103)\qquad d(x)=c(x^\pi),$$

where $c(x)$ is separable and

$$\pi=p^k, \quad \text{or} \quad \pi=1$$

according as $d(x)$ is inseparable over $\mathfrak{F}$ of characteristic p or is separable. Then there exists a nilpotent matrix N of index e such that

$$(104)\qquad B^\pi=N+A, \qquad c(A)=0.$$

Thus $\mathfrak{K}=\mathfrak{F}[A]$ is a separable field and the polynomial $x^\pi-A$ is irreducible in $\mathfrak{F}$.

Our proof depends upon the known lemma:*

LEMMA 17. *Let $c(x)$ be irreducible and separable. Then for every e there exists a polynomial $g_e(x)$ such that $c[g_e(x)]$ is divisible by $c(x)^e$.*

The matrix $T=B^\pi$ has $c(x)^e$ as minimum function and Lemma 17 implies

* Loc. cit. chap. 10. The result trivially follows from our Lemma 8.

the existence of a polynomial $g(T)=A_0$ such that $c(A_0)=0$. Since $c(x)$ is separable and irreducible the algebra $\mathfrak{K}=\mathfrak{F}[A_0]$ is a field. We now prove the following lemma:

LEMMA 18. *The matrix $c(T)=N_0$ is nilpotent of index e and the algebra $\mathfrak{K}[N_0]=\mathfrak{F}[T]$.*

For the minimum function of N_0 is clearly x^e. It remains to prove that $\mathfrak{K}[N_0]$ which is contained in $\mathfrak{F}[T]$ has the same order over $\mathfrak{F}$ and equals $\mathfrak{F}[T]$. It is sufficient to show that $a_0+a_1N_0+\cdots+a_{e-1}N_0^{e-1}=0$ for a_i in $\mathfrak{K}$ if and only if the $a_i=0$. But $a_0N_0^{e-1}=0$, $N_0^{e-1}\neq 0$, $a_0=0$ in $\mathfrak{K}$. Similarly all the other $a_i=0$.

The quantity T now has the form $T=A+N$, where A is in $\mathfrak{K}$, $N=(a_1+a_2N_0+\cdots+a_{e-1}N_0^{e-2})N_0$ is nilpotent. But $c(T)=c(A)+Nc_1(N)=N_0$, $c(A)$ is nilpotent and in $\mathfrak{K}$. It follows that $c(A)=0$, $N_0=Nc_1(N)$, $N_0^g=0$ where g is the index of N. Thus $g\geqq e$. Similarly $g\leqq e$ and N has index e. We next prove the following lemma:

LEMMA 19. *Let $x^{p^k}-a$ be reducible in a field $\mathfrak{K}$ containing a. Then $a=b^p$, b in $\mathfrak{K}$.*

For $x^{p^k}-a=g(x)h(x)$, where the constant term of $g(x)$ is the tth power of a root ξ of $x^{p^k}-a=0$, $g(x)$ has degree t. Hence $\xi^{rs}=b_0$ in $\mathfrak{F}$, where s is prime to p, r is a power of p. But then $ss_1\equiv 1\ (p^k)$ and $\xi^r=b_1$ in $\mathfrak{K}$. Since $r<p^k$ we have $p^k=rp^\nu$, $b_1{}^{p^\nu}=a$, $a=b^p$, b in $\mathfrak{K}$.

We may finally show that x^r-A is irreducible. For otherwise $A=D^p$, D in $\mathfrak{K}$, $c(A)=c(D^p)=0$, $c(x^p)$ is inseparable and has a root D in the separable field $\mathfrak{K}$. This is impossible, and our proof of Theorem 31 is complete.

The matrix A has $c(x)$ as minimum function and is similar to

$$\text{diag}\,[G,\cdots,G], \tag{105}$$

where G has $c(x)$ as both minimum function and characteristic function. It is well known that the only matrices commutative with G are polynomials in G with coefficients in $\mathfrak{F}$. If $AB=BA$ and we write $B=(B_{ij})$, we obtain $B_{ij}G=GB_{ij}$, the B_{ij} are in $\mathfrak{F}[G]$. Thus B may be regarded as a matrix with elements in the field $\mathfrak{K}=\mathfrak{F}[G]$.

The matrix N of Theorem 31 is commutative with A and hence may be regarded as a nilpotent matrix of index e and elements in $\mathfrak{K}$. The matrix B is also a matrix over $\mathfrak{K}$ but now has minimum function

$$(x^{p^k}-G)^e. \tag{106}$$

Then the construction of matrices with minimum function $[c(x^{p^k})]^e$ is completed by Theorems 30, 31, and the above.

5. **Fields of matrices.** Let $c(x)$ define a separable field $\mathfrak{K}=\mathfrak{F}[G]$ where G is a matrix in our canonical form given by (94) for the separable irreducible polynomial $c(x)$ used as the $f(x)$ of (93). Write

$$c(x) = (x - \alpha_1) \cdots (x - \alpha_n)$$

with distinct α_i in an algebraic extension of $\mathfrak{F}$. Then the Vandermonde matrix*

$$V = (\alpha_{ij}), \qquad \alpha_{ij} = \alpha_j^{i-1} \qquad (i, j = 1, \cdots, n) \tag{107}$$

is non-singular,

$$VV' = (s_{i+k-2}) \tag{108}$$

has elements $s_j=\sum_{i=1}^{n}\alpha_i^j$ in $\mathfrak{F}$ and determinant the discriminant of $f(x)$. A simple computation gives

$$V^{-1}GV = \alpha = \text{diag}\,[\alpha_1, \cdots, \alpha_n]. \tag{109}$$

Let H be any n-rowed square matrix with elements in $\mathfrak{F}$, and define

$$H_0 = (VV')^{-1}H = (h_{ij}) \qquad (h_{ij} \text{ in } \mathfrak{F}).$$

Then

$$V^{-1}HV = V'H_0V = (\Lambda(\alpha_i, \alpha_j)) \qquad (i, j = 1, \cdots, n), \tag{110}$$

where

$$\Lambda(x, y) = \sum_{i,j}^{1,\cdots,n} x^{i-1}h_{ij}y^{j-1}.$$

Clearly $HG=GH$ if and only if $\Lambda(\alpha_i, \alpha_j)=0$ for $i\neq j$, and $H=h(G)$, where $h(x)=\Lambda(x, x)$.

The result that $HG=GH$ if and only if H is a polynomial in G is true for other types of matrices G as well as those above. In particular it is true for indecomposable nilpotent matrices. Let now A be an n-rowed square matrix with irreducible minimum function $d(x)$ of degree m. This is the case $e=1$ of our theory, and we do not assume that $d(x)$ is separable. If $d(x)=x$, then $A=0$, and we discard this case. Otherwise A is similar in $\mathfrak{F}$ to the matrix

$$\text{diag}\,[G, \cdots, G],$$

where G is given by (94) for $d(x)$, A is now a q-rowed matrix with elements in a field $\mathfrak{F}[G]$. The only matrices commutative with A are q-rowed matrices

* The method of proof of this section has been used many times in the theories of Riemann matrices and of linear transformations. For a partial list of references see the bibliography in the paper referred to in the first footnote on page 387.

with elements in $\mathfrak{F}[G]$, and the minimum function of A over $\mathfrak{F}[G]$ is $x-G$. However the analogous result cannot be obtained when $e>1$ since we cannot in general prove the existence of an inseparable sub-field $\mathfrak{R}$ of our algebra of polynomials in a matrix.

6. **Consequences of the canonical forms.** The transpose A' of any matrix A is similar to A. For certain simple matrices the transformation carrying A into A' assumes an interesting and simple form. We let A have the canonical form

$$A = \begin{pmatrix} 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdot & \cdot & \cdot \cdot \cdot \cdot \cdot \\ 0 & 0 & 0 \cdots 1 \\ a & 0 & 0 \cdots 0 \end{pmatrix} \tag{111}$$

so that the characteristic and minimum functions of A are $f(x)=x^n-a$. Write

$$U = \begin{pmatrix} 0 & 0 \cdots 0 & 1 \\ 0 & 0 \cdots 1 & 0 \\ \cdot & \cdot \cdot \cdot \cdot \cdot & \cdot \\ 0 & 1 \cdots 0 & 0 \\ 1 & 0 \cdots 0 & 0 \end{pmatrix}. \tag{112}$$

Then

$$UA = \begin{pmatrix} a & 0 \cdots 0 & 0 \\ 0 & 0 \cdots 0 & 1 \\ 0 & 0 \cdots 1 & 0 \\ \cdot & \cdot \cdot \cdot \cdot \cdot & \cdot \\ 0 & 1 \cdots 0 & 0 \end{pmatrix}, \qquad UA^2 = \begin{pmatrix} 0 & a & 0 \cdots 0 & 0 \\ a & 0 & 0 \cdots 0 & 0 \\ 0 & 0 & 0 \cdots 0 & 1 \\ 0 & 0 & 0 \cdots 1 & 0 \\ \cdot & \cdot & \cdot \cdot \cdot \cdot \cdot & \cdot \\ 0 & 0 & 1 \cdots 0 & 0 \end{pmatrix}. \tag{113}$$

By an evident computation if

$$f(A) = a_0 + a_1 A + \cdots + a_{n-1} A^{n-1}, \tag{114}$$

then

$$Uf(A) = \begin{pmatrix} a_1 a & a_2 a \cdots a_{n-1} a & a_0 \\ a_2 a & a_3 a \cdots a_0 & a_1 \\ \cdot & \cdot \cdot \cdot \cdot & \cdot \\ a_{n-1} a & a_0 \cdots a_{n-3} & a_{n-2} \\ a_0 & a_1 \cdots a_{n-2} & a_{n-1} \end{pmatrix}. \tag{115}$$

It is clear that UA is symmetric, and since U is symmetric and non-singular,

$$UAU^{-1} = A'. \tag{116}$$

We shall require (115) later. We shall also use the following existence theorem which is a consequence of Theorem 30 for the case $p=2$:

THEOREM 32. *Let $n=2^k$, $f(x)=x^n-a$ be irreducible in $\mathfrak{F}$. Then there exists an n-rowed square matrix A with $f(x)$ as minimum function and such that*

$$E^{-1}A'E = -A, \qquad S^{-1}A'S = A \tag{117}$$

for an alternate matrix E and a symmetric matrix S.

The theorem is true for $n=2$, $k=1$ since

$$E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 1 \\ a & 0 \end{pmatrix}$$

satisfy (117) by direct computation. Assume the theorem true for $k-1$, write $m=2^{k-1}$, and have $S_m^{-1}A_m'S_m=A_m$. By Theorem 30 the matrix

$$A = \begin{pmatrix} 0 & I_m \\ A_m & 0 \end{pmatrix} \tag{118}$$

has $f(x)$ as minimum (and characteristic) function. Then the matrices

$$E = \begin{pmatrix} 0 & S_m \\ -S_m & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & S_m \\ S_m & 0 \end{pmatrix} \tag{119}$$

are alternate and symmetric respectively, and if $\epsilon = \pm 1$,

$$\begin{pmatrix} 0 & S_m \\ \epsilon S_m & 0 \end{pmatrix}\begin{pmatrix} 0 & I_m \\ A_m & 0 \end{pmatrix} = \begin{pmatrix} S_mA_m & 0 \\ 0 & \epsilon S_m \end{pmatrix},$$
$$\begin{pmatrix} 0 & A_m' \\ I_m & 0 \end{pmatrix}\begin{pmatrix} 0 & S_m \\ \epsilon S_m & 0 \end{pmatrix} = \begin{pmatrix} \epsilon S_mA_m & 0 \\ 0 & S_m \end{pmatrix}, \tag{120}$$

since $A_m'S_m=S_mA_m$. Put $\epsilon=-1$ and obtain $EA=-A'E$. The value $\epsilon=1$ gives $SA=A'S$ as desired.

V. ORTHOGONAL EQUIVALENCE IN $\mathfrak{F}$ OF CHARACTERISTIC TWO

1. **The problem.** Our principal interest will be in obtaining a complete determination of the invariant factors of any symmetric matrix whose elements are in a field of characteristic two.* Part of this theory will be concerned with the orthogonal equivalence of symmetric matrices, and we shall

* The field of reference throughout this chapter will be any field of characteristic two and we shall drop the notation $\mathfrak{F}^{(2)}$ and simply use $\mathfrak{F}$.

show why it can happen that two symmetric matrices may be similar but not orthogonally equivalent.

2. **J-orthogonal equivalence.*** Let J be an involution over $\mathfrak{F}$ of the algebra $\mathfrak{M}_n$ of all n-rowed square matrices with elements in a field $\mathfrak{F}$ (with any characteristic). Suppose that

$$A^J = \epsilon A, \qquad B = PAP^{-1} \qquad (\epsilon = \pm 1) \tag{121}$$

for a non-singular matrix P. We may then prove the lemma:

LEMMA 20. *The matrix B of* (121) *has the property*

$$B^J = \epsilon B \tag{122}$$

if and only if P^JP is commutative with A.

For $B^J = (P^J)^{-1}\epsilon A P^J = \epsilon B = \epsilon PAP^{-1}$, $P^JPA = AP^JP$. The converse is trivial.

If B is any matrix similar to A so that $B = PAP^{-1}$, then $B = DAD^{-1}$ if and only if $D = PC$ where C is commutative with A. Then

$$D^JD = C^J(P^JP)C. \tag{123}$$

But D is J-orthogonal if and only if D^JD is the n-rowed identity matrix.

LEMMA 21. *The matrix PAP^{-1} is J-orthogonally equivalent to A if and only if P^JP is congruent to the identity matrix under a transformation* (123) *with C commutative with A.*

Let us study the implications of Lemma 21 in a special case. Assume that A is such that the only matrices commutative with A are polynomials in A with coefficients in $\mathfrak{F}$. Then $P^JP = \phi(A)$ is such a polynomial and $C = C(A)$. If $A^J = A$, then $C^J = C$, and $C^J\phi C = I$ if and only if

$$\phi = C^{-2} \tag{124}$$

is the square of a polynomial in A. Conversely if $\phi = Q^JQ = Q^2$, we have $P^JP = Q^JQ$, $D = PQ^{-1}$ is J-orthogonal, and

$$B = PAP^{-1} = DAD^{-1} \tag{125}$$

is J-orthogonally equivalent to A.

We shall see in an example later that it is possible for a given $\phi(A)$ to have the form $P'P$ but not the form $[Q(A)]^2$. Thus the restriction above is not a redundant consequence of the equation $\phi(A) = P'P$.

* For recent results implying theorems on the J-orthogonal equivalence of matrices over an arbitrary field of characteristic not two, see the papers of John Williamson in the American Journal of Mathematics, vol. 57 (1935), pp. 475–490; vol. 58 (1936), pp. 141–163; and vol. 59 (1937), pp. 399–413.

In the remainder of the chapter we shall assume, unless it is otherwise stated, that $\mathfrak{F}$ has characteristic two. We shall also restrict our attention to the case of ordinary orthogonal equivalence.

3. **Separable fields.** Every separable field $\mathfrak{K}$ over $\mathfrak{F}$ is a simple extension

$$\mathfrak{K} = \mathfrak{F}[A] \tag{126}$$

of all polynomials in A with coefficients in $\mathfrak{F}$, where A is a root of an irreducible separable equation

$$f(x) = x^m - (a_1 x^{m-1} + \cdots + a_m) = 0. \tag{127}$$

We may easily prove the following theorem:

THEOREM 33. *Let $f(x)$ of* (127) *be separable and irreducible in $\mathfrak{F}$ of characteristic two. Then there exists an m-rowed symmetric matrix G with elements in $\mathfrak{F}$ and $f(x)$ as characteristic function.*

For let

$$G_0 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ a_m & a_{m-1} & a_{m-2} & \cdots & a_1 \end{pmatrix}. \tag{128}$$

Then every m-rowed square matrix with $f(x)$ as characteristic function is indecomposable, has $f(x)$ as minimum function, and is similar to G_0. In (109) we saw that $V^{-1}G_0V$ is a diagonal matrix and is thus symmetric. Thus $V'G_0'V'^{-1} = V^{-1}G_0V$,

$$VV'(G_0') = G_0(VV'). \tag{129}$$

The diagonal elements of VV' are

$$s_{i+i-2} = \sum_{j=1}^{m} \alpha_j^{2i-2} = \left(\sum_{j=1}^{m} \alpha_j^{i-1} \right)^2 = (s_{i-1})^2. \tag{130}$$

Hence VV' is a definite matrix. It is this remarkable property which gives us our result, Theorem 33.

We may now write $VV' = RR'$, where R is a non-singular matrix with elements in $\mathfrak{F}$. Then $RR'G_0' = G_0RR'$, $R^{-1}G_0R = R'G_0'R'^{-1}$. Define

$$G = R^{-1}G_0R, \tag{131}$$

and obtain $G' = R'G_0'R'^{-1} = G$. The matrix G is symmetric and is similar to G_0. It is the desired matrix. Notice that

$$W = V^{-1}R, \qquad W'W = I, \qquad WGW^{-1} = \alpha,$$

so that G is orthogonally equivalent in an algebraic extension of $\mathfrak{F}$ to the diagonal matrix α. But then we may prove the lemma:

LEMMA 22. *Let $\mathfrak{F}_0 = \mathfrak{F}[G]$ be the field of symmetric matrices consisting of all polynomials with coefficients in $\mathfrak{F}$ of the matrix G of* (131). *Then the only alternate matrix of $\mathfrak{F}_0$ is the zero matrix. Moreover a matrix of $\mathfrak{F}_0$ is definite if and only if it is the square of a non-zero quantity of $\mathfrak{F}_0$.*

For if $\psi(G)$ is in $\mathfrak{F}_0$, we have

$$W\psi(G)W' = \psi(\alpha).$$

If $\psi(G)$ is alternate so is the diagonal matrix $\psi(\alpha)$. But then $\psi(\alpha) = 0$, $\psi(G) = 0$. Let $\psi(G)$ be definite so that we may write $\psi(G) = H'H$. Then $W = W'^{-1} = (R'V'^{-1})^{-1} = V'R'^{-1}$,

$$\psi(\alpha) = W\psi(G)W' = (TV)'(TV),$$

where $T = HR^{-1}$ has elements t_{ij} in $\mathfrak{F}$. The elements in the first column of TV are

$$t_i(\alpha_1) = \sum_{j=1}^{m} t_{ij}\alpha_1^{j-1} \qquad (i = 1, \cdots, m).$$

Thus the element in the first row and column of $\psi(\alpha)$ is

$$\psi(\alpha_1) = \sum_{i=1}^{m} [t_i(\alpha_1)]^2 = \left[\sum_{i=1}^{m} t_i(\alpha_1)\right]^2 = [\tau(\alpha_1)]^2,$$

where $\tau(\alpha_1)$ is in $\mathfrak{F}[\alpha_1]$. Then $\psi(G) = [\tau(G)]^2$, $\tau(G) \neq 0$ in $\mathfrak{F}_0$.

The only matrices commutative with G are quantities of $\mathfrak{F}_0 = \mathfrak{F}[G]$, a field of symmetric matrices. We let

$$A = \text{diag } [G, \cdots, G] \tag{132}$$

have $n = qm$ rows, so that A is a q-rowed matrix with elements in $\mathfrak{F}_0$. Let

$$\mathfrak{M}_q \quad \text{over} \quad \mathfrak{F}_0 \tag{133}$$

be the set of all q-rowed matrices with elements in $\mathfrak{F}_0$. Then $\mathfrak{M}_q$ is the algebra of all n-rowed matrices commutative with A. Every such matrix has the form

$$B = (B_{ij}) \qquad (B_{ij} \text{ in } \mathfrak{F}_0,\ i, j = 1, \cdots, q)$$

and is an n-rowed matrix with elements in $\mathfrak{F}$. We define

$$B^T = (B_{0ij}), \qquad B_{0ij} = B_{ji} \qquad (i, j = 1, \cdots, q)$$

so that B^T is the transpose of B considered as a q-rowed matrix of $\mathfrak{M}_q$ over $\mathfrak{F}_0$.

But B is an n-rowed square matrix with elements in $\mathfrak{F}$, and

$$B' = (B_{1ij}), \qquad B_{1ij} = B'_{ji} \qquad (i, j = 1, \cdots, q).$$

However every B_{ji} is symmetric, and thus we have proved that $B^T = B'$ for every B of $\mathfrak{M}_q$ over $\mathfrak{F}_0$. We now have the following result:

THEOREM 34. *Let S be an n-rowed symmetric matrix commutative with A. Then S is in $\mathfrak{M}_q$ over $\mathfrak{F}_0$ and is symmetric if and only if it is symmetric in $\mathfrak{M}_q$, is alternate if and only if it is alternate in $\mathfrak{M}_q$, and is definite if and only if it is definite in $\mathfrak{M}_q$.*

For $S^T = S'$ and $S = S'$ if and only if $S = S^T$. If $S = (S_{ij})$ and the S_{ij} are in $\mathfrak{F}_0$, then $S = S'$ and $S_{ii} = 0$ implies that S is alternate. Conversely if S is alternate, we have $S = S'$, and the S_{ii} are alternate. By the lemma above the $S_{ii} = 0$; S is alternate in $\mathfrak{M}_q$. If S is definite in $\mathfrak{M}_q$, the S_{ii} are the squares of quantities of $\mathfrak{F}_0$ and are thus zero or definite; S is definite. Conversely let S be a definite n-rowed matrix. Then the principal sub-matrices S_{ii} are either zero or definite and at least one $S_{ii} \neq 0$. By our lemma each $S_{ii} = (\tau_{ii})^2$, τ_{ii} in $\mathfrak{F}_0$ and not all zero, S is definite in $\mathfrak{M}_q$.

Consider a symmetric field of matrices $\mathfrak{K} = \mathfrak{F}[B]$, B a root of $f(x) = 0$ of (127). Then B is similar to A and

$$B = PAP^{-1}, \qquad A = \text{diag}\,[G, \cdots, G], \tag{134}$$

as in (132). Since A and B are symmetric we may apply Lemma 20 to see that $P'P$ is in $\mathfrak{M}_q$ over $\mathfrak{F}_0$. By Theorem 34 the matrix $P'P$ is definite in $\mathfrak{M}_q$ and $P'P = Q'Q$ where Q is in $\mathfrak{M}_q$, $QA = AQ$. Lemma 21 then states that B is orthogonally equivalent to A, and we have the following theorem:

THEOREM 35. *Any two n-rowed symmetric matrices with the same separable irreducible minimum function over $\mathfrak{F}$ of characteristic two are orthogonally equivalent in $\mathfrak{F}$.*

Theorem 34 may be applied to prove a generalization of Theorem 32.

THEOREM 36. *Let $f(x) = c(x^{2^k})$ be irreducible in $\mathfrak{F}$ of characteristic two, $c(x)$ be separable of degree m, and $q = 2^k > 1$. Then there exists an alternate n-rowed square matrix E and an n-rowed matrix B such that*

$$f(B) = 0, \qquad E^{-1}B'E = B, \tag{135}$$

for every n divisible by $2^k m$.

It is clearly sufficient to give the proof for $n = 2^k m$. We construct an m-rowed symmetric matrix G which is a root of $c(x) = 0$. Use Theorem 32 for the field $\mathfrak{F}_0$, and $a = G$ and obtain an n-rowed square matrix B such that B

considered as a matrix of 2^k rows over $\mathfrak{F}_0$ has $x^{2^k}-G$ as minimum function. Then $B^{2^k}=A$, A the matrix of (132), $f(B)=0$ as desired. Also there exists a matrix E such that $E^{-1}B^JE=B$, $E^J=E$, where E is alternate, J is the involution of the algebra of all 2^k-rowed matrices over $\mathfrak{F}_0$. By Theorem 34 we have E alternate, $E^{-1}B'E=B$.

4. **Application to primary matrices.** Let S be a symmetric primary matrix. We apply Theorem 31 to prove the existence of a polynomial $A=A(S)$ in S such that

$$S^{2^k} = A + N, \tag{136}$$

where N is nilpotent, $c(A)=0$. Here the minimum function of S is $[c(x^{2^k})]^e=0$; N has index e. By an orthogonal transformation we may transform A into the form (132). By Theorem 34 both N and S are symmetric over $\mathfrak{F}_0$, N is nilpotent. Then $S^{2^k}=GI_q+N$, and the minimum and characteristic functions of S over $\mathfrak{F}_0$ are

$$(x^{2^k}-G)^e, \qquad (x^{2^k}-G)^f, \tag{137}$$

where $x^{2^k}-G$ is irreducible in the field $\mathfrak{F}_0$ and S is now a matrix of $q=2^kf$ rows.

Let T be symmetric and similar to S. There is no loss of generality if we take $A(T)$, which is similar to A, equal to A. For by Theorem 35 the matrix $A(T)$ is orthogonally equivalent to A. But then S and T are matrices with elements in $\mathfrak{F}_0$. We now prove the statement:

THEOREM 37. *The matrices S and T are orthogonal in $\mathfrak{F}$ if and only if they are orthogonal considered as matrices over $\mathfrak{F}_0$.*

For $SA=AS$, $T=PSP^{-1}$, so that $P'P$ is commutative with S and hence with A. Then $P'P$ is definite, and we have already seen that $P'P$ is a definite matrix over $\mathfrak{F}_0$. Now S and T are orthogonally equivalent if and only if $C'(P'P)C=I$ for $CS=SC$. Thus $CA=AC$, C is a matrix with elements in $\mathfrak{F}_0$ and is commutative with S; S and T are orthogonally equivalent when considered as matrices over $\mathfrak{F}_0$.

We now apply Theorem 30 to see that S is similar in $\mathfrak{F}_0$ to a matrix S_k, where

$$S_j = \begin{pmatrix} 0 & I_{\nu_j} \\ S_{j-1} & 0 \end{pmatrix}, \qquad \nu_j = 2^{j-1}m \qquad (j = 1, \cdots, k), \tag{138}$$

and

$$S_1 = \begin{pmatrix} 0 & I_{\nu_0} \\ G + N_1 & 0 \end{pmatrix}. \tag{139}$$

By Theorem 30 the matrix N_1 is nilpotent and has the same indices as S. Finally Theorem 37 implies that T is orthogonally equivalent to S over $\mathfrak{F}$ if and only if T is orthogonally equivalent to S over $\mathfrak{F}_0$. We have thus reduced our considerations for primary matrices to the case of matrices with minimum function $(x^{2^k}-a)^e$, a in our field $\mathfrak{F}$. We shall prove that $e>1$; that is, the following theorem:

THEOREM 38. *There exist no inseparable fields of symmetric matrices over $\mathfrak{F}$ of characteristic two.*

For let $\mathfrak{K}$ over $\mathfrak{F}$ be inseparable. Then there exists a symmetric matrix S in $\mathfrak{K}$ such that $S^2=A$, A as in (132). Then

$$S^2 = \text{diag } [G, \cdots, G],$$

where x^2-G is irreducible in $\mathfrak{F}$. But S^2 is definite over $\mathfrak{F}$ since $S^2=S'S$, S is non-singular. By the proof of Lemma 22 the matrix S^2 is definite over $\mathfrak{F}_0$, $G=[\phi(G)]^2$, x^2-G is reducible in $\mathfrak{F}$, and we have a contradiction.

5. **Nilpotent matrices.** We shall construct symmetric nilpotent matrices with arbitary indices. Let U be defined as in (112), N_0 be the matrix A of (111) with $a=0$. Then N_0 is nilpotent of index and order n and is indecomposable.

The matrix

$$Uf(N_0) = \begin{pmatrix} 0 & 0 \cdots 0 & 0 & a_0 \\ 0 & 0 \cdots 0 & a_0 & a_1 \\ 0 & 0 \cdots a_0 & a_1 & a_2 \\ \cdot & \cdot \cdots \cdot & \cdot & \cdot \\ 0 & a_0 \cdots a_{n-4} & a_{n-3} & a_{n-2} \\ a_0 & a_1 \cdots a_{n-3} & a_{n-2} & a_{n-1} \end{pmatrix}, \tag{140}$$

for any polynomial $f(N)$ with coefficients a_i in $\mathfrak{F}$. We may write

$$f(x) \equiv f_1(x^2) + xf_2(x^2).$$

Then $f_i(x^2)$ is the square of a polynomial in x with coefficients in $\mathfrak{F}$ if and only if its coefficients are squares in $\mathfrak{F}$. We now have the theorem:

THEOREM 39. *Let $h(x)=f_1(x^2)$ or $f_2(x^2)$ according as n is odd or even. Then $Uf(N_0)$ is alternate if and only if $h(x)\equiv 0$ and is semi-definite if and only if $h(x)=[g(x)]^2\not\equiv 0$.*

For we observe that the only elements on the main diagonal of $Uf(N_0)$ are zeros and the complete set of coefficients of $h(x)$. Thus $Uf(N_0)$ alternate implies that $h(x)\equiv 0$; $Uf(N_0)$ semi-definite implies that $h(x)$ is a square.

The matrices $f(N_0)$, $Uf(N_0)$ are non-singular if and only if $a_0 \neq 0$. In particular $U(I+N_0)$ is non-singular. By Theorem 39 it is definite. We now write

$$U(I + N_0) = P'P, \qquad N = PN_0P^{-1}.$$

Since $UN_0 = N_0'U$ we have $P'PN_0 = N_0'P'P$, $PN_0P^{-1} = N = P'^{-1}N_0'P' = N'$, and N' is symmetric.

We have thus proved the existence of an indecomposable nilpotent symmetric matrix N of any order n. Every nilpotent matrix is similar to a direct sum of indecomposable nilpotent matrices, we form such a direct sum and obtain the result:

THEOREM 40. *There exist symmetric nilpotent matrices with any indices.*

6. **Primary matrices.** We shall require the following lemma:

LEMMA 23. *Let $a \neq 0$ be in $\mathfrak{F}$, $e_1 \geqq e_2 \geqq \cdots \geqq e_t \geqq 1$, be integers such that $e_1 > 1$. Then there exists a symmetric nilpotent matrix N with $e_1, \cdots, e_t$ as indices and a symmetric matrix Q commutative with N such that*

$$\begin{pmatrix} Q & 0 \\ 0 & Q(a + N) \end{pmatrix} \tag{141}$$

is semi-definite.

For let us take

$$N = \begin{pmatrix} N_1 & 0 \\ 0 & N_2 \end{pmatrix},$$

where the index of N_1 is $e_1 > 1$. Then we choose $N_1 \neq 0$ to have order e_1 and be indecomposable, N_1 to be the matrix

$$PN_0P^{-1}, \qquad P'P = U(I + N_0)$$

as in §5. If

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad Q_1 = f(N_1),$$

then Q is symmetric and commutative with N. Moreover (141) is semi-definite if and only if

$$\begin{pmatrix} Q_1 & 0 \\ 0 & Q_1(a + N_1) \end{pmatrix} \tag{142}$$

is semi-definite. The matrix (142) is congruent to

$$\begin{pmatrix} P'Q_1P & 0 \\ 0 & P'Q_1(a + N_1)P \end{pmatrix}. \tag{143}$$

Now $P'f(N_1)P = P'PP^{-1}f(N_1)P = P'Pf(N_0) = U(I+N_0)f(N_0)$. If e_1 is even, we take $Q_1 = I+N_1$ and have

$$P'f(N_1)P = U(I + N_0)(I + N_0) = U(I + N_0^2)$$

alternate by Theorem 39. Also

$$P'Q_1(a + N_1)P = U(I + N_0^2)(a + N_0) = U(a + aN_0^2 + N_0 + N_0^3)$$

is semi-definite. But then the matrix (143) congruent to (142) is semi-definite, and so is (141). Similarly when e_1 is odd we take $f(N_1) = Q_1 = N_1 + N_1^3$,

$$P'Q_1P = U(N_0 + N_0^3), \quad P'Q_1(a + N_1)P = U(N_0^2 + N_0^4 + aN_0 + aN_0^3).$$

The first of these matrices is alternate, the second is semi-definite since now $e_1 \geqq 3$, and $N_0^2 \neq 0$.

We next prove the lemma:

LEMMA 24. *Let M be a symmetric matrix with elements in an infinite field $\mathfrak{F}$, and let Q be symmetric and commutative with M such that*

$$\begin{pmatrix} Q & 0 \\ 0 & QM \end{pmatrix} \tag{144}$$

is semi-definite. Then the matrix

$$B_0 = \begin{pmatrix} 0 & M \\ I & 0 \end{pmatrix} \tag{145}$$

is similar to a symmetric matrix B, and there exists a matrix $S = S'$ commutative with B such that

$$\begin{pmatrix} S & 0 \\ 0 & SB \end{pmatrix} \tag{146}$$

is semi-definite.

Since $\mathfrak{F}$ is infinite we apply Theorem 8 to obtain a quantity $b \neq 0$ in $\mathfrak{F}$ such that the matrix $b^4I + Q^2M$ is definite. Then

$$Q_0 = \begin{pmatrix} Q & b^2 \\ b^2 & QM \end{pmatrix} \tag{147}$$

is semi-definite since (144) is semi-definite. But

$$|Q_0| = \begin{vmatrix} Q & b^2I \\ b^2I & QM \end{vmatrix} = \begin{vmatrix} 0 & b^2I + b^{-2}Q^2M \\ b^2I & QM \end{vmatrix} = |b^4I + Q^2M| \neq 0, \tag{148}$$

so that Q_0 is a definite symmetric matrix, and we may write

$$Q_0 = P_0' P_0, \tag{149}$$

where P_0 is non-singular. Also

$$Q_0 B_0 = \begin{pmatrix} b^2 & QM \\ QM & b^2 M \end{pmatrix}, \qquad B_0' Q_0 = Q_0 B_0 \tag{150}$$

by direct computation. Then $P_0' P_0 B_0 = B_0' P_0' P_0$.

The matrix $B = P_0 B_0 P_0^{-1}$ is now the desired symmetric matrix. Take $S = (P_0^{-1})' S_0 P_0^{-1}$ congruent to S_0 given by

$$\begin{pmatrix} Q & Q \\ Q & QM \end{pmatrix}. \tag{151}$$

Then S_0 is semi-definite and so is S. The matrix

$$S_0 B_0 = S_0 \begin{pmatrix} 0 & M \\ I & 0 \end{pmatrix} = \begin{pmatrix} Q & QM \\ QM & QM \end{pmatrix} = B_0' S_0 \tag{152}$$

is also semi-definite. Now $SB = BS = (P_0')^{-1} S_0 P_0^{-1} P_0 B_0 P_0^{-1} = (P_0^{-1})' (S_0 B_0) P_0^{-1}$ is congruent to $S_0 B_0$ and is semi-definite. This proves that (146) is semi-definite and completes the proof of our lemma.

We use Lemma 24 to prove a fundamental result:

THEOREM 41. *A primary matrix B with elements in a field $\mathfrak{F}$ of characteristic two is similar in $\mathfrak{F}$ to a symmetric matrix if and only if the minimum function of B is not an irreducible inseparable polynomial.*

For if B has an irreducible inseparable minimum function and B_0 is symmetric and similar to B, the field $\mathfrak{F}[B_0]$ is inseparable, contrary to Theorem 38. Conversely let the minimum function of B be not an irreducible inseparable polynomial. Then the argument of §4 together with Theorem 40 reduces the proof of our theorem to the question of the existence of a symmetric matrix B with arbitrary indices

$$e = e_1 \geqq e_2 \geqq \cdots \geqq e_t \geqq 1, \qquad e > 1, \tag{153}$$

and with

$$(x^{2k} - a)^e \qquad (a \text{ in } \mathfrak{F}) \tag{154}$$

as minimum function, $x^{2k} - a$ irreducible in $\mathfrak{F}$. By Theorem 40 there exists a symmetric nilpotent matrix N with indices $e_1, \cdots, e_t$. The matrix $M = aI_f + N$ has the same indices as N, where $f = e_1 + \cdots + e_t$, and so does

$$\begin{pmatrix} 0 & M \\ I_f & 0 \end{pmatrix} \tag{155}$$

by Theorem 30. We use Lemma 24 and choose the symmetric matrix M so that there exists an M_2 similar to (155) and having the further property of Lemma 24. By Theorem 30 the matrix M_2 has the same indices as N. Its characteristic function is $(x^2-a)^t$. An evident induction yields a matrix M_k which is symmetric and is the desired matrix B.

A field $\mathfrak{F}$ which is perfect has the property that there are no inseparable polynomials over $\mathfrak{F}$. For example every finite field is perfect. But then Lemma 11 and Theorem 41 give the following theorem:

THEOREM 42. *Every square matrix with elements in a perfect field $\mathfrak{F}$ of characteristic two is similar in $\mathfrak{F}$ to a symmetric matrix.*

7. **Orthogonal equivalence of primary matrices.** We have reduced the problem of the orthogonal equivalence of primary symmetric matrices to the case of matrices with characteristic function $(x^{2^k}-a)^t$, $x^{2^k}-a$ irreducible in $\mathfrak{F}$, The case $a=0$ is the case of nilpotent matrices. We may now easily show that two matrices may be similar and not orthogonally equivalent. Since the first part of our problem is that of the orthogonal equivalence of symmetric nilpotent matrices, and since, even in this case, the only criteria are of a necessarily complicated nature, we shall restrict our attention to the nilpotent case. It is the only case occurring when $\mathfrak{F}$ is perfect.

Every symmetric nilpotent matrix A with indices $e_1, \cdots, e_t$ is similar in $\mathfrak{F}$ to

$$N = \operatorname{diag}[N_1, \cdots, N_t], \tag{156}$$

where N_i is nilpotent of index and order e_i and may be taken to be symmetric by Theorem 41. Then N is symmetric and

$$A = L\dot{N}L^{-1}, \qquad L'L = Q, \tag{157}$$

is commutative with N by Lemma 20. This last condition states that

$$Q = (Q_{ij}), \qquad Q_{ij}N_j = N_iQ_{ij} \qquad (i, j = 1, \cdots, t), \tag{158}$$

so that in particular the Q_{ii} are polynomials in N_i. Also by Lemma 21 if

$$A_0 = L_0NL_0^{-1}, \qquad Q_0 = L_0'L_0, \tag{159}$$

and A_0 is similar to A, then A_0 is orthogonally equivalent to A if and only if $Q_0=C'QC$ is congruent to Q under a transformation whose matrix C is commutative with N. These are the formal orthogonal equivalence conditions.

We first study the case where $N=N_1$ is indecomposable. By the proof of Theorem 40 we may choose N so that

$$P'P = U(I + N_0), \qquad N = PN_0P^{-1}, \tag{160}$$

with (140) holding. Now

$$Q = Q(N), \qquad Q_0 = Q_0(N), \qquad C = C(N) \tag{161}$$

are all polynomials in N and our condition is simply

$$Q_0 = C^2 Q. \tag{162}$$

The matrix Q is definite if and only if $P'QP = P'PQ(N_0) = U(I+N_0)Q(N_0)$ is definite. Write

$$f(N_0) = (I + N_0)Q(N_0). \tag{163}$$

Since $I+N_0$ is non-singular there exists a $Q(N_0)$ for any $f(N_0)$. But Theorem 39 gives necessary and sufficient conditions that $Uf(N_0)$ be definite. Moreover if $f_0(N_0) = (I+N_0)Q_0$, then $Uf_0(N_0)$ may be definite for many polynomials $Q_0 \neq QC^2$. In fact if $f = f_1 + N_0 f_2$, then $fC^2 = f_1C^2 + N_0 f_2 C^2 = f_{10} + N_0 f_{20}$ if and only if $f_{10} = f_1C^2, f_{20} = f_2C^2$. For example, if $e = 3$, we may explicitly compute $Uf(N_0)$ which is definite if and only if $a_0 = b_0^2 \neq 0$, $a_2 = b_2^2$,

$$f(N_0) = (b_0 + b_1N)^2 + a_2N_0.$$

We write $c(N_0) = c_0 + c_1N_0 + c_2N_0^2$, $[c(N_0)]^2 = c_0^2 + c_1^2N_0^2$, so that

$$[c(N_0)]^2 f(N_0) = [b_0c_0 + (c_0b_1 + c_1b_0)N_0]^2 + c_0^2a_2N_0.$$

Now $f_0(N_0) = (d_0+d_1N)^2 + d_2N_0 = [c(N_0)]^2 f(N_0)$ for some $c(N_0)$ if and only if $d_0 = b_0c_0$, $d_1 = c_0b_1 + c_1b_0$, $d_2 = c_0^2a_2$. Then $c_0 = d_0b_0^{-1}$ and $c_1 = (d_1 - c_0b_1)b_0^{-1}$ are determined. But d_2 is at our choice and can be given distinct from $c_0^2a_2$. Notice that in general our first condition is $f_{10} = c^2f_1$ which may always have a solution c, but that this solution may not satisfy $f_{20} = f_2c^2$.

THEOREM 43. *There exist similar indecomposable symmetric nilpotent matrices which are not orthogonally equivalent.*

Decomposable nilpotent matrices need not be orthogonally decomposable. A very simple example may be obtained as follows. Let

$$E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad A = \begin{pmatrix} E & E \\ E & E \end{pmatrix}.$$

The matrix A is a nilpotent matrix of index two in $\mathfrak{F}$ of characteristic two and is similar in $\mathfrak{F}$ to

$$\begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix}, \qquad N = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

If A were orthogonally equivalent to a direct sum, the components would be necessarily nilpotent of index two, hence A would be orthogonal to

$$\begin{pmatrix} aN & 0 \\ 0 & bN \end{pmatrix}, \tag{164}$$

a, b not zero and in $\mathfrak{F}$. For the only two-rowed nilpotent symmetric matrices are multiples of the N above, and A has rank two, $ab \neq 0$. But (164) is non-alternate, while A is alternate and cannot even be congruent to (164).

8. **Reduction to primary components.** Consider first the question of reducing a J-symmetric matrix to primary components. We let $\mathfrak{F}$ be an arbitrary field and let J be an involution defined by

$$A^J = E^{-1}A'E, \tag{165}$$

where $E' = \epsilon E$ is non-singular, $\epsilon = \pm 1$. Suppose that

$$A^J = \delta A, \qquad \delta = \pm 1, \tag{166}$$

and that

$$A = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1}, \tag{167}$$

where G and H are relatively prime according to our definition. Then $EA^J = A'E = \delta EA$,

$$\delta EP\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1} = P'^{-1}\begin{pmatrix} G' & 0 \\ 0 & H' \end{pmatrix}P'E, \tag{168}$$

so that

$$P'EP\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} = \begin{pmatrix} \delta G' & 0 \\ 0 & \delta H' \end{pmatrix}P'EP. \tag{169}$$

By Lemma 12 we have

$$P'EP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \qquad E_1' = \epsilon E_1, \qquad E_2' = \epsilon E_2. \tag{170}$$

Also $E_1G = \delta G'E_1$, $E_2H = \delta H'E_2$, so that we have the first part of the following theorem:

THEOREM 44. *Let $E' = \epsilon E$, $A^J = E^{-1}A'E = \delta A$, where $\epsilon = \pm 1$, $\delta = \pm 1$ and the square matrix A has relatively prime components G and H. Then*

$$A = P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1}, \quad P'EP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \quad E_1' = \epsilon E_1, \quad E_2' = \epsilon E_2, \tag{171}$$

and the matrices $G^{J_1} \equiv E_1^{-1}G'E_1 = \delta G$, $H^{J_2} \equiv E_2^{-1}H'E_2 = \delta H$. The components G and H are not unique, but if

$$A = P_0 \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} P_0^{-1}, \tag{172}$$

then $G_0 = Q_1^{-1} G Q_1$, $H_0 = Q_2^{-1} H Q_2$, *and the replacement of* G *and* H *by similar matrices replaces* E_1 *and* E_2 *by congruent matrices such that*

$$P_0' E P_0 = \begin{pmatrix} Q_1' E_1 Q_1 & 0 \\ 0 & Q_2' E_2 Q_2 \end{pmatrix}. \tag{173}$$

The last part of the above is due to Lemma 12 and

$$P^{-1} P_0 \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1} P_0, \qquad P_0 = P \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}. \tag{174}$$

We now let A_0 be a matrix similar to A and such that $A_0^J = \epsilon A_0$. Then

$$A_0 = R \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} R^{-1}, \qquad R'ER = \begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix}. \tag{175}$$

If $A_0 = DAD^{-1}$, where $D^J D = I$, then

$$DP \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} (DP)^{-1} = R \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} R^{-1}, \tag{176}$$

so that

$$DP = R \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}, \qquad G = C_1^{-1} G_0 C_1, \qquad H_0 = C_2^{-1} H_0 C_2. \tag{177}$$

Moreover $E^{-1} D' E D = I$,

$$(DP)' E (DP) = P' E P = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} = \begin{pmatrix} C_1' L_1 C_1 & 0 \\ 0 & C_2' L_2 C_2 \end{pmatrix}. \tag{178}$$

Hence necessarily L_1 and E_1 are congruent; L_2 and E_2 are congruent. Hence we may replace G_0 and H_0 by similar matrices and take $E_1 = L_1$, $E_2 = L_2$. We may now prove the following theorem:

THEOREM 45. *Let* $A_0^J = \delta A_0$ *be similar to* A *of Theorem* 44. *Then* A_0 *is* J*-orthogonally equivalent to* A *only if*

$$A_0 = R \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} R^{-1}, \qquad R'ER = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \tag{179}$$

so that $G_0^{J_1} = \delta G_0$, $H_0^{J_2} = \delta H_0$. *Moreover* A_0 *is* J*-orthogonally equivalent to* A *if and only if* G_0 *is* J_1*-orthogonally equivalent to* G, *and* H_0 *is* J_2*-orthogonally equivalent to* H.

This theorem reduces the problem of J-orthogonal equivalence to the similar problem for primary matrices. For proof we need only take $L_1 = E_1$, $L_2 = E_2$ above and obtain $C_1' E_1 C_1 = E_1$, $C_1^J C_1 = I_1$ and similarly $C_2^J C_2 = I_2$. Thus $G_0 = C_1 G C_1^J$, $H_0 = C_2 H C_2^J$ as desired.

We apply Theorem 45 to our case of orthogonal equivalence in $\mathfrak{F}$ of characteristic two. Suppose that $A = A'$ has components G and H. Then

$$P'IP = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}. \tag{180}$$

One of E_1 and E_2 must be definite by Theorem 7. Assume that E_1 is definite so that we may take $E_1 = I_1$ to be an identity matrix. Then two cases arise. In the first case E_2 is definite; we may take $E_2 = I_2$ and have $P'P = I$; P is an orthogonal matrix and A is orthogonally equivalent to

$$\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}$$

where both G and H are symmetric. Moreover

$$A_0 = P_0 \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} P_0^{-1}, \qquad P_0^{-1} P_0 = I,$$

is orthogonally equivalent to A if and only if G and G_0 are orthogonally equivalent, and H and H_0 are orthogonally equivalent.

We next consider the only remaining case, that where E_2 is an alternate matrix. Then H has even order $2s$, and we may take

$$E_2 = \begin{pmatrix} 0 & I_s \\ I_s & 0 \end{pmatrix}.$$

The matrix H is J_2-symmetric by Theorem 42, and

$$H^{J_2} = E_2^{-1} H' E_2.$$

However G is symmetric. Finally if A_0 is similar to A, then necessarily

$$A_0 = P_0 \begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} P_0^{-1}, \qquad P_0^{-1} P_0 = \begin{pmatrix} I_1 & 0 \\ 0 & E_2 \end{pmatrix},$$

if A_0 is to be orthogonally equivalent to A. We apply this result to prove the following theorem:

THEOREM 46. *There exist symmetric matrices A which have symmetric relatively prime components G, H and are such that A is not orthogonally equivalent to any direct sum*

$$\begin{pmatrix} G_0 & 0 \\ 0 & H_0 \end{pmatrix} \tag{181}$$

for G_0 similar to G, H_0 similar to H.

For let us construct any symmetric matrix G^* whose characteristic function is prime to that of H, where

$$H = \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 0 & I_s \\ I_s & 0 \end{pmatrix}.$$

Then $H^{J_2} = E_2^{-1} H' E_2 = H$ by direct computation. Also

$$\begin{pmatrix} I_{n-2s} & 0 \\ 0 & E_2 \end{pmatrix} = P'P$$

is definite. Hence

$$A = P \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix} P^{-1}$$

is symmetric and has G and H as components. But if A were orthogonally equivalent to (181), we could apply Theorem 44 to see that E_2 is congruent to the identity matrix I_{2s}. But this is impossible since E_2 is alternate.

We now use the form of (180) to prove the theorem:

THEOREM 47. *A matrix A with elements in a field $\mathfrak{F}$ of characteristic two is similar to a symmetric matrix if and only if the minimum function of A is not a product of distinct inseparable irreducible polynomials.*

For every symmetric A is similar to a direct sum of its relatively prime primary components. By (180) and Theorem 45 one of these primary components B_i is similar to a symmetric matrix. But then Theorem 41 implies that the minimum function of B_i, which is a factor of that of A, is not an inseparable irreducible polynomial. Conversely let A have minimum function

$$f(x) = g(x) \cdot h(x),$$

where $g(x)$ and $h(x)$ are relatively prime, $h(x)$ is the product of all irreducible inseparable factors of $f(x)$ whose second power does not divide $f(x)$. By Theorem 41 corresponding to each distinct power $[g_i(x)]^{e_i}$ of an irreducible polynomial occurring in the factorization of $g(x)$ into such factors there corresponds a symmetric matrix G_i which is a primary component of A. Their direct sum is a symmetric component G of A. By Theorem 36 correspond-

* In particular we might take G to define a symmetric separable field, S nilpotent.

ing to each irreducible factor $h_i(x)$ of $h(x)$ there is an indecomposable matrix H_i with $h_i(x)$ as characteristic function and such that E_i is alternate, $E_i^{-1}H_i' E_i = H_i$. The direct sum of the H_i is a matrix H which is a component of A such that A is similar to

$$\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}, \qquad G = G', \qquad E^{-1}H'E = E,$$

where E is the alternate matrix which is the direct sum of the E_i. But

$$\begin{pmatrix} I & 0 \\ 0 & E \end{pmatrix} = P'P; \qquad P\begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}P^{-1}$$

is symmetric by Theorem 44 and is similar to A.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.