CHAPTER VI

SYMMETRIC, SKEW, AND HERMITIAN MATRICES

6.01 Hermitian matrices. If we denote by $\bar{x}$ the matrix which is derived from x by replacing each coordinate by its conjugate imaginary, then x is called a *hermitian* matrix if

$$\bar{x} = x'. \tag{1}$$

We may always set $x = x_1 + ix_2$ where x_1 and x_2 are real and (1) shows that, when x is hermitian,

$$x_1' = x_1, x_2' = -x_2, \tag{2}$$

so that the theory of real symmetric and real skew matrices is contained in that of the hermitian matrix. The following are a few properties which follow immediately from the definition; their proof is left to the reader.

If x and y are hermitian and a is arbitrary, then

$$x + y,\ \bar{x},\ x',\ ax\bar{a}',\ xy + yx,\ i(xy - yx),$$

are all hermitian.

Any matrix x can be expressed uniquely in the form $a + ib$ where $2a = x + \bar{x}'$, $2b = -i(x - \bar{x}')$ are hermitian.

The product of two commutative hermitian matrices is hermitian. In particular, any integral power of a hermitian matrix x is hermitian; and, if $g(\lambda)$ is a scalar polynomial with real coefficients, $g(x)$ is hermitian.

THEOREM 1. *If $a, b, c, \cdots$ are hermitian matrices such that $a^2 + b^2 + c^2 + \cdots = 0$, then $a, b, c, \cdots$ are all* 0.

If $\Sigma a^2 = 0$, its trace is 0; but $\Sigma a^2 = \Sigma a\bar{a}'$ and the trace of the latter is the sum of the squares of the absolute values of the coordinates of $a, b, \cdots$; hence each of these coordinates is 0.

THEOREM 2. *The roots of a hermitian matrix are real and its elementary divisors are simple.*

Let x be a hermitian matrix and $g(\lambda)$ its reduced characteristic function. Since $g(x) = 0$, we have $0 = \bar{g}(\bar{x}) = \bar{g}(x')$ and, since x and x' have the same reduced characteristic function, it follows that $g(\lambda) \equiv \bar{g}(\lambda)$, that is, the coefficients of g are real. Suppose that $\xi_1 = \alpha + i\beta$ $(\beta \neq 0)$ is a root of $g(\lambda)$; then $\xi_2 = \alpha - i\beta \neq \xi_1$ is also a root, and we may set

$$g(\lambda) = (\lambda - \xi_1)(g_1(\lambda) + ig_2(\lambda)) = (\lambda - \xi_2)(g_1(\lambda) - ig_2(\lambda)) \tag{3}$$

where g_1, g_2 are real polynomials of lower degree than g, neither of which is identically 0 since g is real and ξ_1 complex. Now

$$[g_1(x)]^2 + [g_2(x)]^2 = [g_1(x) + ig_2(x)][g_1(x) - ig_2(x)]$$

and this product is 0 since from (3) $\lambda - \xi_1$ is a factor of $g_1(\lambda) - ig_2(\lambda)$ and $(\lambda - \xi_1)(g_1(\lambda) + ig_2(\lambda)) = g(\lambda)$. But, since the coefficients of g_1 and g_2 are real, the matrices $g_1(x)$, $g_2(x)$ are hermitian and, seeing that the sum of their squares is 0, they both vanish by Theorem 1. This is however impossible since $g_1(\lambda)$ is of lower degree than the reduced characteristic function of x. Hence x cannot have a complex root.

To prove that the elementary divisors are simple it is only necessary to show that $g(\lambda)$ has no multiple root. Let

$$g(\lambda) = (\lambda - \xi)^r h(\lambda), \quad h(\xi) \neq 0.$$

If $r > 1$, set $g_1(\lambda) = (\lambda - \xi)^{r-1}h(\lambda)$; then $[g_1(\lambda)]^2$ has $g(\lambda)$ as a factor so that the square of the hermitian matrix $g_1(x)$ is 0. Hence by Theorem 1, $g_1(x)$ is itself 0, which is impossible since the degree of g_1 is less than that of g. It follows that r cannot be greater than 1, which completes the proof of the theorem.

Since the elementary divisors are simple, the canonical form of x is a diagonal matrix. Suppose that $n - r$ roots are 0 and that the remaining roots are $\xi_1, \xi_2, \cdots, \xi_r$; these are of course not necessarily all different. The canonical form is then

$$\begin{matrix} \xi_1 & & & & & & \\ & \xi_2 & & & & & \\ & & \ddots & & & & \\ & & & \xi_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0. \end{matrix}$$

The following theorem is contained in the above results.

THEOREM 3. *A hermitian matrix of rank r has exactly $n - r$ zero roots.*

It also follows immediately that the characteristic equation of x has the form

$$x^n - a_1x^{n-1} + \cdots + (-1)^r a_r x^{n-r} = 0 \quad (a_r \neq 0)$$

where a_i is the elementary symmetric function of the ξ's of degree i. Since a_r is the sum of the principal minors of x of order r, we have

THEOREM 4. *In a hermitian matrix of rank r at least one principal minor of order r is not* 0.

In view of the opening paragraph of this section Theorems 1–4 apply also to real symmetric matrices; they apply also to real skew matrices except that Theorem 2 must be modified to state that the roots are pure imaginaries.

6.02 The invariant vectors of a hermitian matrix. Let H be a hermitian matrix, α_1, α_2 two different roots, and a_1, a_2 the corresponding invariant vectors so chosen that $Sa_i\bar{a}_i = 1$; then, since $Ha_1 = \alpha_1 a_1$, $\bar{H}\bar{a}_1 = \alpha_1\bar{a}_1$, we have

$$\alpha_1 Sa_2\bar{a}_1 = Sa_2\bar{H}\bar{a}_1 = S\bar{H}'a_2\bar{a}_1 = \alpha_2 Sa_2\bar{a}_1$$

and, since $\alpha_1 \neq \alpha_2$, we must have $Sa_2\bar{a}_1 = 0$. Again, if α is a repeated root of order s and $a_1, a_2, \cdots, a_s$ a corresponding set of invariant vectors we may choose these vectors (cf. §1.09) so that $Sa_i\bar{a}_j = \delta_{ij}$. The invariant vectors may therefore be so chosen that they form a unitary set and

$$H = \Sigma\alpha_i a_i S\bar{a}_i. \tag{4}$$

If U is the matrix defined by

$$Ue_i = a_i \quad (i = 1, 2, \cdots, n), \tag{5}$$

then

$$U\bar{U}' = 1, \tag{6}$$

so that U is unitary, and if A is the diagonal matrix $\sum_1^r \alpha_i e_i Se_i$, then

$$H = UA\bar{U}' = UAU^{-1}. \tag{7}$$

We may therefore say:

THEOREM 5. *A hermitian matrix can be transformed to its canonical form by a unitary matrix.*

If H is a real symmetric matrix, the roots and invariant vectors are real, and hence U is a real orthogonal matrix. Hence

THEOREM 6. *A real symmetric matrix can be transformed to its canonical form by a real orthogonal matrix.*

If T is a real skew matrix, $h = iT$ is hermitian. The non-zero roots of T are therefore pure imaginaries and occur in pairs of opposite sign. The invariant vectors corresponding to the zero roots are real and hence by the proof just given they may be taken orthogonal to each other and to each of the other invariant vectors. Hence, if the rank of T is r, we can find a real orthogonal matrix which transforms it into a form in which the last $n - r$ rows and columns are zero.

Let $i\alpha$ be a root of T which is not 0 and $a = b + ic$ a corresponding invariant vector; then $Ta = i\alpha a$ so that

$$Tb = -\alpha c, \quad Tc = \alpha b.$$

Hence

$$-\alpha Sc^2 = ScTb = -SbTc = -\alpha Sb^2, \quad -\alpha Sbc = SbTb = 0,$$

which gives

$$Sb^2 = Sc^2, \quad Sbc = 0.$$

We can then choose a so that $Sb^2 = Sc^2 = 1$ and can therefore find a real orthogonal matrix which transforms T into

$$\begin{matrix} 0 & \alpha_1 & & & \\ -\alpha_1 & 0 & & & \\ & & 0 & \alpha_2 & \\ & & -\alpha_2 & 0 & \\ & & & & \ddots \end{matrix} = \Sigma\alpha_j(e_{2j-1}Se_{2j} - e_{2j}Se_{2j-1}). \tag{8}$$

We have therefore the following theorem.

THEOREM 7. *If T is a real skew matrix, its non-zero roots are pure imaginaries and occur in pairs of opposite sign; its rank is even; and it can be transformed into the form* (8) *by a real orthogonal matrix.*

6.03 **Unitary and orthogonal matrices.** The following properties of a unitary matrix follow immediately from its definition by equation (6).

The product of two unitary matrices is unitary.

The transform of a hermitian matrix by a unitary matrix is hermitian.

The transform of a unitary matrix by a unitary matrix is unitary.

If H_1 and H_2 are hermitian, a short calculation shows that

$$U_1 = \frac{1 - iH_1}{1 + iH_1}, \qquad U_2 = \frac{iH_2 - 1}{iH_2 + 1} \tag{9}$$

are unitary (the inverses used exist since a hermitian matrix has only real roots). Solving (9) for H_1 and H_2 on the assumption that the requisite inverses exist we get

$$H_1 = \frac{i(U_1 - 1)}{U_1 + 1}, \qquad H_2 = \frac{i(U_2 + 1)}{U_2 - 1}.$$

These are hermitian when U_1 and U_2 are unitary, and therefore any unitary matrix which has no root equal to -1 can be put in the first form while the second can be used when U has no root equal to 1.

THEOREM 8. *The absolute value of each root of a unitary matrix equals* 1.

Let $\alpha + i\beta$ be a root and $a + ib$ a corresponding invariant vector; then

$$U(a + ib) = (\alpha + i\beta)(a + ib), \; \overline{U}(a - ib) = (\alpha - i\beta)(a - ib).$$

Hence

$$Sa^2 + Sb^2 = S(a + ib)(a - ib) = S(a + ib)U'\overline{U}(a - ib) = SU(a + ib)\overline{U}(a - ib)$$
$$= (\alpha^2 + \beta^2)S(a + ib)(a - ib) = (\alpha^2 + \beta^2)S(a^2 + b^2),$$

so that $\alpha^2 + \beta^2 = 1$.

Corollary.

$$U^{-1}(a + ib) = (\alpha - i\beta)(a + ib),$$
$$U'(a - ib) = \overline{U}^{-1}(a - ib) = (\alpha + i\beta)(a - ib). \tag{10}$$

THEOREM 9. *The elementary divisors of a unitary matrix are simple.*

For, if we have
$U(a_1 + ib_1) = (\alpha + i\beta)(a_1 + ib_1)$, $U(a_2 + ib_2) = (\alpha + i\beta)(a_2 + ib_2) + (a_1 + ib_1)$,
then from (10)

$$(\alpha + i\beta)S(a_1 - ib_1)(a_2 + ib_2) = SU'(a_1 - ib_1)(a_2 + ib_2) = S(a_1 - ib_1)U(a_2 + ib_2)$$
$$= (\alpha + i\beta)S(a_1 - ib_1)(a_2 + ib_2) + S(a_1 - ib_1)(a_1 + ib_1)$$

which is impossible since $S(a_1 - ib_1)(a_1 + ib_1) = Sa_1^2 + Sb_1^2 \neq 0$.

The results of this section apply immediately to real[1] orthogonal matrices; it is however convenient to repeat (9).

THEOREM 10. *If U is a real orthogonal matrix, it can be expressed in the form $(1 + T)/(1 - T)$ if it has no root equal to 1 and in the form $(T - 1)/(T + 1)$ if it has no root equal to -1, the matrix T being a real skew matrix in both cases; and any real matrix of this form which is not infinite, is a real orthogonal matrix.*

6.04 Hermitian and quasi-hermitian forms. Let H be a hermitian matrix and $x = u + iv$ a vector of which u and iv are the real and imaginary parts; then the bilinear form $f = S\bar{x}Hx$ is called a *hermitian form*. Such a form is real since

$$\bar{f} = Sx\overline{H}\bar{x} = SxH'\bar{x} = S\bar{x}Hx = f.$$

In particular, if x and H are real, f is a real quadratic form and, if $H = iT$ is a pure imaginary, T is skew and $f = 0$.

If we express H in terms of its invariant vectors, say $H = \Sigma\alpha_i a_i S\bar{a}_i$ and then put $x = \Sigma\xi_i a_i$, the form f becomes $f = \Sigma\alpha_i\xi_i\bar{\xi}_i$. This shows that, if all the roots of H are positive, the value of f is positive for all values of x; H and f are then said to be *positive definite*. Similarly if all the roots are negative, H and f are *negative definite*. If some roots are 0 so that f vanishes for some value of $x \neq 0$, we say that H and f are *semi-definite*, positive or negative as the case may be. It follows immediately that, when H is semi-definite, $S\bar{x}Hx$ can only vanish if $Hx = 0$.

[1] The first part of the theorem applies also to complex orthogonal matrices.

THEOREM 11. *If H and K are hermitian and H is definite, the elementary divisors of $H\lambda - K$ are real and simple.*

Since $H\lambda - K$ and $-(H\lambda - K)$ are equivalent, we may suppose that H is positive definite. Its roots are then positive so that

$$H^{\frac{1}{2}} = \Sigma\alpha_i^{\frac{1}{2}}a_iS\bar{a}_i$$

has real roots and hence is also hermitian so that $H^{-\frac{1}{2}}KH^{-\frac{1}{2}}$ is hermitian. But

$$H\lambda - K = H^{\frac{1}{2}}(\lambda - H^{-\frac{1}{2}}KH^{-\frac{1}{2}})H^{\frac{1}{2}}$$

so that $H\lambda - K$ is equivalent to $\lambda - H^{-\frac{1}{2}}KH^{-\frac{1}{2}}$ which has real and simple elementary divisors by Theorem 2.

In order to include the theory of complex symmetric matrices we shall now define a type of matrix somewhat more general than the hermitian matrix and closely connected with it. If $A = A(\lambda)$ is a matrix whose coefficients are analytic functions of a scalar variable λ, we shall call it *quasi-hermitian* if

$$A'(\lambda) = A(-\lambda). \tag{11}$$

For convenience we shall set $A''(\lambda)$ for $A(-\lambda)$ with a similar convention for vector functions.

If $A = B + \lambda C$, B and C being functions of λ^2, then $A'' = B - \lambda C$ so that, if A is quasi-hermitian, B is symmetric and C skew just as in the case of a hermitian matrix except that now B and C are not necessarily real. If A is any matrix,

$$2P' \equiv A' + A'' = 2P'', \qquad 2Q' \equiv (A' - A'')/\lambda = 2Q''$$

so that any matrix can be expressed in the form $P + \lambda Q$ where P and Q are quasi-hermitian.

If $x = u + \lambda v$, where u and v are vectors which are functions of λ^2 and if A is quasi-hermitian, then

$$f(\lambda) = Sx''Ax = f(-\lambda) \tag{12}$$

is called a *quasi-hermitian form*. Again, if $|1 + \lambda A| \not\equiv 0$, and we set $Q = (1 - \lambda A)/(1 + \lambda A)$, then

$$Q' = \frac{(1 - \lambda A')}{1 + \lambda A'} = \frac{1 - \lambda A''}{1 + \lambda A''} = (Q'')^{-1}$$

so that

$$Q'Q'' = 1. \tag{13}$$

We shall call such a matrix *quasi-orthogonal.*

6.05 Reduction of a quasi-hermitian form to the sum of squares. We have seen in §5.06 that any matrix A of rank r can be expressed in the form

$$A = \sum_{s=1}^{r} A_sy_s \frac{SA_s'x_s}{Sx_sA_sy_s} \tag{14}$$

where

$$A_{s+1} = A_s - A_s y_s \frac{SA'_s x_s}{Sx_s A_s y_s}, \quad A_1 = A, \quad Sx_s A_s y_s \neq 0,$$

and the null space of A_{s+1} is obtained by adding $(y_1, y_2, \cdots, y_s)$ to the null space of A and the null space of A'_{s+1} by adding $(x_1, x_2, \cdots, x_s)$ to that of A'.

Suppose now that A is quasi-hermitian and replace y_s, x_s by z_s, z''_s and set $z_s = u_s + \lambda v_s$, $A_s = B_s + \lambda C_s$ so that

$$Sz''_s A_s z_s = Su_s B_s u_s + \lambda^2(2Su_s C_s v_s - Sv_s B_s v_s)$$

and, so long as A_s is not 0, we can clearly choose z_s so that $Sz''_s A_s z_s \not\equiv 0$. Each matrix A_s is then quasi-hermitian since $A'_s = A''_s$, and

$$A = \sum_1^r \frac{A_s z_s S A''_s z''_s}{Sz''_s A_s z_s}. \tag{15}$$

If x is an arbitrary vector and

$$f = f(\lambda^2) = SxAx'', \varphi_s(\lambda) = SxA_s z_s = \psi_s(\lambda^2) + \lambda\chi_s(\lambda^2)$$

then ψ_s and χ_s are linear functions of the coordinates of x which are linearly independent and

$$f = \sum_s \frac{SxA_s z_s S A''_s z''_s \cdot x''}{Sz''_s A_s z_s} = \sum \frac{\varphi_s(\lambda)\varphi''_s(\lambda)}{Sz''_s A_s z_s} = \sum \frac{\psi_s^2(\lambda^2) - \lambda^2\chi_s^2(\lambda^2)}{Sz''_s A_s z_s} \tag{16}$$

which is the required expression for $f(\lambda^2)$ in terms of squares.

If A is hermitian, then $\lambda = i$ and ψ_s, χ_s, $Sz''_s A_s z_s = S\bar{z}_s A_s z_s$ are real and, if $S\bar{z}_s A_s z_s = \alpha_s^{-1}$, (16) becomes

$$f = \sum_1^r \alpha_s \varphi_s \bar{\varphi}_s. \tag{17}$$

If $\lambda = 0$, then A is symmetric and

$$f = SxAx = \sum_1^r \alpha_s \varphi_s^2 \tag{18}$$

where the terms are all real if A is real.

In terms of the matrices themselves these results may be expressed as follows.

THEOREM 12. *If A is a hermitian matrix of rank r, there exist an infinity of sets of vectors a_s and real constants α_s such that*

$$A = \sum_1^r \alpha_s a_s S\bar{a}_s; \tag{19}$$

and, if A is symmetric, there exists an infinity of sets of vectors a_s and constants α_s such that

$$A = \sum_{1}^{r} \alpha_s a_s S a_s \tag{20}$$

a_s and α_s being real if A is real.

If π of the α's in (19) are positive and ν are negative, the difference $\sigma = \pi - \nu$ is called the *signature* of A. A given hermitian matrix may be brought to the form (19) in a great variety of ways but, as we shall now show, the signature is the same no matter how the reduction is carried out. Let K_1 be the sum of the terms in (19) for which α_s is positive and $-K_2$ the sum of the terms for which it is negative so that $A = K_1 - K_2$; the matrices K_1 and K_2 are positive semi-definite and, if k_1 and k_2 are their ranks, we have $r = k_1 + k_2$. Suppose that by a different method of reduction we get $A = M_1 - M_2$ where M_1 and M_2 are positive semi-definite matrices of ranks m_1 and m_2 and $m_1 + m_2 = r$; and suppose, if possible, that $k_2 < m_2$. The orders of the null spaces of K_2 and M_1 relative to the right ground of A are $r - k_2$ and $r - m_1 = m_2$ and, since $r - k_2 + m_2 > r$, there is at least one vector x in the ground of A which is common to both these null spaces, that is,

$$Ax = K_1 x = -M_2 x \neq 0,$$

and hence $S\bar{x}K_1x = -S\bar{x}M_2x$. But both K_1 and M_2 are positive semi-definite; hence we must have $S\bar{x}K_1x = 0$ which by §6.04 entails $K_1x = 0$. We have therefore arrived at a contradiction and so must have $k_2 = m_2$ which is only possible when the signature is the same in both cases.

In the case of a skew matrix the reduction given by (16) is not convenient and it is better to modify it as follows. Let $A' = -A$ and set

$$\begin{aligned} A_{s+1} &= A_s + \frac{A_s y_s S A_s x_s}{S x_s A_s y_s} - \frac{A_s x_s S A_s y_s}{S x_s A_s y_s}, \\ A_1 &= A, \quad S x_s A_s y_s \neq 0. \end{aligned} \tag{21}$$

So long as $A_s \neq 0$, the condition $Sx_sA_sy_s \neq 0$ can always be satisfied by a suitable choice of x_s and y_s and it is easily proved as in §5.06 that the null space of A_{s+1} is obtained from that of A_s by adding x_s, y_s; also A_s is skew so that we must necessarily have $x_s \neq y_s$. It follows that the rank of A is even and

$$A = \sum_{1}^{r/2} \frac{A_s x_s S A_s y_s - A_s y_s S A_s x_s}{S x_s A_s y_s} = \Sigma \alpha_s (a_{2s-1} S a_{2s} - a_{2s} S a_{2s-1}) \tag{22}$$

where each term in the summation is a skew matrix of rank 2 and

$$\alpha_s^{-1} = S x_s A_s y_s, \qquad a_{2s-1} = A_s x_s, \qquad a_{2s} = A_s y_s.$$

This form corresponds to the one given in Theorem 12 for symmetric matrices.

If we put

$$T = \sum_{1}^{r/2} (e_{2s-1}Se_{2s} - e_{2s}Se_{2s-1}) = -T'$$

$$R = \sum_{1}^{r/2} \alpha_s(e_{2s-1}Se_{2s} - e_{2s}Se_{2s-1}) = -R' \tag{23}$$

$$P = \sum_{1}^{r/2} (\alpha_s a_{2s}Se_{2s} + a_{2s-1}Se_{2s-1}), \qquad Q = \sum_{1}^{r/2} (a_{2s}Se_{2s} + a_{2s-1}Se_{2s-1})$$

then (22) may be put in the form $A = PTP' = QRQ'$. When $r = n$, the determinant of T equals 1 and therefore $| A | = | P |^2$. The following theorem summarizes these results.

THEOREM 13. *If A is a skew matrix of rank r, then* (i) *r is even;* (ii) *A can be expressed by rational processes in the form*

$$A = \sum_{1}^{r/2} \alpha_s(a_{2s-1}Sa_{2s} - a_{2s}Sa_{2s-1}) = PTP' = QRQ' \tag{24}$$

where P, Q, R and T are given by (23)*;* (iii) *if $r = n$, the determinant of A is a perfect square, namely $| P |^2$;* (iv) *if x and y are any vectors and $w = \Sigma\alpha_s \mid a_{2s-1}a_{2s} \mid$, then*

$$SxAy = S \mid xy \mid w. \tag{25}$$

The following theorem contains several known properties of hermitian matrices.

THEOREM 14. *If $T(A)$ is an associated matrix for which $T'(A) = T(A')$, then, when A is quasi-hermitian, $T(A)$ is also quasi-hermitian.*

For $A' = A''$ gives $T'(A) = T(A') = T(A'') = T''(A)$.

Particular cases of interest are: If A is hermitian, $T(A)$ is hermitian. If $T(\mu A) = \mu^s T(A)$ and A is skew, then $T(A)$ is skew if s is odd, symmetric if s is even.

6.06 The Kronecker method of reduction. Let $A = \sum_{1}^{r} x_iSy_i$ be a quasi-hermitian matrix of rank r; then

$$\Sigma y_iSx_i = A' = A'' = \Sigma x_i''Sy_i'', \tag{26}$$

from which it follows that y_i is linearly dependent on $x_1'', x_2'', \cdots, x_r''$, say

$$y_i = \sum_{j=1}^{r} q_{ij}x_j'', \qquad | q_{ij} | \neq 0, \qquad (i = 1, 2, \cdots, r).$$

Using this value of y_i we have

$$A = \Sigma q_{ij} x_i S x''_j, \qquad A' = \Sigma q_{ji} x''_i S x_j, \qquad A'' = \Sigma q''_{ij} x''_i S x_j$$

and therefore

$$q_{ij} = q''_{ji}. \tag{27}$$

Further, since $| q_{ij} | \neq 0$, we can find s_{ij} $(i, j = 1, 2, \cdots, r)$ so that

$$\sum_j q_{ij} s_{jk} = \delta_{ik}$$

and then (27) gives $s_{ij} = s''_{ji}$.

Let $x_1, \cdots, x_r, x_{r+1}, \cdots, x_n$ be a basis and $z_1, z_2, \cdots, z_n$ the reciprocal basis. Then, if $w_i = \sum_1^r s_{ji} z''_j$, the basis reciprocal to $y''_1, \cdots, y''_r, x_{r+1}, \cdots, x_n$ is $w''_1, \cdots, w''_r, z_{r+1}, \cdots, z_n$. Hence

$$P = \sum_1^r w_i S z_i = \Sigma s_{ij} z''_i S z_j$$

is quasi-hermitian. Further, if $u = \sum_1^r \xi_i x_i$, then $Su''Pu = \Sigma \xi''_i \xi_j s_{ij}$; and we can choose u so that this form is not 0. We also have

$$AP = \sum_1^r x_i S y_i \sum_1^r w_j S z_j = \sum_1^r x_i S z_i,$$

whence $APu = u$.

Let

$$A_{s+1} = A_s - \frac{u_s S u''_s}{S u''_s P_s u_s}, \qquad A_1 = A, \qquad P_1 = P, \tag{28}$$

where P_s is formed from A_s in the same way as P is from A and u_s is a vector of the left ground of A_s such that $Su''_s P_s u_s \neq 0$; also, as above, $A_s P_s u = u$ for any vector u in the left ground of A_s and A_s in quasi-hermitian. The vector u''_s belongs to the right ground of A_s and therefore every vector of the null space of A_s lies in the null space of A_{s+1}; also

$$\begin{aligned} A_{s+1} P_s u_s &= A_s P_s u_s - u_s \frac{S u''_s P_s u_s}{S u''_s P_s u_s} \\ &= u_s - u_s = 0. \end{aligned}$$

Hence the null space of A_{s+1} is derived from that of A_s by adding $P_s u_s$ to it. It then follows as in §6.06 that A can be expressed in the form

$$A = \sum_1^r \frac{u_s S u''_s}{S u''_s P_s u_s} \tag{29}$$

which is analogous to (16) and may be used in its place in proving Theorem 12.

We may also note that, if Q is the matrix defined by $Q'e_j = x_j$ ($j = 1, 2, \cdots, n$), then

$$A = Q' \sum_1^r q_{ij} e_i S e_j Q'' = Q'BQ''$$

where B is the quasi-hermitian matrix $\sum_1^r q_{ij} e_i S e_j$. It may be shown by an argument similar to that used for hermitian matrices that a basis for the x's may be so chosen that Q is quasi-orthogonal provided A is real.

6.07 Cogredient transformation. If $SxAy$ and $SxBy$ are two bilinear forms, the second is said to be derived from the first by a cogredient transformation if there exists a non-singular matrix P such that $SxAy \equiv SPxBPy$, that is,

$$A = P'BP. \tag{30}$$

When this relation holds between A and B, we shall say they are *cogredient*.

From (30) we derive immediately $A' = P'B'P$ and therefore, if

$$R = \frac{A + A'}{2} = R', \qquad S = \frac{A - A'}{2} = -S',$$

$$U = \frac{B + B'}{2} = U', \qquad V = \frac{B - B'}{2} = -V',$$

then

$$R + \lambda S = P'(U + \lambda V)P$$

so that $R + \lambda S$ and $U + \lambda V$ are strictly equivalent.

Suppose conversely that we are given that $R + \lambda S$ and $U + \lambda V$, which are quasi-hermitian, are strictly equivalent so that there exist constant non-singular matrices p, q such that

$$R + \lambda S = p(U + \lambda V)q$$

or

$$R = pUq, \qquad S = pVq; \tag{31}$$

then, remembering that R and U are symmetric, S and V skew, we ha

$$R = q'Up', \qquad S = q'Vp'$$

Equating these two values of R and S, respectively, we get

$$(q')^{-1}pU = Up'q^{-1}, \qquad (q')^{-1}pV = Vp'q^{-1}$$

or, if W stands for U or V indifferently, and

$$J = (q')^{-1}p, \tag{32}$$

we have

$$JW = WJ',$$

repeated application of which gives

$$J^r W = W(J')^r.$$

From this it follows that, if $f(\lambda)$ is a scalar polynomial,

$$f(J)W = Wf(J') = W(f(J))'. \tag{33}$$

In particular, since $|J| \neq 0$, we may choose $f(\lambda)$ so that $f(J)$ is a square root of J and, denoting this square root by K, we have $KW = WK'$ or

$$W = K^{-1}WK', \qquad K^2 = J, \qquad (W = U \text{ or } V).$$

Using this in (31) we have

$$R = pK^{-1}UK'q, \qquad S = pK^{-1}VK'q$$

and from (32) $p = q'J = q'K^2$ or

$$pK^{-1} = q'K = (K'q)'.$$

Hence, if we put $P' = q'K$, there follows

$$R = P'UP, \qquad S = P'VP$$

or

$$A = R + S = P'(U + V)P = P'BP.$$

We therefore have the following theorem, which is due to Kronecker.

THEOREM 15. *A necessary and sufficient condition that A and B be cogredient is that $A + \lambda A'$ and $B + \lambda B'$ shall be strictly equivalent.*

If A and B are symmetric, these polynomials become $A(1 + \lambda)$ and $B(1 + \lambda)$ which are always strictly equivalent provided the ranks of A and B are the same. Hence quadratic forms of the same rank are always cogredient, as is also evident from Theorem 12 which shows in addition that P may be taken real if the signatures are the same.

The determination of P from (31) is unaltered if we suppose S symmetrical instead of skew, or R skew instead of symmetrical. Hence

THEOREM 16. *If R, S, U, V are all symmetric or all skew, and if $R + \lambda S$ and $U + \lambda V$ are strictly equivalent, we can find a constant non-singular matrix P such that*

$$R + \lambda S = P'(U + \lambda V)P,$$

that is, the corresponding pairs of forms are cogredient.

In the case of a hermitian form $S\bar{x}Ax$ changing x into Px replaces A by $\bar{P}'AP$ and we have in place of (30)

$$A = \bar{P}'BP. \tag{34}$$

If we put $B = \Sigma\beta_s b_s S\bar{b}_s$, then

$$\bar{P}'BP = \Sigma\beta_s \bar{P}'b_s S\bar{b}_s P = \Sigma\beta_s \bar{P}'b_s S P'\bar{b}_s = \Sigma\beta_s c_s S\bar{c}_s.$$

where $c_s = \bar{P}'b_s$. Equation (34) can therefore hold only if the signature as well as the rank is the same for B as for A. Conversely, if $A = \Sigma\alpha_s a_s S\bar{a}_s$ and A and B have the same signature and rank the notation may be so arranged that α_s and β_s have the same signs for all s; then any matrix for which

$$\bar{P}'b_s = \left(\frac{\alpha_s}{\beta_s}\right)^{\frac{1}{2}} a_s \quad (s = 1, 2, \cdots, r), \qquad |P| \neq 0$$

where r is the common rank of A and B, clearly satisfies (33).[2] Hence

THEOREM 17. *Two hermitian forms are cogredient if, and only if, they have the same rank and signature.*

The reader will readily prove the following extension of Theorem 16 by the aid of the artifice used in the proof of Theorem 11

THEOREM 18. *If A, B, C, D are hermitian matrices such that $A + \lambda B$ and $C + \lambda D$ are* (i) *equivalent* (ii) *both definite for some value of λ, there exists a constant non-singular matrix P such that*

$$A + \lambda B = \bar{P}'(C + \lambda D)P.$$

6.08 **Real representation of a hermitian matrix.** Any matrix $H = A + iB$ in which A and B are real matrices of order n can be represented as a real matrix of order $2n$. For the matrix of order 2

$$i_2 = \begin{Vmatrix} 0 & -1 \\ 1 & 0 \end{Vmatrix}$$

satisfies the equation $i_2^2 = -1$ and, on forming the direct product of the original set of matrices of order n and a set of order 2 in which i_2 lies, we get a set of order $2n$ in which H is represented by

$$\mathfrak{H} = A + i_2 B = \begin{Vmatrix} A & -B \\ B & A \end{Vmatrix}$$

As a verification of this we may note that

$$\begin{Vmatrix} A & -B \\ B & A \end{Vmatrix} \begin{Vmatrix} C & -D \\ D & C \end{Vmatrix} = \begin{Vmatrix} AC - BD & -(AD + BC) \\ AD + BC & AC - BD \end{Vmatrix}$$

which corresponds to

$$(A + iB)(C + iD) = AC - BD + i(AD + BC).$$

[2] The proof preceding Theorem 15 generalizes readily up to equation (33); at that point, however, if $K = f(J)$, we require $\bar{K}' = f(\bar{J}')$, which is only true when the coefficients of $f(\lambda)$ are real.

This representation has the disadvantage that a complex scalar $\alpha + i\beta$ is represented by

$$\begin{Vmatrix} \alpha & -\beta \\ \beta & \alpha \end{Vmatrix}$$

which is not a scalar matrix although it is commutative with every matrix of the form $\mathfrak{H}$. Consequently, if H has a complex root, this root does not correspond to a root of $\mathfrak{H}$. If, however, all the roots of H are real, the relation $HK = \alpha K$ is represented by $\mathfrak{H}\mathfrak{K} = \alpha\mathfrak{K}$ when α is real so that α is a root of both H and $\mathfrak{H}$.

To prove the converse of this it is convenient to represent the vector $x + iy$ in the original space by (x, y) in the extended space. Corresponding to

$$(A + iB)(x + iy) = Ax - By + i(Bx + Ay)$$

we then have

$$\begin{Vmatrix} A & -B \\ B & A \end{Vmatrix} (x, y) = (Ax - By, Bx + Ay).$$

If therefore $\mathfrak{H}$ has a real root α and (x, y) is a corresponding invariant vector so that

$$\mathfrak{H}(x, y) = \alpha(x, y) = (\alpha x, \alpha y),$$

we have

$$Ax - By = \alpha x, \qquad Bx + Ay = \alpha y,$$

which gives

$$(A + iB)(x + iy) = \alpha(x + iy).$$

It follows that invariant vectors in the two representations correspond provided they belong to real roots. This gives

THEOREM 19. *To every real root of $H = A + iB$ there corresponds a real root of*

$$\mathfrak{H} = \begin{Vmatrix} A & -B \\ B & A \end{Vmatrix}$$

and vice-versa.

In this representation $\bar{H}$ and H' correspond to

$$\begin{Vmatrix} A & B \\ -B & A \end{Vmatrix} \qquad \begin{Vmatrix} A' & -B' \\ B' & A' \end{Vmatrix},$$

respectively, and hence, if H is hermitian, $B' = -B$ so that $\mathfrak{H}$ is symmetric. The theory of hermitian matrices of order n can therefore be made to depend on that of real symmetric matrices of order $2n$. For example, if we have proved of real symmetric matrices that they have real roots and simple elementary divisors, it follows that the same is true of hermitian matrices, thus reversing the order of the argument made in §6.01.