CHAPTER V

COMPOUND MATRICES

5.01 In chapter I it was found necessary to consider the adjoint of A which is a matrix whose coordinates are the first minors of $| A |$. We shall now consider a more general class of matrices, called compound matrices, whose coordinates are minors of $| A |$ of the rth order; before doing so, however, it is convenient to extend the definition of Sxy to apply to vectors of higher grade.

5.02 **The scalar product** Let $x_i = \Sigma \xi_{ij} e_j$, $y_i = \Sigma \eta_{ij} e_j$ $(i = 1, 2, \cdots)$ be arbitrary vectors, then, by equation (37) §1.11 we have

$$(1) \qquad | x_1 x_2 \cdots x_r | = \sum_{(i)}^{*} | \xi_{1i_1} \xi_{2i_2} \cdots \xi_{ri_r} || e_{i_1} e_{i_2} \cdots e_{i_r} |,$$

and hence it is natural to extend the notion of the scalar product by setting

$$(2) \quad S | x_1 x_2 \cdots x_r || y_1 y_2 \cdots y_r | = \sum_{(i)}^{*} | \xi_{1i_1} \xi_{2i_2} \cdots \xi_{ri_r} || \eta_{1i_1} \eta_{2i_2} \cdots \eta_{ri_r} |.$$

We then have the following lemma which becomes the ordinary rule for multiplying together two determinants when $r = n$.

LEMMA 1.

$$(3) \qquad S | x_1 x_2 \cdots x_r || y_1 y_2 \cdots y_r | = | Sx_i y_j |.$$

For $S | x_1 x_2 \cdots x_r || e_{i_1} e_{i_2} \cdots e_{i_r} | = | \xi_{1i_1} \xi_{2i_2} \cdots \xi_{ri_r} |$, hence

$$S | x_1 x_2 \cdots x_r || y_1 e_{i_2} \cdots e_{i_r} | = \sum_{i_1} \eta_{1i_1} | \xi_{1i_1} \cdots \xi_{ri_r} |$$

$$= \left| \left(\sum_{i_1} \eta_{1i_1} \xi_{1i_1} \right) \xi_{2i_2} \cdots \xi_{ri_r} \right| = | Sx_1 y_1, \xi_{2i_2} \cdots \xi_{ri_r} |;$$

again

$$S | x_1 x_2 \cdots x_r || y_1 y_2 e_{i_3} \cdots e_{i_r} | = \sum_{i_2} \eta_{2i_2} S | x_1 \cdots x_r || y_1 e_{i_2} \cdots e_{i_r} |$$

$$= \sum_{i_2} \eta_{2i_2} | Sx_1 y_1 \, \xi_{2i_2} \cdots \xi_{ri_r} |$$

$$= | Sx_1 y_1 \, Sx_2 y_2 \, \xi_{3i_3} \cdots \xi_{ri_r} |.$$

The lemma follows easily by a repetition of this process.

The Laplace expansion of a determinant can clearly be expressed as a scalar product. This is most easily done by introducing the notion of the complement of a vector relative to the fundamental basis. If $i_1, i_2, \cdots, i_r$ is a sequence of distinct integers in natural order each less than or equal to n and $i_{r+1}, \cdots, i_n$ the remaining integers up to and including n, also arranged in natural order, the complement of $| e_{i_1} e_{i_2} \cdots e_{i_r} |$ relatively to the fundamental basis is defined as[1]

$$(4) \qquad | e_{i_1} e_{i_2} \cdots e_{i_r} |_c = (-1)^{\Sigma i_\alpha + r(r+1)/2} | e_{i_{r+1}} e_{i_{r+2}} \cdots e_{i_n} |$$

and the complement of $| x_1 x_2 \cdots x_r |$ by

$$(5) \qquad | x_1 x_2 \cdots x_r |_c = \sum_{(i)}{}^{*} | \xi_{1i_1} \xi_{2i_2} \cdots \xi_{ri_r} || e_{i_1} e_{i_2} \cdots e_{i_r} |_c,$$

which is a vector of grade $n - r$.

Laplace's expansion of a determinant in terms of minors of order r can now be expressed in the following form.

LEMMA 2.

$$(6) \qquad S | x_1 x_2 \cdots x_r |_c | x_{r+1} x_{r+2} \cdots x_n | = | \xi_{11} \xi_{22} \cdots \xi_{nn} | = | S x_i e_j |$$
$$= S | x_1 \cdots x_n || e_1 \cdots e_n | = (-1)^{r(n-r)} S | x_1 x_2 \cdots x_r || x_{r+1} \cdots x_n |_c.$$

Further as an immediate consequence of (5) we have

LEMMA 3.

$$(7) \qquad S | x_1 x_2 \cdots x_r |_c | y_1 y_2 \cdots y_r |_c = S | x_1 x_2 \cdots x_r || y_1 y_2 \cdots y_r |.$$

5.03 **Compound matrices.** If $A = \Sigma a_{ij} e_{ij}$, then, as in (1),

$$| Ax_1 Ax_2 \cdots Ax_r | = \sum_{(j)}{}^{*} | \xi_{1j_1} \cdots \xi_{rj_r} || Ae_{j_1} \cdots Ae_{j_r} |.$$

But $Ae_j = \sum_i a_{ij} e_i$; so a second application of (1) gives

$$| Ax_1 Ax_2 \cdots Ax_r | = \sum_{(i)}{}^{*} \sum_{(j)}{}^{*} | \xi_{1j_1} \cdots \xi_{rj_r} || a_{i_1 j_1} \cdots a_{i_r j_r} || e_{i_1} \cdots e_{i_r} |.$$

But the determinants $| \xi_{1j_1} \cdots \xi_{rj_r} |$ are the coordinates of the r-vector $| x_1 x_2 \cdots x_r |$; hence $| Ax_1 \cdots Ax_r |$ is a linear vector form in $| x_1 x_2 \cdots x_r |$ in the corresponding space of $\binom{n}{r}$ dimensions. We denote this vector function or matrix by $C_r(A)$ and write

$$(8) \qquad | Ax_1 Ax_2 \cdots Ax_r | = C_r(A) | x_1 x_2 \cdots x_r |.$$

We shall call $C_r(A)$ the rth *compound* of A. Important particular cases are

$$(8') \qquad C_1(A) = A, \quad C_n(A) = | A |,$$

[1] The Grassmann notation cannot be conveniently used here since it conflicts with the notation for a determinant. It is sometimes convenient to define the complement of $| e_1 e_2 \cdots e_n |$ as 1.

and, if k is a scalar,

(8″) $$C_r(k) = k^r.$$

THEOREM 1.

(9) $$C_r(AB) = C_r(A)C_r(B).$$

For

$$\begin{aligned} | ABx_1ABx_2 \cdots ABx_r | &= C_r(A) \, | Bx_1Bx_2 \cdots Bx_r | \\ &= C_r(A)C_r(B) \, | x_1x_2 \cdots x_r |. \end{aligned}$$

Corollary. If $| A | \neq 0$, then

(10) $$[C_r(A)]^{-1} = C_r(A^{-1}).$$

THEOREM 2.

(11) $$[C_r(A)]' = C_r(A').$$

For $S \, | x_1x_2 \cdots x_r | \, C_r(A) \, | y_1y_2 \cdots y_r | = | Sx_iAy_j | = | SA'x_iy_j |$

$= S \, | A'x_1 \cdots A'x_r || y_1 \cdots y_r | = S \, | y_1 \cdots y_r | \, C_r(A') \, | x_1 \cdots x_r |.$

THEOREM 3. *If* $A = \sum_1^m a_iSb_i$, *then*

(12) $$C_r(A) = \sum_{(i)}^{*} | a_{i_1}a_{i_2} \cdots a_{i_r} | \, S \, | b_{i_1}b_{i_2} \cdots b_{i_r} |.$$

This theorem follows by direct substitution for A in the left-hand side of (8) It gives a second proof for Theorem 2.

If $r = m$, (12) consists of one term only, and this term is 0 unless m is the rank of A, a property which might have been made the basis of the definition of rank. In particular, if $X = \sum_1^r e_iSx_i$, $Y = \sum_1^r y_iSe_i$, then $C_r(X) = | e_1e_2 \cdots e_r | \, S \, | x_1x_2 \cdots x_r |$, $C_r(Y) = | y_1y_2 \cdots y_r | \, S \, | e_1e_2 \cdots e_r |$ so that $C_r(XY) = | e_1e_2 \cdots e_r | \, S \, | x_1x_2 \cdots x_r || y_1y_2 \cdots y_r | \, S \, | e_1e_2 \cdots e_r |$. But $XY = \sum_{i,j} e_iSx_iy_jSe_j$ so that $C_r(XY) = | Sx_iy_i || e_1e_2 \cdots e_r | \, S \, | e_1e_2 \cdots e_r |$. Comparing these two forms of $C_r(XY)$ therefore gives another proof of the first lemma of §5.02.

If we consider the complement of $| Ax_1Ax_2 \cdots Ax_r |$ we arrive at a new matrix $C^r(A)$ of order $\binom{n}{r}$ which is called the *rth supplementary compound* of A. From (7) and (12) we have

(13) $$\begin{aligned} | Ax_1Ax_2 \cdots Ax_r |_c &= \sum_i^{*} | a_{i_1} \cdots a_{i_r} |_c \, S \, | b_{i_1} \cdots b_{i_r} |_c \, | x_1 \cdots x_r |_c \\ &= C^r(A) \, | x_1x_2 \cdots x_r |_c \end{aligned}$$

and derive immediately the following which are analogous to Theorems 1 and 2.

THEOREM 4.

$$C^r(AB) = C^r(A)C^r(B). \tag{14}$$

THEOREM 5.

$$[C^r(A)]' = C^r(A'). \tag{15}$$

The following theorems give the connection between compounds and supplementary compounds and also compounds of compounds.

THEOREM 6.

$$C^r(A')C_{n-r}(A) = | A | = C^{n-r}(A)C_r(A'). \tag{16}$$

This is the Laplace expansion of the determinant $| A |$. Using equation (6) and setting $| e |$ for $| e_1 e_2 \cdots e_n |$ we have

$$\begin{aligned} | A | S | x_1 x_2 \cdots x_r |_c | x_{r+1} \cdots x_n | &= | A | S | x_1 \cdots x_n | | e | \\ &= S | Ax_1 \cdots Ax_n | | e | \\ &= S | Ax_1 \cdots Ax_r |_c | Ax_{r+1} \cdots Ax_n | \\ &= SC^r(A) | x_1 \cdots x_r |_c C_{n-r}(A) | x_{r+1} \cdots x_n | \\ &= S | x_1 \cdots x_r |_c C^r(A')C_{n-r}(A) | x_{r+1} \cdots x_n | \end{aligned}$$

and, since the x's are arbitrary, the first part of the theorem follows. The second part is proved in a similar fashion.

Putting $r = n - 1$ in (16) gives the following corollary.

Corollary. $\quad \operatorname{adj} A = C^{n-1}(A').$

THEOREM 7.

$$| C_r(A) | = | A |^{\binom{n-1}{r-1}} = | C^r(A) |. \tag{17}$$

For from (16) with A' in place of A, and from the fact that the order of $C_r(A)$ is $\binom{n}{r}$, we have

$$| A |^{\binom{n}{r}} = | C^r(A)C_{n-r}(A') | = | C^r(A) | \, |C_{n-r}(A') |$$

and, since $| A |$ is irreducible when the coordinates of A are arbitrary variables, it follows that $| C^r(A) |$ is a power of $| A |$. Considerations of degree then show that the theorem is true when the coordinates are variables and, since the identity is integral, it follows that it is also true for any particular values of these variables.

THEOREM 8.

$$| A |^{\binom{n-1}{r}} C_s(C_r(A)) = | A |^s C^{\binom{n}{r}-s}(C^{n-r}(A)) \tag{18}$$

$$| A |^{\binom{n-1}{r}} C_s(C^r(A)) = | A |^s C^{\binom{n}{r}-s}(C_{n-r}(A)). \tag{19}$$

Using (15), (16) and (17) we get

$$C_s(C^{n-r}(A'))C^{\binom{n}{r}-s}(C^{n-r}(A)) = |C^{n-r}(A)| = |A|^{\binom{n-1}{r}}$$

therefore

$$\begin{aligned} |A|^{\binom{n-1}{r}} C_s(C_r(A)) &= C_s(C_r(A))C_s(C^{n-r}(A'))C^{n-s}(C^{n-r}(A)) \\ &= C_s(C_r(A)C^{n-r}(A'))C^{n-s}(C^{n-r}(A)) \\ &= C_s(|A|)C^{n-s}(C^{n-r}(A)) \\ &= |A|^sC^{n-s}(C^{n-r}(A)). \end{aligned}$$

Similarly

$$C_s(C_{n-r}(A'))C_s(C^r(A)) = C_s(|A|) = |A|^s$$

and therefore

$$\begin{aligned} |A|^sC^{\binom{n}{r}-s}(C_{n-r}(A)) &= C^{\binom{n}{r}-s}(C_{n-r}(A))C_s(C_{n-r}(A'))C_s(C^r(A)) \\ &= |C_{n-r}(A)|\, C_s(C^r(A)) \\ &= |A|^{\binom{n-1}{r}} C_s(C^r(A)). \end{aligned}$$

An important particular case is $C_s(C^{n-1}(A)) = |A|^{s-1}C^{n-s}(A)$ whence

(20) $$C_s(\text{adj } A) = C_s(C^{n-1}(A')) = |A|^{s-1}C^{n-s}(A').$$

5.04 **Roots of compound matrices.** If A has simple elementary divisors and its roots are $\lambda_1, \lambda_2, \cdots, \lambda_n$, the corresponding invariant vectors being $a_1, a_2, \cdots, a_n$, then the roots of $C_r(A)$ are the products $\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}$ in which no two subscripts are the same and the subscripts are arranged in, say, numerical order; and the invariant vector corresponding to $\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}$ is $|a_{i_1}a_{i_2}\cdots a_{i_r}|$. For there are $\binom{n}{r}$ distinct vectors of this type and

$$C_r(A)\,|a_{i_1}a_{i_2}\cdots a_{i_r}| = |Aa_{i_1}\cdots Aa_{i_r}| = \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}\,|a_{i_1}a_{i_2}\cdots a_{i_r}|.$$

Similarly for $C^r(A)$ the roots and invariants are $\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}$ and $|a_{i_1}a_{i_2}\cdots a_{i_r}|_c$.

It follows from considerations of continuity that the roots are as given above even when the elementary divisors are not simple.

5.05 **Bordered determinants.** Let $A = \|a_{ij}\| = \sum_{j=1}^{n} a_jSe_j,\ a_j = \sum_i a_{ij}e_i$, be any matrix and associate with it two sets of vectors

$$X:\ x_i = \sum_{j=1}^{n} \xi_{ij}e_j, \qquad (i = 1, 2, \cdots, r)$$

$$Y:\ y_i = \sum_{j=1}^{n} \eta_{ij}e_j.$$

Consider the bordered determinant

$$(21)\qquad \Delta_r = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \xi_{11} & \cdots & \xi_{r1} \\ a_{21} & a_{22} & \cdots & a_{2n} & \xi_{12} & \cdots & \xi_{r2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & \xi_{1n} & \cdots & \xi_{rn} \\ \eta_{11} & \eta_{12} & \cdots & \eta_{1n} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \eta_{r1} & \eta_{r2} & \cdots & \eta_{rn} & 0 & \cdots & 0 \end{vmatrix} = \begin{vmatrix} A & X \\ Y' & 0_r \end{vmatrix}$$

where $r < n$, and 0_r is a square block of 0's with r rows and columns.

If we introduce r additional fundamental units $e_{n+1}, \cdots, e_{n+r}$, Δ_r can be regarded as the determinant of a matrix $\mathfrak{A}$ of order $n + r$, namely,

$$\mathfrak{A} = \sum_1^n a_i S e_i + \sum_1^r x_i S e_{n+i} + \sum_1^r e_{n+i} S y_i = \sum_1^{n+2r} c_i S d_i.$$

If now we form $| \mathfrak{A} | = S \mid e \mid C_{n+r}(\mathfrak{A}) \mid e \mid$ as in §5.03, we have

$$C_{n+r}(\mathfrak{A}) = \sum_{(i)}{}^{*} \mid c_{i_1} \cdots c_{i_{n+r}} \mid S \mid d_{i_1} \cdots d_{i_{n-r}} \mid \quad (i = 1, 2, \cdots, n + 2r).$$

In this form any $\mid c_{i_1} \cdots c_{i_{n+r}} \mid$ which contains more than n out of $a_1, \cdots, a_n$, $x_1, \cdots, x_r$ is necessarily 0; also, if it does not contain all the x's, the corresponding $\mid d_{i_1}, \cdots, d_{i_{n+r}} \mid$ will contain more than n out of $e_1, \cdots, e_n$, $y_1, \cdots, y_r$ and is consequently 0. We therefore have

$$C_{n+r}(\mathfrak{A}) = \sum_{(i)}{}^{*} \mid a_{i_1} a_{i_2} \cdots a_{i_{n-r}} x_1 x_2 \cdots x_r e_{n+1} \cdots e_{n+r} \mid$$

$$\times S \mid e_{i_1} e_{i_2} \cdots e_{i_{n-r}} y_1 y_2 \cdots y_r e_{n+1} \cdots e_{n+r} \mid \quad (i = 1, 2, \cdots, n)$$

and hence, passing back to space of n dimensions,

$$| \mathfrak{A} | = \sum_i{}^{*} S \mid e \mid\mid a_{i_1} \cdots a_{i_{n-r}} x_1 \cdots x_r \mid S \mid e_{i_1} \cdots e_{i_{n-r}} y_1 \cdots y_r \mid\mid e \mid$$

$$= \Sigma^{*} S \mid x_1 \cdots x_r \mid\mid a_{i_1} \cdots a_{i_{n-r}} \mid_c S \mid e_{i_1} \cdots e_{i_{n-r}} \mid_c \mid y_1 \cdots y_r \mid$$

$$= S \mid x_1 \cdots x_r \mid C^{n-r}(A) \mid y_1 \cdots y_r \mid.$$

This relation shows why the bordered determinant is frequently used in place of the corresponding compound in dealing with the theory of forms.

5.06 **The reduction of bilinear forms.** The Lagrange method of reducing quadratic and bilinear forms to a normal form is, as we shall now see, closely connected with compounds.

If A is any matrix, not identically 0, there exist vectors x_1, y_1 such that $Sx_1Ay_1 \neq 0$; then, setting $A = A_1$ for convenience, the matrix

$$A_2 = A_1 - A_1y_1 \frac{SA_1'x_1}{Sx_1A_1y_1}$$

has its rank exactly 1 less than that of A. For, if $A_1z = 0$, then

$$A_2z = A_1z - A_1y_1 \frac{SA'x_1 \cdot z}{Sx_1A_1y_1} = A_1z - A_1y_1 \frac{Sx_1A_1z}{Sx_1A_1y_1} = 0$$

and, conversely if $A_2z = 0$, then

$$A_1z = A_1y_1 \frac{Sx_1A_1z}{Sx_1A_1y_1} = kA_1y_1,$$

say, or $A_1(z - ky_1) = 0$. The null-space of A_2 is therefore obtained from that of A_1 by adding y_1 to its basis, which increases the order of this space by 1 since $A_1y_1 \neq 0$.

If $A_2 \neq 0$, this process may be repeated, that is, there exist x_2, y_2 such that $Sx_2A_2y_2 \neq 0$ and the rank of

$$A_3 = A_2 - A_2y_2 \frac{SA_2'x_2}{Sx_2A_2y_2}$$

is 1 less than that of A_2. If the rank of A is r, we may continue this process by setting

(22) $$A_{s+1} = A_s - A_sy_s \frac{SA_s'x_s}{Sx_sA_sy_s} \qquad (s = 1, 2, \cdots, r)$$

where $Sx_sA_sy_s \neq 0$ and $A_1 = A$, $A_{r+1} = 0$; we then have

(23) $$A = \sum_{s=1}^{r} A_sy_s \frac{SA_s'x_s}{Sx_sA_sy_s} = \sum_{1}^{r} \mathfrak{A}_s$$

where $\mathfrak{A}_s = A_sy_s \dfrac{SA_s'x_s}{Sx_sA_sy_s}$ is a matrix of rank 1. Generally speaking, one may take $x_s = y_s$ and it is of some interest to determine when this is not possible. If $SxBx = 0$ for every x, we readily see that B is skew. For then $Se_iBe_i = Se_jBe_j = S(e_i + e_j)B(e_i + e_j) = 0$ and therefore

$$0 = S(e_i + e_j)B(e_i + e_j) = Se_iBe_i + Se_jBe_j + Se_iBe_j + Se_jBe_i,$$

that is, $Se_iBe_j = -Se_jBe_i$ and hence $B' = -B$. Hence we may take $x_s = y_s$ so long as $A_s \neq -A_s'$.

5.07 We shall now derive more explicit forms for the terms in (23) and show how they lead to the Sylvester-Francke theorems on compound determinants.

Let $x^1, x^2, \cdots, x^r, y^1, y^2, \cdots, y^r$ be variable vectors and set

(24) $$\begin{aligned} J &= S \mid x_sx^1x^2 \cdots x^r \mid C_{r+1}(A_s) \mid y_sy^1y^2 \cdots y^r \mid \\ &= S \mid x_sx^1x^2 \cdots x^r \mid \mid A_sy_sA_sy^1 \cdots A_sy^r \mid; \end{aligned}$$

then from (22)

$$J = S \mid x_s x^1 \cdots x^r \mid\mid A_s y_s A_{s+1} y^1 \cdots A_{s+1} y^r \mid$$
$$= \mid S x_s A_s y_s S x^1 A_{s+1} y^1 \cdots S x^r A_{s+1} y^r \mid.$$

If the x's denote rows in this determinant, the first row is

$$S x_s A_s y_s,\ S x_s A_{s+1} y^1,\ \cdots,\ S x_s A_{s+1} y^r$$

each term of which is 0 except the first, since x_s lies in the null-space of A'_{s+1}, and $S x_s A_s y_s \neq 0$. Hence

$$J = S x_s A_s y_s \mid S x^1 A_{s+1} y^1 \cdots S x^r A_{s+1} y^r \mid \tag{25}$$

and therefore from (24)

$$S \mid x_s x^1 \cdots x^r \mid C_{r+1}(A_s) \mid y_s y^1 \cdots y^r \mid \tag{26}$$
$$= S x_s A_s y_s S \mid x^1 \cdots x^r \mid C_r(A_{s+1}) \mid y^1 \cdots y^r \mid.$$

Repeated application of this relation gives

$$S \mid x_s x_{s+1} \cdots x_{s+t-1} x^1 x^2 \cdots x^r \mid C_{r+t}(A_s) \mid y_s \cdots y_{s+t-1} y^1 \cdots y^r \mid \tag{27}$$
$$= S x_s A_s y_s S x_{s+1} A_{s+1} y_{s+1} \cdots S x_{s+t-1} A_{s+t-1} y_{s+t-1} S \mid x^1 \cdots x^r \mid$$
$$\cdot\, C_r(A_{s+t}) \mid y^1 \cdots y^r \mid,$$

a particular case of which is

$$S \mid x_1 x_2 \cdots x_{s-1} x \mid C_s(A) \mid y_1 \cdots y_{s-1} y \mid \tag{27'}$$
$$= S x_1 A_1 y_1 \cdots S x_{s-1} A_{s-1} y_{s-1} S x A_s y.$$

To simplify these and similar formulae we shall now use a single letter to indicate a sequence of vectors; thus we shall set $X_{s,\, s+t-1}$ for $x_s x_{s+1} \cdots x_{s+t-1}$ and Y^r for $y^1 y^2 \cdots y^r$; also $C_{r,\, s}$ for $C_r(A_s)$. Equations (26) and (27) may then be written

$$S \mid x_s X^r \mid C_{r+1,\, s} \mid y_s Y^r \mid = S x_s A_s y_s S \mid X^r \mid C_{r,\, s+1} \mid Y^r \mid, \tag{26a}$$

$$S \mid X_{s,\, s+t-1} X^r \mid C_{r+t,\, s} \mid Y_{s,\, s+t-1} Y^r \mid = \prod_{i=s}^{s+t-1} S x_i A_i y_i S \mid X^r \mid C_{r,\, s+t} \mid Y^r \mid. \tag{27a}$$

We get a more convenient form for (26a), namely

$$S \mid X_{s,\, t} X^r \mid C_{r+t-s+1,\, s} \mid Y_{s,\, t} Y^r \mid \tag{28}$$
$$= S x_s A_s y_s S \mid X_{s+1,\, t} X^r \mid C_{r+t-s,\, s+1} \mid Y_{s+1,\, t} Y^r \mid$$

by replacing r by $r + t - s$ and then changing $x^1 x^2 \cdots x^{r+t-s}$ into $x_{s+1} \cdots x_t x^1 \cdots x^r$ along with a similar change in the y's. Putting $s = 1, 2, \cdots, t$ in succession and forming the product of corresponding sides of the equations so obtained from (28) we get after canceling the common factors, which are not identically 0 provided that $r + t$ is not greater than the rank of A,

$$(29) \qquad S \mid X_t X^r \mid C_{r+t,1} \mid Y_t Y^r \mid = \prod_1^t Sx_i A_i y_i \cdot S \mid X^r \mid C_{r,\, t+1} \mid Y^r \mid,$$

or from (27′)

$$(30) \qquad S \mid X_t X^r \mid C_{r+t} \mid Y_t Y^r \mid = S \mid X_t \mid C_t \mid Y_t \mid S \mid X^r \mid C_{r,\, t+1} \mid Y^r \mid$$

which may also be written in the form

$$(30') \qquad K \equiv \frac{S \mid X_t X^r \mid C_{r+t} \mid Y_t Y^r \mid}{S \mid X_t \mid C_t \mid Y_t \mid} = \mid Sx^i A_{t+1} y^j \mid;$$

in particular

$$(31) \qquad \frac{S \mid X_t x \mid C_{t+1}(A) \mid Y_t y \mid}{S \mid X_t \mid C_t(A) \mid Y_t \mid} = SxA_{t+1}y.$$

This gives a definition of A_{t+1} which may be used in place of (22); it shows that this matrix depends on $2t$ vector parameters. It is more convenient for some purposes to use the matrix $A^{(t)}$ defined by

$$(32) \qquad SxA^{(t)}y = S \mid X_t x \mid C_{t+1}(A) \mid Y_t y \mid .$$

From (31) we then have $Sx^i A_{t+1} y^j = Sx^i A^{(t)} y^j / S \mid X_t \mid C_t \mid Y_t \mid$ and therefore from (30′)

$$(33) \qquad K = \frac{\mid Sx^i A^{(t)} y^j \mid}{[S \mid X_t \mid C_t \mid Y_t \mid]^r} = \frac{S \mid X^r \mid C_r(A^{(t)}) \mid Y^r \mid}{[S \mid X_t \mid C_t(A) \mid Y_t \mid]^r}.$$

Hence

$$(34) \qquad S \mid X_t X^r \mid C_{r+t}(A) \mid Y_t Y^r \mid = \frac{S \mid X^r \mid C_r(A^{(t)}) \mid Y^r \mid}{[S \mid X_t \mid C_t(A) \mid Y_t \mid]^{r-1}}$$

which is readily recognized as Sylvester's theorem if the x's are replaced by fundamental units and the integral form of (33) is used.

5.08 **Invariant factors.** We shall now apply the above results in deriving the normal form of §3.02. We require first, however, the following lemma.

LEMMA 4. *If $A(\lambda)$ is a matric polynomial, there exists a constant vector y and a vector polynomial x such that $SxAy$ is the highest common factor of the coordinates of A.*

Let $y = \Sigma \eta_i e_i$ be a vector whose coordinates are variables independent of λ. Let α_1 be the H. C. F. of the coordinates of $A = \| a_{ij} \|$ and set

$$A = \alpha_1 B, \; By = \Sigma \eta_j b_{ij} e_j = \Sigma \beta_i e_i.$$

There is no value λ_1 of λ independent of the η's for which every $\beta_i = 0$; for if this were so, $\lambda - \lambda_1$ would be a factor of each b_{ij} and α_1 could not then be the H. C. F. of the a_{ij}. Hence the resultant of $\beta_1, \beta_2, \cdots, \beta_n$ as polynomials in λ is not identically 0 as a polynomial in the η's; there are therefore values of the η's for which this resultant is not 0, and for these values the β's have no factor common to all. There then exist scalar polynomials $\xi_1, \xi_2, \cdots, \xi_n$ such that $\Sigma\xi_i\beta_i = 1$ and therefore, if $x = \Sigma\xi_i e_i$, we have $SxBy = 1$ or $SxAy = \alpha_1$.

Returning now to the form of A given in §5.06, namely

$$A = \sum_1^r \frac{A_s y_s S A_s' x_s}{S x_s A_s y_s},$$

we can as above choose x_s, y_s in such a manner that $Sx_sA_sy_s = \alpha_s$ is the highest common factor of the coordinates of A_s and, when this is done, $v_s = A_sy_s/\alpha_s$, $u_s = A_s'x_s/\alpha_s$ are integral in λ. We then have

$$A = \sum_1^r \frac{A_s y_s S A_s' x_s}{\alpha_s} = \Sigma\alpha_s v_s S u_s. \tag{35}$$

Moreover $A_s y_i = 0 = A_s' x_i$ when $i < s$ and therefore in

$$S\,|\,x_1 \;\cdots\; x_r\,|\,|\,A_1y_1A_2y_2 \;\cdots\; A_ry_r\,| = |\,Sx_iA_jy_j\,| = |\,SA_j'x_iy_j\,|$$

all terms on one side of the main diagonal are 0 so that it reduces to $Sx_1A_1y_1 \cdots Sx_rA_ry_r = \alpha_1\alpha_2 \cdots \alpha_r$. Hence, dividing by $\alpha_1 \cdots \alpha_r$ and replacing A_jy_j/α_j by v_j as above, we see that $|\,x_1 \;\cdots\; x_r\,|$ and $|\,v_1 \;\cdots\; v_r\,|$ are not 0 for any value of λ, and therefore the constituent vectors in each set remain linearly independent for all values of λ. It follows in the same way that the sets $u_1, \cdots, u_r$ and $y_1, \cdots, y_r$, respectively, are also linearly independent for all values of λ, that is, these four sets are elementary sets. It follows from Theorem 5 §4.03, that we can find elementary polynomials P and Q such that

$$Pv_i = e_i = Q'u_i \quad (i = 1, 2, \cdots, r),$$

and hence

$$PAQ = P\left(\sum_1^r \alpha_s v_s S u_s\right) Q = \sum_1^r \alpha_s e_s S e_s, \tag{36}$$

which is the normal form of §3.02.

5.09 Vector products. Let $x_i = \Sigma\xi_{ij}e_j$ $(i = 1, 2, \cdots, r)$ be a set of arbitrary vectors and consider the set of all products of the form $\xi_{1i_1}\xi_{2i_2} \cdots \xi_{ri_r}$ arranged in some definite order. These products may then be regarded as the coordinates of a hypernumber[2] of order n^r which we shall call the *tensor product*[3]

[2] The term 'hypernumber' is used in place of vector, as defined in §1.01 since we now wish to use the term 'vector' in a more restricted sense.

[3] This product was called by Grassmann the general or indeterminate product.

of $x_1, x_2, \cdots, x_r$ and we shall denote it by $x_1x_2 \cdots x_r$. In particular if we take all the products $e_{i_1}e_{i_2} \cdots e_{i_r}$ $(i_1, i_2, \cdots, i_r = 1, 2, \cdots, n)$ each has all its coordinates zero except one, which has the value 1, and no two are equal. Further

$$x_1x_2 \cdots x_r = \Sigma\xi_{1i_1}\xi_{2i_2} \cdots \xi_{ri_r}e_{i_1}e_{i_2} \cdots e_{i_r}.$$

If we regard the products $e_{i_1}e_{i_2} \cdots e_{i_r}$ as the basis of the set of hypernumbers, we are naturally led to consider sums of the type

$$w = \Sigma\omega_{i_1i_2 \cdots i_r}e_{i_1}e_{i_2} \cdots e_{i_r}$$

where the ω's are scalars; and we shall call such a hypernumber a *tensor* of *grade* r. It is readily seen that the product $x_1x_2 \cdots x_r$ is distributive and homogeneous with regard to each of its factors, that is,

$$x_1(\lambda x_2 + \mu y_2)x_3 \cdots x_r = \lambda x_1x_2 \cdots x_r + \mu x_1y_2x_3 \cdots x_r.$$

The product of two tensors of grade r and s is then defined by assuming the distributive law and setting

$$(e_{i_1}e_{i_2} \cdots e_{i_r})(e_{j_1}e_{j_2} \cdots e_{j_s}) = e_{i_1} \cdots e_{i_r}e_{j_1} \cdots e_{j_s}.$$

It is easily shown that the product so defined is associative; it is however not commutative as is seen from the example

$$\begin{aligned} x_1x_2 - x_2x_1 &= \Sigma\Sigma(\xi_{1i_1}\xi_{2i_2} - \xi_{1i_2}\xi_{2i_1})e_{i_1}e_{i_2} \\ &= \sum_{(i)}{}^{*} \begin{vmatrix} \xi_{1i_1} & \xi_{1i_2} \\ \xi_{2i_1} & \xi_{2i_2} \end{vmatrix} (e_{i_1}e_{i_2} - e_{i_2}e_{i_1}). \end{aligned}$$

Here the coefficients of $e_{i_1}e_{i_2} - e_{i_2}e_{i_1}$ $(i_1 < i_2)$ are the coordinates of $| x_1x_2 |$ so that this tensor might have been defined in terms of the tensor product by setting

$$| x_1x_2 | = x_1x_2 - x_2x_1.$$

In the same way, if we form the expression[4]

$$f(x_1, x_2, \cdots, x_r) = \Sigma\,\mathrm{sgn}(i_1, i_2, \cdots, i_r)x_{i_1}x_{i_2} \cdots x_{i_r},$$

and expand it in terms of the coordinates of the x's and the fundamental units, it is readily shown that the result is

$$\sum_{(i)}{}^{*} | \xi_{1i_1}\xi_{2i_2} \cdots \xi_{ri_r} | f(e_{i_1}, e_{i_2}, \cdots, e_{i_r}).$$

[4] The determinant of a square array of vectors x_{ij} $(i, j = 1, 2, \cdots, r)$ may be defined as

$$| x_{ij} | = \Sigma\mathrm{sgn}(i_1, i_2, \cdots, i_r)x_{1i_1}x_{2i_2} \cdots x_{ri_r}.$$

In this definition the row marks are kept in normal order and the column marks permuted; a different expression is obtained if the rôles of the row and column marks are interchanged but, as these determinants seem to have little intrinsic interest, it is not worth while to develop a notation for the numerous variants of the definition given above.

Here the scalar multipliers are the same as the coordinates of $| x_1 x_2 \cdots x_r |$ and hence the definition of §1.11 may now be replaced by

$$| x_1 x_2 \cdots x_r | = \Sigma \mathrm{sgn}(i_1, i_2, \cdots, i_r) x_{i_1} x_{i_2} \cdots x_{i_r},$$

which justifies the notation used. We then have

$$(37) \qquad | x_1 x_2 \cdots x_r | = \sum_{(i)}^{*} | \xi_{1i_1} \xi_{2i_2} \cdots \xi_{ri_r} | \, | e_{i_1} e_{i_2} \cdots e_{i_r} |.$$

It is easily seen that the tensors $| e_{i_1} e_{i_2} \cdots e_{i_r} |$ are linearly independent and (37) therefore shows that they form a basis for the set of vectors of grade r. Any expression of the form

$$\Sigma \xi_{i_1 i_2 \cdots i_r} | e_{i_1} e_{i_2} \cdots e_{i_r} |$$

is called a vector of grade r and a vector of the form (37) is called a pure vector of grade r.

5.10 The direct product. If $A_i = \| a_{pq}^{(i)} \|$ $(i = 1, 2, \cdots, r)$ is a sequence of matrices of order n, then

$$(38) \qquad A_1 x_1 A_2 x_2 \cdots A_r x_r = \sum_{i,j} a_{i_1 j_1}^{(1)} a_{i_2 j_2}^{(2)} \cdots a_{i_r j_r}^{(r)} \xi_{1 j_1} \xi_{2 j_2} \cdots \xi_{r j_r} e_{i_1} e_{i_2} \cdots e_{i_r}$$
$$= \mathfrak{A}(x_1 x_2 \cdots x_r)$$

where $\mathfrak{A}$ is a linear homogeneous tensor function of $x_1 x_2 \cdots x_r$, that is, a matrix in space of n^r dimensions. This matrix is called the direct product[5] of $A_1, A_2, \cdots, A_r$ and is denoted by $A_1 \times A_2 \times \cdots \times A_r$. Obviously

$$(39) \qquad A_1 B_1 \times A_2 B_2 \times \cdots = (A_1 \times A_2 \times \cdots)(B_1 \times B_2 \times \cdots),$$

and the form of (38) shows that

$$(40) \qquad (A_1 \times A_2 \times \cdots)' = A_1' \times A_2' \times \cdots.$$

From (39) we have, on putting $r = 1$ for convenience,

$$A_1 \times A_2 \times A_3 = (A_1 \times 1 \times 1)(1 \times A_2 \times 1)(1 \times 1 \times A_3).$$

Making $A_i = 1$ $(i = 2, 3, \cdots, r)$ in (38) we have

$$A_1 x_1 x_2 \cdots, x_r = \Sigma a_{i_1 j_1}^{(1)} \xi_{1 j_1} \xi_{2 i_2} \cdots \xi_{r i_r} e_{i_1} e_{i_2} \cdots e_{i_r}$$

and hence the determinant of the corresponding matrix equals $| A_1 |^{n^{r-1}}$. Treating the other factors in the same way we then see that

$$(41) \qquad | A_1 \times A_2 \times \cdots \times A_r | = | A_1 A_2 \cdots A_r |^{n^{r-1}}.$$

Again if as in §5.04 we take x_1 as an invariant vector of A_1, x_2 as an invariant vector of A_2, and so on, and denote the roots of A_i by λ_{ij}, we see that the roots

[5] This definition may be generalized by taking $x_1, x_2, \cdots$ as vectors in different spaces of possibly different orders. See also §7.03.

of $A_1 \times A_2 \times \cdots \times A_r$ are the various products $\lambda_{1j_1}\lambda_{2j_2} \cdots \lambda_{rj_r}$. When the roots of each matrix are distinct, this gives equation (41) and, since this is an integral relation among the coefficients of the A's, it follows that it is true in general.

An important particular case arises when each of the matrices in (38) equals the same matrix A; the resultant matrix is denoted by $\Pi_r(A)$, that is

$$\Pi_r(A) = A \times A \times \cdots \qquad (r \text{ factors}). \tag{42}$$

It is sometimes called the *product transformation*. Relations (39), (40), and (41) then become

$$\Pi_r(AB) = \Pi_r(A)\Pi_r(B),\ \Pi_r(A)' = \Pi_r(A'),\ |\,\Pi_r(A)\,| = |\,A\,|^{rn^{r-1}}. \tag{43}$$

5.11 **Induced or power matrices.** If $x_1, x_2, \cdots, x_r$ are arbitrary vectors, the symmetric expression obtained by forming their products in every possible order and adding is called a *permanent*. It is usually denoted by $\overset{+}{|}\, x_1x_2 \cdots x_r \,\overset{+}{|}$ but it will be more convenient here to denote it by $\{x_1x_2 \cdots x_r\}$; and similarly, if α_{ij} is a square array of scalars, we shall denote by $\{\alpha_{11}\alpha_{22} \cdots \alpha_{rr}\}$ the function $\Sigma\alpha_{1i_1}\alpha_{2i_2} \cdots \alpha_{ri_r}$ in which the summation stretches over every permutation of $1, 2, \cdots, r$.

If some of the x's are equal, the terms of $\{x_1x_2 \cdots x_r\}$ become equal in sets each of which has the same number of terms. If the x's fall into s groups of $i_1, i_2, \cdots, i_s$ members, respectively, the members in each group being equal to one another, then

$$\frac{\{x_1x_2 \cdots x_r\}}{i_1!\,i_2! \cdots i_s!} \qquad (\Sigma i_j = r)$$

has integral coefficients. For the present we shall denote this expression by $\{x_1x_2 \cdots x_r\}^*$, but sometimes it will be more convenient to use the more explicit notation

$$\begin{Bmatrix} x_1 & x_2 & \cdots & x_s \\ i_1 & i_2 & \cdots & i_s \end{Bmatrix}$$

in which i_1 of the x's equal x_1, i_2 equal x_2, etc.; this notation is, in fact, that already used in §2.08, for instance,

$$\{x \quad x \quad y\} = 2x^2y + 2xyx + 2yx^2$$

$$\begin{Bmatrix} x & y \\ 2 & 1 \end{Bmatrix} = x^2y + xyx + yx^2 = \tfrac{1}{2}\{xxy\}.$$

The same convention applies immediately to $\{\alpha_{11}\alpha_{22} \cdots \alpha_{rr}\}$.

In the notation just explained we have

$$\{x_1x_2 \cdots x_r\} = \Sigma''\{\xi_{1i_1}\xi_{2i_2} \cdots \xi_{ri_r}\}^*\{e_{i_1}e_{i_2} \cdots e_{i_r}\} \tag{44}$$

where the summation Σ'' extends over all combinations $i_1i_2 \cdots i_r$ of the first n integers repetition being allowed. This shows that the set of all permanents of grade r has the basis $\{e_{i_1}e_{i_2} \cdots e_{i_r}\}$ of order $(n + r - 1)!/r!(n - 1)!$. From (44) we readily derive

$$(45)\quad \{Ax_1Ax_2 \cdots Ax_r\} = \sum_{i,j}{}'' \{a_{i_1j_1}a_{i_2j_2} \cdots a_{i_rj_r}\}^*\{\xi_{1j_1} \cdots \xi_{rj_r}\}^*\{e_{i_1} \cdots e_{i_r}\}$$

which is a linear tensor form in $\{x_1x_2 \cdots x_r\}$. We may therefore set

$$(46)\qquad \{Ax_1Ax_2 \cdots Ax_r\} = P_r(A)\{x_1x_2 \cdots x_r\},$$

where $P_r(A)$ is a matrix of order $(n + r - 1)!/r!(n - 1)!$ whose coordinates are the polynomials in the coordinates of A which are given in (45); this matrix is called the rth *induced* or *power* matrix of A. As with $C_r(A)$ and $\Pi_r(A)$ it follows that

$$(47)\qquad \begin{gathered} P_r(AB) = P_r(A)P_r(B), \quad P_r(A)' = P_r(A'), \\ | P_r(A) | = | A |^{\binom{n+r-1}{r-1}}; \end{gathered}$$

also the roots of $P_r(A)$ are the various products of the form $\lambda_1^{\alpha_1}\lambda_2^{\alpha_2} \cdots \lambda_r^{\alpha_r}$ for which $\Sigma\alpha_i = r$.

5.12 **Associated matrices.** The matrices considered in the preceding sections have certain common properties; the coordinates of each are functions of the variable matrix A and, if $T(A)$ stands for any one of them, then

$$(48)\qquad T(AB) = T(A)T(B).$$

Following Schur, who first treated the general problem of determining all such matrices, we shall call any matrix with these properties an *associated* matrix. If S is any constant matrix in the same space as $T(A)$, then $T_1(A) = ST(A)S^{-1}$ is clearly also an associated matrix; associated matrices related in this manner are said to be *equivalent*.

Let the orders of A and $T(A)$ be n and m respectively and denote the corresponding identity matrices by 1_n and 1_m; then from (48)

$$(49)\qquad T^2(1_n) = T(1_n), \qquad T(1_n)T(A) = T(A) = T(A)T(1_n).$$

If s is the rank of $T(1_n)$, we can find a matrix S which transforms $T(1_n)$ into a diagonal matrix with s 1's in the main diagonal and zeros elsewhere; and we may without real loss of generality assume that $T(1_n)$ has this form to start with, and write

$$T(1_n) = \left\| \begin{matrix} 1_s & 0 \\ 0 & 0 \end{matrix} \right\|$$

The second equation of (49) then shows that $T(A)$ has the form

$$T(A) = \left\| \begin{matrix} T_s(A) & 0 \\ 0 & 0 \end{matrix} \right\|$$

and we shall therefore assume that $s = m$ so that $T(1_n) = 1_m$. It follows from this that $| T(A) | \not\equiv 0$ so that $T(A)$ is not singular for every A; we shall then say that T is non-singular.

A non-singular associated matrix $T(A)$ is reducible (cf. §3.10) if it can be expressed in the form $T(A) = T_1(A) + T_2(A)$ where, if $E_1 = T_1(1_n)$, $E_2 = T_2(1_n)$, so that $E_1 + E_2 = 1_m$, then

$$T_1(A) = E_1T(A)E_1, \quad T_2(A) = E_2T(A)E_2$$
$$E_1T(A)E_2 = 0 = E_2T(A)E_1$$

so that

$$E_1^2 = E_1, E_2^2 = E_2$$
$$E_1E_2 = 0 = E_2E_1$$

and there is therefore an equivalent associated matrix $t(A)$ which has the form

$$t(A) = \left\| \begin{matrix} t_1(A) & 0 \\ 0 & t_2(A) \end{matrix} \right\|$$

When $T(A)$ is reducible in this manner we have

$$\begin{aligned} T_1(AB) &= E_1T(AB)E_1 = E_1T(A)T(B)E_1 \\ &= E_1T(A)(E_1 + E_2)T(B)E_1 \\ &= E_1T(A)E_1T(B)E_1 = T_1(A)T_1(B) \end{aligned}$$

so that $T_1(A)$ and $T_2(A)$ are separately associated matrices. We may therefore assume $T(A)$ irreducible without loss of generality since reducible associated matrices may be built up out of irreducible ones by reversing the process used above.

5.13 We shall now show that, if λ is a scalar variable, then $T(\lambda)$ is a power of λ. To begin with we shall assume that the coordinates of $T(\lambda)$ are rational functions in λ and that $T(1)$ is finite; we can then set $T(\lambda) = T_1(\lambda)/f(\lambda)$ where $f(\lambda)$ is a scalar polynomial whose leading coefficient is 1 and the coordinates of $T_1(\lambda)$ are polynomials whose highest common factor has no factor in common with $f(\lambda)$. If μ is a second scalar variable, (48) then gives

$$\frac{T_1(\lambda)T_1(\mu)}{f(\lambda)f(\mu)} = \frac{T_1(\lambda\mu)}{f(\lambda\mu)},$$

hence $f(\lambda\mu)$ is a factor of $f(\lambda)f(\mu)$, from which it follows readily that $f(\lambda\mu) = f(\lambda)f(\mu)$; so that $f(\lambda)$ is a power of λ and also

(50) $$T_1(\lambda\mu) = T_1(\lambda)T_1(\mu).$$

We also have $f(1) = 1$ and hence $T_1(1_n) = T(1_n) = 1_m$.

Let $T_1(\lambda) = F_0 + \lambda F_1 + \cdots + \lambda^s F_s (F_s \neq 0)$; then from (50)

$$F_0 + \lambda\mu F_1 + \cdots + \lambda^s\mu^s F_s = (F_0 + \lambda F_1 + \cdots)(F_0 + \mu F_1 + \cdots)$$

which gives

$$F_i^2 = F_i,\ F_iF_j = 0\ (i \neq j), \qquad (i, j = 0, 1, \cdots, s).$$

Now

$$T_1(\lambda)T(A) = f(\lambda)T(\lambda)T(A) = f(\lambda)T(\lambda A) = T(A)T_1(\lambda);$$

therefore

$$\Sigma F_iT(A)\lambda^i = \Sigma T(A)F_i\lambda^i$$

and hence on comparing powers of λ we have

$$F_iT(A) = T(A)F_i$$

and, since $\Sigma F_i = T_1(1) = 1_m$ and we have assumed that $T(A)$ is irreducible, it follows that every $F_i = 0$ except F_s, which therefore equals 1_m. Hence $T_1(\lambda) = \lambda^s$ and, since $f(\lambda)$ is a power of λ, we may set

(51) $$T(\lambda) = \lambda^r.$$

Since $T(\lambda A) = T(\lambda)T(A) = \lambda^rT(A)$, we have the following theorem.

THEOREM 9. *If $T(A)$ is irreducible, and if $T(\lambda)$ is a rational function of the scalar variable λ, then $T(\lambda) = \lambda^r$ and the coordinates of $T(A)$ are homogeneous functions of order r in the coordinates of A.*

The restriction that $T(\lambda)$ is rational in λ is not wholly necessary. For instance, if q is any whole number and ϵ a corresponding primitive root of 1, then $T^q(\epsilon) = 1_m$ and from this it follows without much difficulty that $T(\epsilon) = \epsilon^s$ where s is an integer which may be taken to be the same for any *finite* number of values of q. It follows then that, if $T(\lambda) = \| t_{ij}(\lambda) \|$, the functions $t_{ij}(\lambda)$ satisfy the equation

$$t_{ij}(\epsilon\lambda) = \epsilon^s t_{ij}(\lambda)$$

and under very wide assumptions as to the nature of the functions t_{ij} it follows from this that $T(\lambda)$ has the form λ^r. Again, if we assume that $T(\lambda) = \lambda^\alpha \sum_{-\alpha}^{\alpha} T_r\lambda^r$, then $T(\lambda)T(\mu) = T(\lambda\mu)$ gives immediately

$$T_r\mu^{r+\alpha} = T(\mu)$$

so that only one value of r is admissible and for this value $T_r = 1$ as before.

5.14 If the coordinates of $T(A)$ are rational functions of the coordinates a_{ij} of A, so that r is an integer, we can set $T(A) = T_1(A)/f(A)$ where the coordinates of $T_1(A)$ are integral in the a_{ij} and $f(A)$ is a scalar polynomial in these variables which has no factor common to all the coordinates of $T_1(A)$. As in (50) we then have

$$T_1(AB) = T_1(A)T_1(B),\ f(AB) = f(A)f(B).$$

It follows from the theory of scalar invariants that $f(A)$ can be taken as a positive integral power of $| A |$; we shall therefore from this point on assume that the coordinates of $T(A)$ are homogeneous polynomials in the coordinates of A unless the contrary is stated explicitly. We shall call r the *index* of $T(A)$.

THEOREM 10. *If $T(A)$ is an associated matrix of order m and index r, and if the roots of A are $\alpha_1, \alpha_2, \cdots, \alpha_n$, then the roots of $T(A)$ have the form $\alpha_1^{r_1}\alpha_2^{r_2} \cdots \cdot \alpha_n^{r_n}$ where $\Sigma r_i = r$. The actual choice of the exponents r depends on the particular associated matrix in question but, if $\alpha_1^{r_1}\alpha_2^{r_2} \cdots \alpha_n^{r_n}$ is one root, all the distinct quantities obtained from it by permuting the α's are also roots.*

If the roots of A are arbitrary variables, then A is similar to a diagonal matrix $A_1 = \Sigma\alpha_i e_{ii}$. We can express $T(A_1)$ as a polynomial[6] in the α's, say

$$T(A_1) = \Sigma\alpha_1^{r_1}\alpha_2^{r_2} \cdots \alpha_n^{r_n} F_{r_1 r_2 \cdots r_n} \tag{52}$$

where the F's are constant matrices. If now $B = \Sigma\beta_i e_{ii}$ is a second variable diagonal matrix, the relation $T(A_1B) = T(A_1)T(B)$ gives as in (50)

$$\begin{aligned} F^2_{r_1 r_2 \cdots r_n} &= F_{r_1 r_2 \cdots r_n}, \\ F_{r_1 r_2 \cdots r_n} F_{s_1 s_2 \cdots s_n} &= 0 \ ((r_1, r_2, \cdots) \neq (s_1, s_2 \cdots)) \end{aligned} \tag{53}$$

and hence $T(A_1)$ can be expressed as a diagonal matrix with roots of the required form; these roots may of course be multiple since the rank of $F_{r_1 \cdots r_n}$ is not necessarily 1, the elementary divisors are, however, simple.

Since the associated matrices of similar matrices are similar, it follows that the roots of the characteristic equation of $T(A)$ are given by those terms in (52) for which $F_{r_1 r_2 \cdots r_n} \neq 0$; and, since this equation has coefficients which are polynomials in the coordinates of A, the roots of $T(A)$ remain in this form even when the roots of A are not necessarily all different.

The rest of the theorem follows from the fact that the trace of $T(A_1)$ equals that of $T(A)$ which is rational in the coordinates of A and is therefore symmetric in the α's.

THEOREM 11. *The value of the determinant of $T(A)$ is $| A |^{rm/n}$ and rm/n is an integer.*

For $T(A)T(\mathrm{adj}A) = T(| A |) = | A |^r$ and therefore $| T(A) |$ is a power of $| A |$, say $| A |^s$. But $T(A)$ is a matrix of order m whose coordinates are polynomials in the coordinates of A. Hence $sn = mr$ and rm/n is an integer.

5.15 Transformable systems. From a scalar point of view each of the associated matrices discussed in §§5.03–5.11 can be characterized by a set of scalar functions f_k $(k = 1, 2, \cdots, m)$ of one or more sets of variables (ξ_{ij},

[6] If we merely assume that $T(A_1)$ is a convergent series of the form (52), equation (53) still holds. It follows that there are only a finite number of terms in (52) since (53) shows that there is no linear relation among those $F_{r_1 \cdots r_n}$ which are not zero. Let F_i be the sum of those $F_{r_1 \cdots r_n}$ for which Σr_j has a fixed value ρ_i; then $T(\lambda) = \Sigma\lambda^{\rho_i}F_i$, and as before only one value of ρ_1 is admissible when $T(A)$ is irreducible.

$j = 1, 2, \cdots, n$), $i = 1, 2, \cdots, r$, which have the following property: if the ξ's are subjected to a linear transformation

$$\xi'_{ij} = \sum_{s=1}^{n} a_{js}\xi_{is} \qquad (j = 1, 2, \cdots, n; i = 1, 2, \cdots, r)$$

and if f'_k is the result of replacing ξ_{ij} by ξ'_{ij} in f_k, then

$$f'_k = \sum_{s=1}^{m} \alpha_{ks} f.$$

where the α's are functions of the a_{ij} and are independent of the ξ's. For instance, corresponding to $C_2(A)$ we have

$$f_i \equiv f_{pq} = \begin{vmatrix} \xi_{1p} & \xi_{1q} \\ \xi_{2p} & \xi_{2q} \end{vmatrix} \qquad (p, q = 1, 2, \cdots, n; p < q)$$

for which

$$\alpha_{ij} \equiv \alpha_{pq,\, rs} = \begin{vmatrix} a_{pr} & a_{qr} \\ a_{ps} & a_{qs} \end{vmatrix}.$$

We may, and will, always assume that there are no constant multipliers such that $\Sigma\lambda_i f_i = 0$. Such systems of functions were first considered by Sylvester; they are now generally called *transformable systems*.

If we put $T(A) = \| \alpha_{ij} \|$, we have immediately $T(AB) = T(A)T(B)$, and consequently there is an associated matrix corresponding to every transformable system. Conversely, there is a transformable system corresponding to an associated matrix. For if $X = \| \xi_{ij} \|$ is a variable matrix and c an arbitrary constant vector in the space of $T(A)$, then the coordinates of $T(X)c$ form a transformable system since $T(A)T(X)c = T(AX)c$ and c can be so determined that there is no constant vector b such that $SbT(X)c \equiv 0$.

The basis f_k $(k = 1, 2, \cdots, m)$ may of course be replaced by any basis which is equivalent in the sense of linear dependence, the result of such a change being to replace $T(A)$ by an equivalent associated matrix. If in particular there exists a basis

$$g_1, g_2, \cdots, g_{k_1}, h_1, h_2, \cdots, h_{k_2} \qquad (k_1 + k_2 = k)$$

such that the g's and the h's form separate transformable systems, then $T(A)$ is reducible; and conversely, if $T(A)$ is reducible, there always exists a basis of this kind.

5.16 **Transformable linear sets.** If we adopt the tensor point of view rather than the scalar one, an associated matrix is found to be connected with a linear set $\mathfrak{F}$ of constant tensors, derived from the fundamental units e_i, such that, when e_i is replaced by Ae_i $(i = 1, 2, \cdots, n)$ in the members of the basis of $\mathfrak{F}$, then the new tensors are linearly dependent on the old; in other words the set $\mathfrak{F}$ is invariant as a whole under any linear transformation A of the fundamental units. For instance, in the case of $C_2(A)$ cited above, $\mathfrak{F}$ is the linear set defined by

$$| e_i e_j | \qquad (i, j = 1, 2, \cdots, n; i < j).$$

We shall call a set which has this property a *transformable linear set.*

Let $u_1, u_2, \cdots, u_m$ be a transformable linear set of tensors of grade r and let u'_i be the tensor that results when e_j is replaced by Ae_j $(j = 1, 2, \cdots, n)$ in u_i. Since the set is transformable, we have

$$u'_i = \sum_j \alpha_{ji} u_j = T(A) u_i \qquad (i = 1, 2, \cdots, m)$$

where the α_{ij} are homogeneous polynomials in the coordinates of A of degree r. If we employ a second transformation B, we then have

$$u''_i = T(A)T(B)u_i, \qquad u''_i = T(AB)u_i \qquad (i = 1, 2, \cdots, m)$$

and therefore $T(A)$ is an associated matrix.

We have now to show that there is a transformable linear set corresponding to every associated matrix. In doing this it is convenient to extend the notation Suv to the case where u and v are tensors of grade r. Let E_i $(i = 1, 2, \cdots, n^r)$ be the unit tensors of grade r and

$$u = \Sigma \psi_i E_i, \; v = \Sigma \varphi_i E_i$$

any tensors of grade r; we then define Suv by

$$Suv = \left(\sum_1^{n^r} \psi_i \varphi_i \right) \Big/ r!$$

where the numerical divisor is introduced solely in order not to disagree with the definition of §5.02.

Let $x_i = \Sigma \xi_{ij} e_j$ $(i = 1, 2, \cdots, r)$ be a set of variable vectors and X_i $(i = 1, 2, \cdots, s)$ the set of tensors of the form $x_1^{j_1} x_2^{j_2} \cdots x_r^{j_r}$ $(\Sigma j_i = r)$; we can then put any product $\xi_{11}^{\beta_{11}} \xi_{12}^{\beta_{12}} \cdots \xi_{rn}^{\beta_{rn}}$ for which $\Sigma \beta_{ij} = r$ in the form kSE_iX_j, k being a numerical factor. This can be done in more than one way as a rule; in fact, if $\sum_j \beta_{ij} = \beta_i$, then

$$\xi_{11}^{\beta_{11}} \cdots \xi_{1n}^{\beta_{1n}} = \frac{1}{\beta_1 !} S \, e_1^{\beta_{11}} \cdots e_n^{\beta_{1n}} \, x_1^{\beta_1}$$

and from the definition of Suv it is clear that the factors in $e_1^{\beta_{11}} \cdots e_n^{\beta_{1n}}$ can be permuted in any way without altering the value of the scalar. It follows that

$$\xi_{11}^{\beta_{11}} \cdots \xi_{1n}^{\beta_{1n}} = \frac{1}{\beta_{11}! \, \beta_{12}! \cdots \beta_{1n}!} S \begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{11} & \cdots & \beta_{1n} \end{Bmatrix} x_1^{\beta_1}$$

and repeating this process we get

$$k_1\xi_{11}^{\beta_{11}}\xi_{12}^{\beta_{12}} \cdots \xi_{rn}^{\beta_{rn}} = S\begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{11} & \cdots & \beta_{1n}\end{Bmatrix} \cdots \begin{Bmatrix} \cdots & e_n \\ \cdots & \beta_{rn}\end{Bmatrix} x_1^{\beta_1}x_2^{\beta_2} \cdots x_r^{\beta_r}$$

where k_1 is a numerical factor whose value is immaterial for our present purposes.

If f is any homogeneous polynomial in the variables ξ_{ij} of degree ρ, it can be expressed uniquely in the form

$$f = \Sigma\Sigma\varphi_{\beta_{11}\cdots\beta_{rn}} S\begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{11} & \cdots & \beta_{1n}\end{Bmatrix} \cdots \begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{r1} & \cdots & \beta_{rn}\end{Bmatrix} x_1^{\beta_1} \cdots x_r^{\beta_r}$$

where the inner summation extends over the partitions of β_i into $\beta_{i1}, \beta_{i2}, \cdots, \beta_{in}$ $(i = 1, 2, \cdots, r)$ and the outer over all values of $\beta_1, \beta_2, \cdots, \beta_r$ for which $\Sigma\beta_i = \rho$. We may therefore write

$$f = \sum_1^s SF_jX_j$$

where, as above, $X_j = x_1^{\beta_1}x_2^{\beta_2} \cdots x_r^{\beta_r}$ and

$$F_j \equiv F_{\beta_1\beta_2\cdots\beta_r} = \Sigma\varphi_{\beta_{11}\cdots\beta_{rn}} \begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{11} & \cdots & \beta_{1n}\end{Bmatrix} \cdots \begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{r1} & \cdots & \beta_{rn}\end{Bmatrix}.$$

The expression of f in this form is unique. In the first place, $F_j \neq 0$ unless each $\varphi_{\beta_{11}\cdots\beta_{rn}}$ is zero, since the set of tensors of the form

$$\begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{11} & \cdots & \beta_{1n}\end{Bmatrix} \cdots \begin{Bmatrix} e_1 & \cdots & e_n \\ \beta_{r1} & \cdots & \beta_{rn}\end{Bmatrix} \quad (\Sigma\beta_{ij} = \rho)$$

are clearly linearly independent. Further, if $\Sigma SF_jX_j \equiv 0$, then each SF_jX_j is zero since each gives rise to terms of different type in the ξ_{ij}; and finally the form of F_j shows that $SF_jX_j = 0$ only if $F_j = 0$ since in

$$SF_jX_j = k_1\Sigma\varphi_{\beta_{11}\cdots\beta_{rn}}\xi_{11}^{\beta_{11}} \cdots \xi_{rn}^{\beta_{rn}}$$

each term of the summation is of different type in the ξ_{ij}.

Let (f_k) be a transformable system; we can now write uniquely

$$f_k = \sum_j SF_{kj}X_j \quad (k = 1, 2, \cdots, m) \tag{54}$$

and we may set

$$F = \sum_1^{n^r} f_iE_i = \sum_{i,j} E_iSF_{ij}X_j$$

where $f_i \equiv 0$ when $i > m$. If we transform the x's by $A = \| a_{ij} \|$ and denote $\Pi_r(A)$ temporarily by Π, then X_j becomes ΠX_j and F is transformed into F^* where

$$F^* = \sum_{i,j} E_iSF_{ij}\Pi X_j = \sum_{i,j} E_iS\,\Pi'F_{ij}\cdot X_j. \tag{55}$$

But the f's form a transformable system and hence by this transformation f_i becomes

$$f'_i = \sum_k \alpha_{ik} f_k$$

so that

$$F^* = \sum_{k,i} \alpha_{ik} f_k E_i = \sum_i E_i S \sum_k \alpha_{ik} \sum_j F_{kj} X_j. \tag{56}$$

Comparing (55) and (56) we have

$$\sum_j S \left[\sum_k \alpha_{ik} F_{kj} - \Pi' F_{ij} \right] X_j = 0 \tag{57}$$

and therefore, as was proved above, each of the terms of the summation is zero, that is,

$$\Pi' F_{ij} = \sum_k \alpha_{ik} F_{kj} \tag{58}$$

and therefore, if j is kept fixed, the linear set

$$(F_{1j}, F_{2j}, \cdots) \tag{59}$$

is transformable provided F_{1j}, F_{2j}, $\cdots$ are linearly independent.

If there is no j for which the set (59) is linearly independent we proceed as follows. Let $f_{ij} = SF_{ij}X_j$ so that

$$\begin{aligned} f_1 &= f_{11} + f_{12} + \cdots + f_{1s} \\ f_2 &= f_{21} + f_{22} + \cdots + f_{2s} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ f_m &= f_{m1} + f_{m2} + \cdots + f_{ms}. \end{aligned} \tag{60}$$

If the removal of any column of this array leaves the new f_i so defined linearly independent, they form a transformable system which defines the same associated matrix as the original system; we shall therefore suppose that the removal of any column leads to linear relations among the rows, the coefficients of these relations being constants. Remove now the first column; then by non-singular constant combinations of the rows we can make certain of them, say the first m_1, equal 0, the remainder being linearly independent. On applying the same transformation to the rows of (60), which leaves it still a transformable system, we see that we may replace (60) by an array of the form

$$\begin{aligned} f_1 &= f_{11} \\ &\cdots\cdots\cdots \\ f_{m_1} &= f_{m_1 1} \\ f_{m_1+1} &= f_{m_1+1,1} + f_{m_1+1,2} + \cdots + f_{m_1+1,s} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_m &= f_{m1} + f_{m2} + \cdots + f_{ms} \end{aligned} \tag{61}$$

where $f_{m_1+i} - f_{m_1+i,1}$ $(i = 1, 2, \cdots, m - m_1)$ are linearly independent. It follows that $f_1, \cdots, f_{m_1}$ are transformed among themselves and so form a transformable system. For these functions are transformed in the same way as $f_{11}, f_{21} \cdots, f_{m_1 1}$, and if the last $m - m_1$ rows of (61) were involved in the transformation, this would mean that $f_{11}, \cdots, f_{m_1 1}$, when transformed, would depend on $f_{m_1+1,i}$ etc., which is impossible owing to the linear independence of $f_{m_1+i} - f_{m_1+i,1}$ $(i = 1, 2, \cdots, m - m_1)$.

Corresponding to the first column of (61) we have tensors $F_{11}, F_{21}, \cdots, F_{m1}$ and we may suppose this basis so chosen that F_{i1} $(i = 1, 2, \cdots, p)$ are linearly independent and $F_{j1} = 0$ for $j > p$; and this can be done without disturbing the general form of (61). If $p = m$, we have a transformable system of the type we wish to obtain and we shall therefore assume that $p < m$. We may also suppose the basis so chosen that $S\bar{F}_{i1}F_{j1} = \delta_{ij}$ $(i, j = 1, 2, \cdots, p)$ as in Lemma 2, §1.09. It follows from what we have proved above that $F_{11}, F_{21}, \cdots, F_{m_1 1}$ is a transformable set.

Let A be a real matrix, the corresponding transformation of the F's being, as in (58),

$$F^*_{i1} = \sum_j \alpha_{ij}F_{j1} = \Pi' F_{i1}, \quad (i = 1, 2, \cdots, p); \tag{62}$$

we then have

$$\bar{F}^*_{i1} = \sum_j \bar{\alpha}_{ij}\bar{F}_{j1} = \Pi'(A)\bar{F}_{i1} \tag{63}$$

so that the $\bar{F}_{i1}$ also forms a transformable set. Since $F_{11}, \cdots, F_{m_1 1}$ form a transformable set, α_{ij} and $\bar{\alpha}_{ij}$ are 0 when $i > m_1$ and $j \leq m_1$ no matter what matrix A is. Now

$$\alpha_{ij} = S\bar{F}_{j1}F^*_{i1} = S\bar{F}_{j1}\Pi'(A)F_{i1} = S\Pi(A)\bar{F}_{j_1}F_{i_1} = S\Pi'(A')\bar{F}_{j_1}F_{i_1}$$

which equals 0 for $i \leq m_1$, $j > m_1$ since by (63) $\Pi'(A')\bar{F}_{j1}$ is derived from $\bar{F}_{j1}$ by the transformation A' on the x's and for $j \leq m_1$ is therefore linearly dependent on $\bar{F}_{j1}$ $(j = 1, 2, \cdots, m_1)$. Hence the last $m - m_1$ rows in (61) also form a transformable system, which is only possible if the system $f_1, f_2, \cdots, f_m$ is reducible. If $T(A)$ is irreducible, the corresponding transformable system is irreducible and it follows now that there also corresponds to it an irreducible transformable set of tensors.

5.17 We have now shown that to every associated matrix $T(A)$ of index r and order m there corresponds a transformable linear set of constant tensors $F_1, F_2, \cdots, F_m$ of grade r whose law of transformation is given by (62). Also since $\Pi'(A) = \Pi(A')$, we have

$$\Pi F_i = \Sigma\alpha'_{ik}F_k, \quad \Pi\bar{F}_i = \Sigma\bar{\alpha}'_{ik}\bar{F}_k \tag{64}$$

where $T(A') = \| \alpha'_{ij} \|$.

Since $F_1, F_2, \cdots, F_m$ are linearly independent, we can find a supplement to this set in the set of all tensors of grade r, say

$$G_1, G_2, \cdots, G_\mu \qquad (\mu = n^r - m)$$

such that

$$(65) \qquad S\bar{F}_i G_j = 0.$$

It is convenient also to choose bases for both sets such that

$$(65') \qquad S\bar{F}_i F_j = \delta_{ij} = S\bar{G}_i G_j.$$

Since the two sets together form a basis for the space of Π, we can set

$$\Pi' G_j = \Sigma \beta_{kj} F_k + \Sigma \gamma_{kj} G_k$$

and this gives

$$\beta_{ij} = S\bar{F}_i \Pi' G_j = S G_j \Pi \bar{F}_i$$

which is 0 from (64) and (65), hence the G's are transformed among themselves by Π'. This means, however, that Π' is reducible, and when it is expressed in terms of the basis $(F_1, \cdots, F_m, G_1, \cdots, G_\mu)$, the part corresponding to $(F_1, \cdots, F_m)$ has the form $\| \alpha_{ij} \|$ and is therefore similar to $T(A)$. Hence:

THEOREM 12. *Every irreducible associated matrix $T(A)$ of index r is equivalent to an irreducible part of $\Pi_r(A)$, and conversely.*

5.18 **Irreducible transformable sets.** If F is a member of a transformable linear set $\mathfrak{F} = (F_1, F_2, \cdots, F_m)$, the total set of tensors derived from F by all linear transformations of the fundamental units clearly form a transformable linear set which is contained in $\mathfrak{F}$, say $\mathfrak{F}_1$; and we may suppose the basis of $\mathfrak{F}$ so chosen that $\mathfrak{F}_1 = (F_1, F_2, \cdots, F_k)$ and $S\bar{F}_i F_j = \delta_{ij}$ $(i, j = 1, 2, \cdots, m)$. Let G be an element of $(F_{k+1}, \cdots, F_m)$ and G' a transform of G so that

$$G' = \sum_{i=1}^{m} \gamma_i F_i.$$

Then $S\bar{F}_i G' = \gamma_i$. But $S\bar{F}_i G' = S\bar{F}'_i G$, where F'_i is the transform of F_i obtained by the transverse of the transformation which produced G' from G so that $\bar{F}'_i$ is in $\mathfrak{F}_1$ for $i \leq k$. Hence $\gamma_i = 0$ for $i = 1, 2, \cdots, k$, that is, $(F_{k+1}, \cdots, F_m)$ is also a transformable set; and so, when the original set is irreducible, we must have $\mathfrak{F}_1 = \mathfrak{F}$. If we say that F generates $\mathfrak{F}$, this result may be stated as follows.

LEMMA 5. *An irreducible transformable linear set is generated by any one of its members.*

We may choose F so that it is homogeneous in each e_i; for if we replace, say, e_1 by λe_1, then F has the form $\Sigma \lambda^k H_k$ and by the same argument as in §5.13, any H_k which is not 0 is homogeneous in e_1 and belongs to $\mathfrak{F}$. A repetition of this argument shows that we may choose F to be homogeneous in each of the fundamental units which occur in it. If r is the grade of F, we may assume that F depends on $e_1, e_2, \cdots, e_s$, and, if $k_1, k_2, \cdots, k_s$ are the corresponding degrees of homogeneity, then $\Sigma k_i = r$ and, when convenient, we may arrange the notation so that $k_1 \geq k_2 \geq \cdots \geq k_s$.

If we now replace e_1 in F by $e_1 + \lambda e_i$ $(i > s)$, the coefficient H of λ is not 0, since $i > s$, and H becomes $k_1 F$ when e_1 is replaced by e_1; it therefore forms a generator of $\mathfrak{F}$ in which the degree of e_1 is one less than before. It follows that, when $r \leq n$, we may choose a generator which is linear and homogeneous in r units $e_1, e_2, \cdots, e_r$. It is also readily shown that such a tensor defines an irreducible transformable linear set if, and only if, it forms an irreducible set when the transformations of the units are restricted to permuting the first r e's among themselves. Further, since the choice of fundamental units is arbitrary, we may replace them by variable vectors $x_1, x_2, \cdots, x_r$. For instance, the transformable sets associated with Π_r, P_r and C_r are $x_1 x_2 \cdots x_r$, $\{x_1 x_2 \cdots x_r\}$ and $| x_1 x_2 \cdots x_r |$, respectively, and of these the first is reducible and the other two irreducible.

5.19 It is not difficult to calculate directly the irreducible transformable sets for small values of r by the aid of the results of the preceding paragraph. If we denote $x_1, x_2, \cdots$ by $1, 2, \cdots$, the following are generators for $r = 2, 3$.

	generator		order
		$r = 2$	
2.1	$\{12\}$		$n(n + 1)/2$
2.2	$\|12\|$		$n(n - 1)/2$
		$r = 3$	
3.1	$\{123\}$		$n(n + 1)(n + 2)/6$
3.2	$\|1\|23\|\|$		$n(n^2 - 1)/3$
3.3	$\|1\{23\}\|$		$n(n^2 - 1)/3$
3.4	$\|123\|$		$n(n - 1)(n - 2)/6$.

This method of determining the generators directly is tedious and the following method is preferable.[7] Any generator has the form

$$w_1 = \Sigma \omega_{i_1 i_2 \cdots i_r} x_{i_1} x_{i_2} \cdots x_{i_r}$$

and if $q_{i_1 \cdots i_r}$ denotes the substitution $\begin{pmatrix} 1, & 2, & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}$, we may write

$$\begin{aligned} w_1 &= \Sigma \omega_{i_1 i_2 \cdots i_r} q_{i_1 i_2 \cdots i_r} x_1 x_2 \cdots x_r \\ &= q_1 (x_1 x_2 \cdots x_r) \end{aligned}$$

where q_1 may be regarded (see chap. 10) as an element of the algebra S whose units are the operators q of the symmetric group on r letters. Now w_1 generates a transformable set and hence, if $w_i = q_i(x_1 \cdots x_r)$ $(i = 1, 2, \cdots)$ is a

[7] Fuller details of the actual determination of the generators will be found in Weyl: *Gruppentheorie und Quantentheorie*, 2 ed. chap. 5.

basis of the set, and Q is the set of elements $q_1, q_2, \cdots$ in S, then the set of elements $Qq = (q_1q, q_2q, \cdots)$ must be the same as the set Q, that is, in the terminology of chapter 10, Q is a semi-invariant subalgebra of S; conversely any such semi-invariant subalgebra gives rise to a transformable set and this set is irreducible if the semi-invariant subalgebra is minimal, that is, is contained in no other such subalgebra.

It follows now from the form derived for a group algebra such as S that we get all independent generators as follows. In the first place the operators of S can be divided into sets[8] S_k $(k = 1, 2, \cdots, t)$ such that (i) the product of an element of S_k into an element of S_j $(k \neq j)$ is zero; (ii) in the field of complex numbers a basis for each S_k can be chosen which gives the algebra of matrices of order n_k^2; and in an arbitrary field S is the direct product of a matric algebra and a division algebra; (iii) there exists a set of elements $u_{k1}, u_{k2}, \cdots, u_{k\nu_k}$ in S_k such that $\sum_i u_{ki}$ is the identity of S_k and $u_{ki}^2 = u_{ki} \neq 0$, $u_{ki}u_{kj} = 0$ $(i \neq j)$ and such that the set of elements $u_{ki}S_ku_{ki}$ is a division algebra, which in the case of the complex field contains only one independent element; (iv) the elements of S_k can be divided into ν_k sets $u_{ki}S_k$ $(i = 1, 2, \cdots)$ each of which is a minimal semi-invariant subalgebra of S and therefore corresponds to an irreducible transformable set.

[8] It is shown in the theory of groups that t equals the number of partitions of r.