CHAPTER VII

COMMUTATIVE MATRICES

7.01 We have already seen in §2.08 how to find all matrices commutative with a given matrix x which has no repeated roots. We shall now treat the somewhat more complicated case in which x is not so restricted. If

$$xy = yx \tag{1}$$

then $x^r y = yx^r$ so that, if $f(\lambda)$ is a scalar polynomial, then $f(x)y = yf(x)$. In particular, if f_i is a principal idempotent element of x, then $f_i y = yf_i$. Remembering that $\Sigma f_i = 1$ we may set

$$x = \Sigma f_i x = \Sigma x_i, \qquad y = \Sigma f_i y = \Sigma y_i,$$

and also, by §2.11, $x_i = \lambda_i f_i + z_i$, where z_i is nilpotent. Since $y_i x_j = 0 = x_i y_j$ $(i \neq j)$, the determination of all matrices y which satisfy (1) is reduced to finding y so that

$$y_i x_i = x_i y_i, \qquad y_i = f_i y = y f_i.$$

We can therefore simplify the notation by first assuming that x has only one principal idempotent element, 1, and one root which may be taken to be 0 without loss of generality; x is then nilpotent.

Let $e_1, e_2, \cdots, e_s$ be the partial idempotent elements of x and let their ranks be $n_1, n_2, \cdots, n_s$; x is then composed of blocks of the form

$$\begin{matrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{matrix} \qquad (n_i \text{ rows and columns})$$

provided the fundamental basis is suitably chosen. To simplify the notation further we divide the array of n^2 units e_{ij} into smaller arrays formed by separating off the first n_1 rows, then the next n_2 rows, and so on, and then making a similar division of the columns (see figure 1). And when this is done, we shall denote the units in the block in which the ith set of rows meets the jth set of columns by

$$e^{ij}_{pq} \qquad (i, j = 1, 2, \cdots, s; p = 1, 2, \cdots, n_i; q = 1, 2, \cdots, n_j).$$

It is also convenient to put $e^{ij}_{pq} = 0$ for $p > n_i$ or $q > n_j$.

	n_1	n_2	n_3	$\cdots$
n_1	11	12	13	$\cdots$
n_2	21	22	23	$\cdots$
n_3	31	32	33	$\cdots$
	.	.	.	$\cdots$

FIG. 1

The expression for x is now

$$x = \sum_{i=1}^{s} \sum_{p=1}^{n_i-1} e^{ii}_{p,p+1} = \sum_{i=1}^{s} x_i$$

and we may set

$$y = \sum_{i,j,p,q} \eta^{ij}_{pq} e^{ij}_{pq} = \sum_{i,j} y_{ij}$$

where

$$y_{ij} = e_i y e_j = \sum_{p=1}^{n_i} \sum_{q=1}^{n_j} \eta^{ij}_{pq} e^{ij}_{pq}.$$

The equation $xy = yx$ is then equivalent to

$$x_i y_{ij} = y_{ij} x_j \qquad (i, j = 1, 2, \cdots, s). \tag{2}$$

If we now suppress for the moment the superscripts i, j, which remain constant in a single equation in (2), we may replace (2) by

$$\sum_{p=1}^{n_i-1} e_{\nu,\,p+1} \sum_{l=1}^{n_i} \sum_{m=1}^{n_j} \eta_{lm} e_{lm} = \sum_{l=1}^{n_i} \sum_{m=1}^{n_j} \eta_{lm} e_{lm} \sum_{q=1}^{n_j-1} e_{q,q+1}$$

or

$$\sum_{m=1}^{n_j} \sum_{p=1}^{n_i-1} \eta_{p+1,\,m} e_{pm} = \sum_{l=1}^{n_i} \sum_{q=1}^{n_j-1} \eta_{lq} e_{l,q+1}. \tag{3}$$

Equating corresponding coefficients then gives

$$\eta_{p+1,\,q+1} = \eta_{pq}. \tag{4}$$

Since $q \geq 1$ on the right of (3), it follows that $\eta_{p+1,1} = 0$ $(p = 1, 2, \cdots, n_i - 1)$ and, since $p \leq n_i - 1$ on the left, $\eta_{n_i q} = 0$ $(q = 1, 2, \cdots, n_j - 1)$ and hence from (4)

$$\eta_{p+t,\,t} = 0 = \eta_{n_i - t,\,q-t} \tag{5}$$

where $p = 0, 1, \cdots, n_i - t$, $q = t + 1, s + 2, \cdots, n_j - 1$, $t = 0, 1, \cdots$.

From (4) we see that in y_{ij} all coordinates in an oblique line parallel to the main diagonal of the original array have the same value; from the first part of (5) those to the left of the oblique AB through the upper left hand corner

are zero, as are also those to the left of the oblique CD through the lower right hand corner; the coordinates in the other obliques are arbitrary except that, as already stated, the coordinates in the same oblique are equal by (4). This state of affairs is made clearer by figure 2 where all coordinates are 0 except those in the shaded portion.

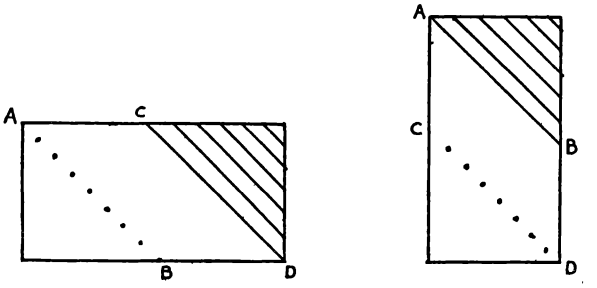


FIG. 2

As an example of this take

$$x = \begin{pmatrix}
\alpha & 1 & & & & & & & & \\
 & \alpha & & & & & & & & \\
 & & \alpha & 1 & & & & & & \\
 & & & \alpha & 1 & & & & & \\
 & & & & \alpha & & & & & \\
 & & & & & \alpha & 1 & & & \\
 & & & & & & \alpha & 1 & & \\
 & & & & & & & \alpha & 1 & \\
 & & & & & & & & \alpha & 1 \\
 & & & & & & & & & \alpha
\end{pmatrix}$$

The above rules then give for y

$$(6) \qquad \begin{pmatrix}
a_0 & a_1 & . & b_0 & b_1 & . & . & . & c_0 & c_1 \\
. & a_0 & . & . & b_0 & . & . & . & . & c_0 \\
d_0 & d_1 & e_0 & e_1 & e_2 & . & . & f_0 & f_1 & f_2 \\
. & d_0 & . & e_0 & e_1 & . & . & . & f_0 & f_1 \\
. & . & . & . & e_0 & . & . & . & . & f_0 \\
g_0 & g_1 & h_0 & h_1 & h_2 & i_0 & i_1 & i_2 & i_3 & i_4 \\
. & g_0 & . & h_0 & h_1 & . & i_0 & i_1 & i_2 & i_3 \\
. & . & . & . & h_0 & . & . & i_0 & i_1 & i_2 \\
. & . & . & . & . & . & . & . & i_0 & i_1 \\
. & . & . & . & . & . & . & . & . & i_0
\end{pmatrix}$$

where the dots represent 0.

If we arrange the notation so that $n_1 \leq n_2 \leq \cdots \leq n_s$, a simple enumeration shows that the number of independent parameters in y is

$$(2s - 1)n_1 + (2s - 3)n_2 + \cdots + n_s.$$

We have therefore the following theorem which is due to Frobenius.

THEOREM 1. *If the elementary divisors of x are $(\lambda - \lambda_i)^{n_{ij}}$, $i = 1, 2, \cdots, r$, $j = 1, 2, \cdots, s_i$, where $\lambda_1, \lambda_2, \cdots, \lambda_r$ are all different and $n_{i1} \leq n_{i2} \leq \cdots \leq n_{is_i}$, then the general form of a matrix commutative with x depends on*

$$\sum_{i=1}^{r} \sum_{j=1}^{s_i} (2s - 2j + 1)n_{ij}$$

independent parameters.

7.02 **Commutative sets of matrices.** The simple condition $xy = yx$ may be replaced by the more stringent one that y is commutative with every matrix which is commutative with x. To begin with we shall merely assume that y is commutative with each of a particular set of partial idempotent elements e_i; as in the previous section we may assume that x has only one principal idempotent element.

In order that $e_i y = y e_i$ for every i it is necessary and sufficient that $y_{ij} = 0$ when $i \neq j$; if $u_1, u_2, \cdots, u_s$ are the partial nilpotent elements of x corresponding to $e_1, e_2, \cdots, e_s$ and we set $m_i = n_i - 1$, this gives for y

$$y = \sum_i (\eta_{i0} e_i + \eta_{i1} u_i + \cdots + \eta_{im_i} u_i^{m_i}). \tag{7}$$

If we now put $z = \sum_i (\beta_i e_i + u_i)$, where no $\beta_i = 0$, and if $g(\lambda)$ is any scalar polynomial, then (cf. §2.11)

$$g(z) = \Sigma g(\beta_i e_i + u_i) = \Sigma(g(\beta_i)e_i + g'(\beta_i)u_i + \cdots + g^{(m_i)}(\beta_i)u_i^{m_i}/m_i!)$$

and when y is given, we can always find $g(\lambda)$ so that

$$\eta_{ik} = g^{(k)}(\beta_i)/k!$$

provided the β's are all different. Hence every y, including x itself, can be expressed as a polynomial in z.

We now impose the more exacting condition that y is permutable with every matrix permutable with x. Let n_{ij} $(i \neq j)$ be the matrix of the same form as y_{ij} in §7.01 but with zero coordinates everywhere except in the principal oblique; for example in (6) u_{23} is obtained by putting $f_0 = 1$ and making every other coordinate 0. We then have

$$e_i u_{ij} = u_{ij} e_j, \qquad u_i u_{ij} = u_{ij} u_j.$$

Hence $yu_{ij} = u_{ij}y$ gives $y_{ii}u_{ij} = u_{ij}y_{jj}$ and therefore from (7)

$$(\eta_{i0}e_i + \eta_{i1}u_i + \cdots + \eta_{im_i}u_i^{m_i})u_{ij} = u_{ij}(\eta_{j0}e_j + \eta_{j1}u_j + \cdots + \eta_{jm_j}u_j^{m_j})$$
$$= (\eta_{j0}e_i + \eta_{j1}u_i + \cdots + \eta_{jm_j}u_i^{m_j})u_{ij}$$

from which we readily derive for all i, j and k

$$\eta_{ik} = \eta_{jk}$$

with the understanding that η_{ik} does not actually occur when $k > m_i$. When r is the matrix used in deriving (6), these conditions give in place of (6)

$$\begin{matrix}
a_0 & a_1 & . & . & . & . & . & . & . & . \\
. & a_0 & . & . & . & . & . & . & . & . \\
. & . & a_0 & a_1 & a_2 & . & . & . & . & . \\
. & . & . & a_0 & a_1 & . & . & . & . & . \\
. & . & . & . & a_0 & . & . & . & . & . \\
. & . & . & . & . & a_0 & a_1 & a_2 & a_3 & a_4 \\
. & . & . & . & . & . & a_0 & a_1 & a_2 & a_3 \\
. & . & . & . & . & . & . & a_0 & a_1 & a_2 \\
. & . & . & . & . & . & . & . & a_0 & a_1 \\
. & . & . & . & . & . & . & . & . & a_0.
\end{matrix} \tag{8}$$

Comparing this form with (7) we see that y is now a scalar polynomial in x, which in the particular case given above becomes $g(x - \alpha)$ where

$$g(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + a_4\lambda^4.$$

The results of this section may be summarized as follows.

THEOREM 2. *Any matrix which is commutative, not only with x, but also with every matrix commutative with x, is a scalar polynomial in x.*

7.03 Rational methods. Since the solution of $xy - yx = 0$ for y can be regarded as equivalent to solving a system of linear homogeneous equations, the solution should be expressible rationally in terms of suitably chosen parameters; the method of §7.01, though elementary and direct, cannot therefore be regarded as wholly satisfactory. The following discussion, which is due to Frobenius, avoids this difficulty but is correspondingly less explicit.

As before let $xy = yx$ and set $a = \lambda - x$; also let $b = L^{-1}aM^{-1}$ be the normal form of a. If u is an arbitrary polynomial in λ and we set

$$P = L^{-1}(au + y)L, \qquad Q = M(ua + y)M^{-1},$$

then

$$Pb = PL^{-1}aM^{-1} = L^{-1}(au + y)aM^{-1} = L^{-1}a(ua + y)M^{-1} = bQ.$$

Conversely, if $Pb = bQ$ and, using the division transformation, we set

$$LPL^{-1} = av + y, \qquad M^{-1}QM = v_1a + y_1, \tag{9}$$

where y and y_1 are constants, then

$$0 = Pb - bQ = L^{-1}(av + y)aM^{-1} - L^{-1}a(v_1a + y_1)M^{-1}$$

or $a(v - v_1)a = ay_1 - ya$. Here the degree on the left is at least 2 and on the right only 1 and hence by the usual argument both sides of the equation vanish. This gives

$$ava = av_1a, \qquad ay_1 = ya$$

whence $v_1 = v$ and, since $a = \lambda - x$, also $y_1 = y$ so that $xy = yx$.

Hence we can find all matrices commutative with x by finding all solutions of

$$Pb = bQ. \tag{10}$$

Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be the invariant factors[1] of a and $n_1, n_2, \cdots, n_n$ the corresponding degrees so that b is the diagonal matrix $\Sigma\alpha_i e_{ii}$, and let $P = \| P_{ij} \|$, $Q = \| Q_{ij} \|$; then

$$P_{ij}\alpha_j = \alpha_i Q_{ij}. \tag{11}$$

By the division transformation we may set

$$P_{ij} = R_{ij}\alpha_i + p_{ij}, \qquad Q_{ij} = S_{ij}\alpha_j + q_{ij}$$

and then from (10) we have

$$R_{ij} = S_{ij}, \qquad p_{ij}\alpha_j = \alpha_i q_{ij}$$

or, if $p = \| p_{ij} \|$, $q = \| q_{ij} \|$,

$$pb = bq. \tag{12}$$

Hence $P = p$, $Q = q$ is a solution of (10) for which the degree of p_{ij} is less than that of α_i and the degree of q_{ij} is less than that of α_j. It is then evident that, when the general solution p, q of (12) is found, then the general solution of (10) has the form

$$P = bR + p, \qquad Q = Rb + q$$

where R is an arbitrary matric polynomial in λ. We are however not concerned with R; for

$$LPL^{-1} = LbRL^{-1} + LpL^{-1} = aM^{-1}RL^{-1} + LpL^{-1}$$

so that in (9) the value of y depends on p only.

[1] Since we may add a scalar to x we may clearly assume that the rank of a is n.

The general solution of (12) is given by

$$(13)\qquad \begin{aligned} p_{ij} &= \frac{\alpha_i}{\alpha_j} s_{ij}, \qquad p_{ji} = s_{ji} \\ q_{ij} &= s_{ij}, \qquad q_{ji} = \frac{\alpha_i}{\alpha_j} s_{ji} \end{aligned} \qquad (i \geqq j)$$

where $s_{\lambda\mu}$ is an arbitrary polynomial whose degree is at most $n_\mu - 1$ and which therefore depends on n_μ parameters. It follows that the total number of parameters in the value of y is that already given in §7.01.

7.04 **The direct product.** We shall consider in this section some properties of the direct product which was defined in §5.10.

THEOREM 3. *If f_{ij} $(i, j = 1, 2, \cdots, m)$ is a set of matrices, of order n, for which*

$$(14)\qquad f_{ij}f_{pq} = \delta_{jp}f_{iq}, \qquad \sum_{i=1}^{m} f_{ii} = 1,$$

then m is a factor of n and any matrix of order n can be expressed uniquely in the form $\Sigma a_{ij}f_{ij}$ where each a_{ij} is commutative with every f_{pq}; and, if $n = mr$, the rank of each f_{pq} is r.

For, if x is an arbitrary matrix and we set

$$(15)\qquad a_{ij} = \sum_{k=1}^{m} f_{ki}xf_{jk},$$

a short calculation shows:

(i) $x = \Sigma a_{ij}f_{ij}$;

(ii) $a_{ij}f_{pq} = f_{pq}a_{ij}$ for all i, j, p, q;

(iii) the set $\mathfrak{A}$ of all matrices of the form (15) is closed under the operations of addition and multiplication;

(iv) if $b_{11}, b_{12}, \cdots$ are members of $\mathfrak{A}$, then $\Sigma b_{ij}f_{ij}$ is zero if, and only if, each $b_{ij} = 0$.

If $(a_1, a_2, \cdots, a_l)$ is a basis of $\mathfrak{A}$, it follows that

$$(a_pf_{ij}:\ p = 1, 2, \cdots, l;\ i, j = 1, 2, \cdots, m)$$

is equivalent to the basis $(e_{ij}, i, j = 1, 2, \cdots, n)$ of the set of matrices of order n. This basis contains lm^2 independent elements and hence $n^2 = lm^2$ so that $n = mr$, $l = r^2$. Let r_{ij} be the rank of f_{ij}. Since $f_{ii} = f_{ij}f_{ji}$, it follows from Theorem 8 of chapter I that $r_{ii} \leq r_{ji}$; also from $f_{ji}f_{ii} = f_{ji}$ we have $r_{ji} \leq r_{ii}$; hence $r_{ii} = r_{ji}$ and therefore each r_{ji} has the same value. Finally, since $1 = \Sigma f_{ii}$, and $f_{ii}f_{jj} = 0$ $(i \neq j)$ and $r_{ii} = r_{jj}$, we have $mr_{ii} = n$ and hence each $r_{ij} = r$.

We shall now show that a basis g_{ij} can be chosen for $\mathfrak{A}$ which satisfies the relations (14) with r in place of m. Since the rank of f_{ii} is r, we can set

$$f_{ii} = \sum_1^r \alpha_{ik} S \beta_{ik} \qquad (i = 1, 2, \cdots, m) \tag{16}$$

where the sets of vectors (α_{ik}) and (β_{ik}) $(i = 1, 2, \cdots, m;\ k = 1, 2, \cdots, r)$ each form a basis of the n-space since $\sum_{i=1}^{m} f_{ii} = 1$. If (α'_{ik}), (β'_{ik}) are the corresponding reciprocal sets and

$$p_{ii} = \sum_1^r \beta'_{ik} S \alpha'_{ik} \qquad (i = 1, 2, \cdots, m)$$

we have, since $S\alpha_{ik}\alpha'_{js} = \delta_{ij}\delta_{ks}$,

$$\Sigma p_{ii} = \Sigma f_{ii} \Sigma p_{ii} = \Sigma\Sigma \alpha_{ik} S \beta_{ik} \beta'_{ij} S \alpha'_{ij} = \Sigma \alpha_{ik} S \alpha'_{ik} = 1,$$

and similarly

$$f_{ii} p_{ii} = \sum_k \alpha_{ik} S \alpha'_{ik}, \qquad f_{ii} p_{jj} = 0 \qquad (i \neq j).$$

Hence

$$f_{ii} = f_{ii} \Sigma p_{jj} = \sum_k \alpha_{ik} S \alpha'_{ik}, \tag{17}$$

that is $\beta_{ik} = \alpha'_{ik}$.

Since $f_{ij} = f_{ii} f_{ij} f_{jj}$, the left ground of f_{ij} is the same as that of f_{ii} and its right ground is the same as that of f_{jj}. Let

$$f_{1j} = \Sigma \alpha_{1k} S \gamma_{jk}.$$

The vectors γ_{jk} $(k = 1, 2, \cdots, r)$ then form a basis for the set α'_{jk} $(k = 1, 2, \cdots, r)$ and, since the basis chosen for this set in (16) is immaterial, we may suppose $\gamma_{jk} = \alpha'_{jk}$ $(j = 1, 2, \cdots, m;\ k = 1, 2, \cdots, r)$, that is,

$$f_{1j} = \sum_k \alpha_{1k} S \alpha'_{jk}.$$

Similarly we may set $f_{j1} = \Sigma \alpha_{jk} S \theta_{1k}$ and then since

$$\sum_k \alpha_{1k} S \alpha'_{1k} = f_{11} = f_{1j} f_{j1} = \sum_{k,s} \alpha_{1k} S \alpha'_{jk} \alpha_{js} S \theta_{1s} = \Sigma \alpha_{1k} S \theta_{1k},$$

we have $\theta_{1k} = \alpha'_{1k}$ and therefore

$$f_{j1} = \Sigma \alpha_{jk} S \alpha'_{1k}.$$

and finally

$$f_{ij} = f_{i1}f_{1j} = \sum_{k,s} \alpha_{ik}S\alpha'_{1k}\alpha_{1s}S\alpha'_{js} = \Sigma\alpha_{ik}S\alpha'_{jk}. \tag{18}$$

If we now set $\alpha_{ik} = Pe_{(i-1)r+k}$, then by §1.09 P is non-singular and $\alpha'_{ik} = (P')^{-1}e_{(i-1)r+k}$; hence, if

$$h_{ij} = \sum_{k=1}^{r} e_{(i-1)r+k,\,(j-1)r+k}, \tag{19}$$

we have

$$f_{ij} = Ph_{ij}P^{-1}. \tag{20}$$

Also if

$$k_{ij} = \sum_{s=0}^{m-1} e_{sr+i,\,sr+j}, \tag{21}$$

then

$$k_{ij}h_{p+1,\,q+1} = e_{pr+i,\,qr+j} = n_{p+1,\,q+1}k_{ij} \tag{22}$$

so that the set (e_{ij}) of all matrices of order n may be regarded as the direct product of the sets (h_{ij}) and (k_{ij}). Finally, since any matrix can be expressed in the form $\Sigma b_{ij}h_{ij}$, where the b_{ij} depend on the basis (k_{ij}), it follows that an arbitrary matrix can also be expressed in the form

$$P\Sigma b_{ij}h_{ij}P^{-1} = \Sigma Pb_{ij}P^{-1}f_{ij};$$

$Pb_{ij}P^{-1}$ depends on the basis $(Pk_{ij}P^{-1})$ and hence, if we set

$$g_{ij} = Pk_{ij}P^{-1} \qquad (i, j = 1, 2, \cdots, r)$$

the g's form a basis of $\mathfrak{A}$ which satisfies (14).

7.05 **Functions of commutative matrices.** Let x and y be commutative matrices whose distinct roots are $\lambda_1, \lambda_2, \cdots$ and $\mu_1, \mu_2, \cdots$ respectively and let R_i be the principal idempotent unit of x corresponding to λ_i and similarly S_j the principal idempotent unit of y corresponding to μ_j. Since R_i and S_j are scalar polynomials in x and y, they are commutative. If we set

$$T_{ij} = R_iS_j,$$

those T_{ij} which are not 0 are linearly independent; for if $\Sigma\xi_{ij}T_{ij} = 0$, then

$$0 = R_p\Sigma\xi_{ij}T_{ij}S_q = \xi_{pq}T_{pq},$$

since $R_pR_i = \delta_{pi}R_p$, $S_jS_q = \delta_{jq}S_q$, so that either $\xi_{pq} = 0$ or $T_{pq} = 0$.

From the definition of T_{ij} it follows that $T_{ij}T_{pq} = 0$ when $i \neq p$ or $j \neq q$, and $T_{ij}^2 = T_{ij}$, $\Sigma T_{ij} = 1$; hence

$$x = \sum [\lambda_i + (x - \lambda_i)]T_{ij}, \qquad y = \sum [\mu_j + (y - \mu_j)]T_{ij},$$

where $(x - \lambda_i)T_{ij}$ and $(y - \mu_j)T_{ij}$ are nilpotent. If $\psi(\lambda, \mu)$ is any scalar polynomial then

$$\psi(\lambda\mu) = \psi(\lambda_i, \mu_j) + \sum_{r,s} \psi_{r\,s}^{i\,j}(\lambda - \lambda_i)^r(\mu - \mu_j)^s$$

where $\psi_{r\,s}^{i\,j}$ are scalars, we have therefore

$$\begin{aligned}\psi(x, y) &= \sum_{i,j}\left[\psi(\lambda_i, \mu_j)T_{ij} + \sum_{r,s} \psi_{r\,s}^{i\,j}(x - \lambda_i)^r(y - \mu_j)^s T_{ij}\right] \\ &= \sum_{i,j} \psi(\lambda_i, \mu_j)T_{ij} + \sum_{i,j}\sum_{r,s} \psi_{r\,s}^{i\,j} T_{i\,j}^{r\,s}\end{aligned}$$

where

$$T_{i\,j}^{r\,s} = (x - \lambda_i)^r(y - \mu_j)^s T_{ij}$$

and r runs from 1 to $\rho_i - 1$, where ρ_i is the smallest integer for which $(x - \lambda_i)^{\rho_i} R_i = 0$, and s has a similar range with respect to y. The matrices $T_{i\,j}^{r\,s}$ are commutative and each is nilpotent; and hence any linear combination of them is also nilpotent.

Let

$$z = \Sigma\psi(\lambda_i, \mu_j)T_{ij}, \qquad w = \Sigma\Sigma\psi_{r\,s}^{i\,j} T_{i\,j}^{r\,s};$$

then w, being the sum of commutative nilpotent matrices, is nilpotent. If we take in z only terms for which $T_{ij} \neq 0$, we see immediately that the roots of z are the corresponding coefficients $\psi(\lambda_i, \mu_i)$; and the reduced characteristic function of z is found as in §2.12. We have therefore the following theorem which is due to Frobenius.

THEOREM 4. *If R_i, S_j $(i = 1, 2, \cdots; j = 1, 2, \cdots)$ are the principal idempotent units of the commutative matrices x, y and $T_{ij} = R_iS_j$; and if λ_i, μ_j are the corresponding roots of x and y, respectively; then the roots of any scalar function $\psi(x, y)$ of x and y are $\psi(\lambda_i, \mu_j)$ where i and j take only those values for which $T_{ij} \neq 0$.*

This theorem extends immediately to any number of commutative matrices.

7.06 **Sylvester's identities.** It was shown in §2.08 that, if the roots of x are all distinct, the only matrices commutative with it are scalar polynomials in x; and in doing so certain identities, due to Sylvester, were derived. We shall now consider these identities in more detail.

We have already seen that in

$$f(\lambda) = |\lambda - x| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

the coefficient a_r of λ^{n-r} is $(-1)^r$ times the sum of the principal minors of x of order r; these coefficients are therefore homogeneous polynomials of degree r in the coordinates of x. We shall now denote $(-1)^r a_r$ by $[{}^x_r]$. If x is replaced by $\lambda x + \mu y$, then $[{}^x_r]$ can be expressed as a homogeneous polynomial in λ, μ of degree r, and we shall write

$$\left[\begin{matrix}\lambda x + \mu y \\ r\end{matrix}\right] = \sum_{s=0}^{r} \left\{\begin{matrix}x & y \\ s & r-s\end{matrix}\right\} \lambda^s \mu^{r-s}. \tag{23}$$

We shall further set, as in §2.08,

$$(\lambda x + \mu y)^r = \sum_{s=0}^{r} \left\{\begin{matrix}x & y \\ s & r-s\end{matrix}\right\} \lambda^s \mu^{r-s} \tag{24}$$

where $\{{}^x_s \ {}^y_t\}$ is obtained by multiplying s x's and t y's together in every possible way and adding the terms so obtained.

In this notation the characteristic equation of $\lambda x + \mu y$ is

$$\begin{aligned} 0 &= \sum_{r=0}^{n} (-1)^r \left[\begin{matrix}\lambda x + \mu y \\ r\end{matrix}\right] (\lambda x + \mu y)^{n-r} \\ &= \sum_{r,\,s,\,t} \left[\begin{matrix}x & y \\ s & r-s\end{matrix}\right] \left\{\begin{matrix}x & y \\ t-s & n+s-r-t\end{matrix}\right\} \lambda^t \mu^{n-t}, \end{aligned} \tag{25}$$

where in the second summation $[{}^x_p \ {}^y_q]$ or $\{{}^x_p \ {}^y_q\}$ is to be replaced by 0 if either p or q is negative and $[{}^x_0 \ {}^y_0] = 1$. Since λ is an independent variable, the coefficients of its various powers in (25) are identically 0, and therefore

$$\sum_{r,\,s=0}^{n} (-1)^r \left[\begin{matrix}x & y \\ s & r-s\end{matrix}\right] \left\{\begin{matrix}x & y \\ t-s & n+s-r-t\end{matrix}\right\} = 0 \qquad (t = 0, 1, \cdots, n) \tag{26}$$

a series of identical relations connecting two arbitrary matrices.

These identities can be generalized immediately. If $x_1, x_2, \cdots, x_m$ are any matrices and $\lambda_1, \lambda_2, \cdots,$ scalar variables, we may write

$$\begin{aligned} \left[\begin{matrix}\Sigma\lambda_i x_i \\ r\end{matrix}\right] &= \sum \left[\begin{matrix}x_1 & x_2 & \cdots & x_m \\ r_1 & r_2 & \cdots & r_m\end{matrix}\right] \lambda_1^{r_1} \lambda_2^{r_2} \cdots \lambda_m^{r_m} \\ (\Sigma\lambda_i x_i)^r &= \sum \left\{\begin{matrix}x_1 & x_2 & \cdots & x_m \\ r_1 & r_2 & \cdots & r_m\end{matrix}\right\} \lambda_1^{r_1} \lambda_2^{r_2} \cdots \lambda_m^{r_m} \end{aligned} \qquad (\Sigma r_i = r) \tag{27}$$

and by the same reasoning as before we have

$$\sum_{r} \sum_{r_1, \cdots r_m} (-1)^r \left[\begin{matrix}x_1 & x_2 & \cdots & x_m \\ r_1 & r_2 & \cdots & r_m\end{matrix}\right] \left\{\begin{matrix}x_1 & x_2 & \cdots & x_m \\ s_1 - r_1 & s_2 - r_2 & \cdots & s_m - r_m'\end{matrix}\right\} = 0 \tag{28}$$

$$\left(\sum_1^m r_i = r\right)$$

where $s_1, s_2, \cdots, s_m$ is any partition of n, zero parts included, and as before a bracket symbol is 0 when any exponent is negative.

Since $\begin{bmatrix} \Sigma\lambda_i x_i \\ r \end{bmatrix}$ is the sum of the principal minors of $\Sigma\lambda_i x_i$ of order r, we see that $\begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ r_1 & r_2 & \cdots & r_m \end{bmatrix}$ $(\Sigma r_i = r)$ is formed as follows. Take any principal minor of x_1 of order r and the corresponding minors of $x_2, x_3, \cdots, x_m$ and replace r_2 of its columns by the corresponding columns of x_2, then replace r_3 of the remaining columns by the corresponding ones of x_3, and so on; do this in every possible way for each of the minors of order r of x_1 and add all the terms so obtained.

There is a great variety of relations connecting the scalar functions defined above, a few of which we note here for convenience.

(i)
$$\begin{bmatrix} 1 \\ r \end{bmatrix} = \frac{n!}{r!(n-r)!}, \qquad \begin{bmatrix} x & 1 \\ r & s \end{bmatrix} = \frac{(n-r)!}{s!(n-r-s)!}\begin{bmatrix} x \\ r \end{bmatrix},$$
$$\begin{bmatrix} x \\ n \end{bmatrix} = |x|, \qquad \begin{bmatrix} x & x \\ r & s \end{bmatrix} = \frac{(r+s)!}{r!s!}\begin{bmatrix} x \\ r+s \end{bmatrix}.$$

(ii) The value of $\begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ & & r & \end{bmatrix}$ is unchanged by a cyclic permutation of the x's.

(iii)
$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1 & r_2 & \cdots & r_m \end{bmatrix} = \frac{n!}{\Pi(r_i!)\,(n-\Sigma r_i)!},$$
$$\begin{bmatrix} x_1 & x_2 & \cdots & x_m & 1 \\ r_1 & r_2 & \cdots & r_m & s \end{bmatrix} = \frac{(n-\Sigma r_i)!}{s!(n-s-\Sigma r_i)!}\begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ r_1 & r_2 & \cdots & r_m \end{bmatrix}$$
$$\begin{bmatrix} x & x & \cdots & x & y_1 & \cdots & y_p \\ r_1 & r_2 & \cdots & r_m & s_1 & \cdots & s_p \end{bmatrix} = \frac{(\Sigma r_i)!}{\Pi(r_i!)}\begin{bmatrix} x & y_1 & \cdots & y_p \\ \Sigma r_i & s_1 & \cdots & s_p \end{bmatrix}$$

(iv)
$$\begin{bmatrix} x \\ r \end{bmatrix}\begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \Sigma\begin{bmatrix} xy_{i_1} & xy_{i_2} & \cdots & xy_{i_r} & y_{i_{r+1}} & \cdots & y_{i_n} \\ 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{bmatrix}$$

where the summation extends over the $n!/r!(n-r)!$ ways of choosing r integers out of $1, 2, \cdots, n$, the order being immaterial.

7.07 **Similar matrices.** In addition to the identities discussed in the preceding section Sylvester gave another type, a modification of which we shall now discuss. If x, y, a are arbitrary matrices, we have

(29)
$$x^{r+1}a - ay^{r+1} = x(x^r a + x^{r-1}ay + x^{r-2}ay^2 + \cdots + ay^r) - (x^r a + x^{r-1}ay + x^{r-2}ay^2 + \cdots + ay^r)y$$

or say

$$x^{r+1}a - ay^{r+1} = x(x, a, y)_r - (x, a, y)_r y$$

where

$$(30) \qquad (x, a, y)_r = \sum_{i=0}^{r} x^{r-i} a y^i.$$

Suppose now that x and y satisfy the same equation $f(\lambda) = 0$ where

$$f(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots + a_{m-1}\lambda + a_m$$

x and y being commutative with each a_i and a commutative with every a_i. Let

$$(31) \qquad u = \sum_{0}^{m-1} a_i (x, a, y)_{m-i-1};$$

then

$$(32) \qquad 0 = f(x)a - af(y) = \Sigma a_i(x^{m-i}a - ay^{m-i}) = xu - uy.$$

If $|u| \neq 0$, it follows that $y = u^{-1}xu$, that is, x and y are similar.

It can be shown that a can be chosen so that $|u| \neq 0$ provided x and y have the same invariant factors and $f(\lambda)$ is the reduced characteristic function.