# CHAPTER IV

## VECTOR POLYNOMIALS. SINGULAR MATRIC POLYNOMIALS

4.01 **Vector polynomials.** If a matric polynomial in $\lambda$ is singular, the elements of its nullspace may depend on $\lambda$. We are therefore led to consider vectors whose coordinates are polynomials in a scalar variable $\lambda$; such a vector is called a *vector polynomial.* Any vector polynomial can be put in the form

$$z(\lambda) = z_0\lambda^m + z_1\lambda^{m-1} + \cdots + z_m$$

where $z_0, z_1, \cdots, z_m$ are vectors whose coordinates are independent of $\lambda$ and, if $z_0 \neq 0$, $m$ is called the degree of $z(\lambda)$. In a linear set with a basis composed of vector polynomials we are usually only concerned with those vectors that have integral coordinates when expressed in terms of the basis and, when this is so, we shall call the set an *integral set.* In a basis of an integral set the degree of an element of maximum degree will be called the degree of the basis.

In practice an integral set is often given in terms of a sequence of vectors which are not linearly independent and so do not form a basis. For the present therefore we shall say that the sequence of vector polynomials

$$z_1(\lambda), z_2(\lambda), \cdots, z_k(\lambda) \tag{1}$$

defines the integral set of all vectors of the form $\Sigma\zeta_i(\lambda)z_i(\lambda)$ where $\zeta$'s are scalar polynomials, and show later that this is really an integral set by finding for it an integral basis. The sequence (1) is said to have rank $r$ if $| z_{i_1}z_{i_2} \cdots z_{i_s} |$ vanishes identically in $\lambda$ for all choices of $s$ $z$'s when $s > r$ and is not identically 0 for some choice of the $z$'s when $s = r$.

The theory of integral sets can be expressed entirely in terms of matric polynomials, but it will make matters somewhat clearer not to do so at first. By analogy with matrices we define an elementary transformation of a sequence of vector polynomials as follows. An elementary transformation of the sequence (1) is the operation of replacing it by a sequence $z_1', z_2', \cdots, z_k'$ where:

$$\text{Type I: } z_i' = z_i + \sum_{p \neq i} \zeta_p z_p,\ z_q' = z_q,\ (q \neq i),$$

$$\text{Type II: } z_i' = z_j,\ z_j' = z_i,\ z_q' = z_q\ (q \neq i, j),$$

$$\text{Type III: } z_p' = \rho_p z_p,\ (p = 1, 2, \cdots, k)$$

where the $\zeta$'s are scalar polynomials and the $\rho$'s constants none of which is 0.

The rank of a sequence is not altered by an elementary transformation, and two sequences connected by an elementary transformation are equivalent in the sense that every vector polynomial belonging to the integral set defined by the one also belongs to the integral set defined by the other.

Two sequences which can be derived the one from the other by elementary transformations are said to be equivalent. The corresponding integral sets

may also be said to be equivalent; and if only transformations with constant coefficients are used, the equivalence is said to be strict. Equivalence may also be defined as follows. If $P$ is an elementary matrix which turns any vector of the integral linear set $(z_1, z_2, \cdots, z_k)$ into a vector of the same set, then it is easily shown that this set is equivalent to $(Pz_1, Pz_2, \cdots, Pz_k)$ and conversely; we also say that the linear set $(z_1, z_2, \cdots, z_k)$ is *invariant under* $P$ although the individual elements of the basis are not necessarily unchanged. If the restriction that $P$ leaves $(z_1, z_2, \cdots, z_k)$ invariant is not imposed, the two sets are said to be similar.

4.02 **The degree invariants.** We have seen in the previous section that the sequence in terms of which an integral set is defined may be transformed by elementary transformations without altering the integral set itself. We shall now show how we may choose a normalized basis and determine certain invariants connected with the set. Let the vectors (1), when written in full, be

$$z_i(\lambda) = \lambda^{m_i} z_{i0} + \lambda^{m_i - 1} z_{i1} + \cdots + z_{im_i}, \tag{2}$$

and suppose the notation so arranged that $m_1 \leq m_2 \leq \cdots \leq m_k$. Suppose further that the leading coefficients $z_{10}, z_{20}, \cdots, z_{s-1,0}$ are linearly independent but that

$$z_{s0} = \sum_1^{s-1} \eta_i z_{i0},$$

the $\eta$'s being constants not all 0; then $m_s \geq m_i$ $(i = 1, 2, \cdots, s-1)$ and

$$z_s' = z_s - \sum_1^{s-1} \eta_i \lambda^{m_s - m_i} z_i$$

is either 0 or has a lower degree than $z_s$, and it may replace $z_s$ in the sequence.

After a finite number of elementary transformations of this kind we arrive at a sequence equivalent to (1) which consists of a number $p$ of vector polynomials $x_1, x_2, \cdots, x_p$ in which the leading coefficients are linearly independent followed by $k - p$ zero-vectors. Now if we form $|x_1 x_2 \cdots x_p|$ using the notation of (2) with $x$'s in place of $z$'s, the term of highest degree is $\lambda^{m_1 + \cdots + m_p} \cdot |x_{10} x_{20} \cdots x_{p0}|$, which is not 0 since the leading vectors $x_{10}, x_{20}, \cdots, x_{p0}$ are linearly independent. But the rank of a sequence is not changed by elementary transformations; hence $p = r$ and we have the following theorem.

THEOREM 1. *If $z_1, z_2, \cdots, z_k$ is a sequence of vector polynomials of rank $r$, the set of vectors of the form $\sum_i \zeta_i(\lambda) z_i(\lambda)$, the $\zeta$'s being scalar polynomials, form an integral set with a basis of order $r$ which may be so chosen that the leading coefficients of its constituent vectors are linearly independent.*

When a basis of an integral set satisfies the conditions of this theorem and its elements are arranged in order of ascending degree, we shall call it a *normal* basis.

*Corollary.* If $x_1, x_2, \cdots, x_r$ is a normal basis with the degrees $m_1 \leq m_2 \leq \cdots \leq m_r$, and if $\xi_1, \xi_2, \cdots, \xi_r$ are scalar polynomials, then the degree of the vector polynomial $x = \sum_1^s \xi_i x_i$ $(\xi_s \not\equiv 0)$ is not less than $m_s$.

THEOREM 2. *If $x_1, x_2, \cdots, x_r$ is a normal basis of an integral set and $m_1 \leq m_2 \leq \cdots \leq m_r$ the corresponding degrees, and if $y_1, y_2, \cdots, y_r$ is any other basis with the degrees $n_1 \leq n_2 \leq \cdots \leq n_r$, then*

$$m_1 \leq n_1, m_2 \leq n_2, \cdots, m_r \leq n_r.$$

*Further, the exponents $m_1, m_2, \cdots, m_r$ are the same for all normal bases.*

Let $s$ be the first integer for which $n_s < m_s$ so that $n_i \leq n_s < m_s$ for $i \leq s$. Since $(x_1, x_2, \cdots, x_r)$ is a basis, we may set

$$y_i = \sum_p \xi_{ip}(\lambda) x_p(\lambda) \qquad (i = 1, 2, \cdots, s).$$

Here no value of $p$ greater than $s - 1$ is admissible since the degree $n_i$ of $y_i$ is less than $m_s$. This would mean that the rank of $y_1, y_2, \cdots, y_s$ was less than $s$, which is impossible since they form part of a basis. Hence $m_s \leq n_s$ for all values of $s$.

If both bases are normal, it follows immediately that $m_i \leq n_i$ and also $n_i \leq m_i$, whence $m_i = n_i$, that is, the set of exponents $m_1, m_2, \cdots, m_r$ is the same for all normal bases. We shall call these exponents the *degree* invariants of the integral set.

4.03 **Elementary sets.** If $z_1(\lambda), z_2(\lambda), \cdots, z_r(\lambda)$ is a basis of an integral set, but not necessarily a normal basis, the $r$-vector $|\, z_1 z_2 \cdots z_r \,|$, which we call the determinant of the basis, is not identically 0 but may vanish for certain values of $\lambda$. If it vanishes for $\lambda = \lambda_1$, then $z_1(\lambda_1), z_2(\lambda_1), \cdots, z_r(\lambda_1)$ are linearly dependent, that is, there is a relation $\Sigma \zeta_i z_i(\lambda_1) = 0$; we may assume $\zeta_1 \neq 0$ without loss of generality. It follows that $\Sigma \zeta_i z_i(\lambda)$ has a factor of the form $(\lambda - \lambda_1)^\alpha$, $\alpha \geq 1$, and hence

$$z_1'(\lambda) = \frac{\Sigma \zeta_i z_i(\lambda)}{(\lambda - \lambda_1)^\alpha}$$

is integral; and, since $\zeta_1 \neq 0$, every element of $(z_1(\lambda), z_2(\lambda), \cdots, z_r(\lambda))$ is integrably expressible in terms of $(z_1'(\lambda), z_2(\lambda), \cdots, z_r(\lambda))$. Moreover, since

$$|\, z_1' z_2 \cdots z_r \,| = \frac{\zeta_1}{(\lambda - \lambda_1)^\alpha} |\, z_1 z_2 \cdots z_r \,|,$$

the determinant of the new basis is of lower degree than that of the old and so, if we continue this process, we shall arrive after a finite number of steps at a basis $(x_1(\lambda), x_2(\lambda), \cdots, x_r(\lambda))$ whose elements are linearly independent for all values of $\lambda$. A set which has a basis of this kind will be called an *elementary* integral set; and it is readily shown that every basis of an elementary integral set has the given property, namely, that its elements are linearly independent for every value of $\lambda$. These results are summarized as follows.

THEOREM 3. *Every integral set of order $r$ is contained in an elementary set of the same order.*

We also have

THEOREM 4. *Let $x_1, x_2, \cdots, x_r$ be a basis of an elementary set. If $r < n$, there exists a complementary elementary basis $x_{r+1}, \cdots, x_n$ such that $| x_1 x_2 \cdots x_n | \not\equiv 0$ for any value of $\lambda$ and this basis can be so chosen that its degree does not exceed that of $x_1, x_2, \cdots, x_r$.*

THEOREM 5. *If $x_1, x_2, \cdots, x_r$ is a basis of an elementary set, there exists an elementary matric polynomial $X$ such that $x_i = Xe_i$ $(i = 1, 2, \cdots, r)$.*

For let $y$ be a constant vector which for some value[1] of $\lambda$ is not linearly dependent on $x_1, x_2, \cdots, x_r$ so that we do not have identically $y = \Sigma \eta_i x_i$ for any $\eta$'s which are scalar polynomials. If for some value of $\lambda$, say $\lambda_1$, we have $y = \Sigma \xi_i x_i(\lambda_1)$, the $\xi$'s being constants, then $y - \Sigma \xi_i x_i(\lambda)$ has the factor $\lambda - \lambda_1$ and, as in the proof of Theorem 3, we can modify $y$ step by step till we arrive at a vector polynomial $x_{r+1}$ such that $x_1, x_2, \cdots, x_r, x_{r+1}$ form an elementary basis. The degree at each step of this process does not exceed that of the original basis since only constant multipliers are used. This procedure may be continued till a basis of order $n$ is reached, which proves Theorem 4.

The proof of Theorem 5 is immediate; in fact, using the basis derived in the proof of Theorem 4, $X = \sum_1^n x_i S e_i$ satisfies the required conditions and $| X | = | x_1 x_2 \cdots x_n |$, which does not vanish for any value of $\lambda$.

As a converse to Theorem 5 we have that, *if $X$ is an elementary matrix, then $x_i = Xe_i$ $(i = 1, 2, \cdots, r)$ is a basis of the elementary set $(x_1, x_2, \cdots, x_r)$.*

4.04 If $z_1, z_2, \cdots, z_k$ is a sequence of vector polynomials of rank $r$, we may always assume $k \leq n$ by merely increasing the order of the fundamental space, if necessary. Setting $z_i = \Sigma \zeta_{ji} e_j$, let us consider the matric polynomial

[1] If the question of degree is not important, any vector polynomial satisfying this condition may take the place of a constant vector.

$$Z = \left\| \begin{matrix} \zeta_{11} & \cdots & \zeta_{1k} & 0 & \cdots & 0 \\ \zeta_{21} & \cdots & \zeta_{2k} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \zeta_{n1} & \cdots & \zeta_{nk} & 0 & \cdots & 0 \end{matrix} \right\| = \sum_1^k z_i S e_i.$$

The elementary transformations used in §4.02 in finding a basis of the integral set correspond when applied to $Z$ to a combination of elementary transformations, as defined as §3.01, and because these transformations involve columns only, they correspond to multiplying $Z$ on the right by an elementary polynomial $Q_1$. Similarly, if

$$Z = \Sigma e_i S y_i,\ y_i = \Sigma \zeta_{ij} e_j,$$

the process of finding a basis for $y_1, y_2, \cdots, y_n$, whose rank is $r$, corresponds to multiplying $Z$ on the left by an elementary polynomial $P_1$.

We shall now suppose that $k = r$ so that $Q_1 = 1$; then $P_1 Z$ has the form

$$Z_1 = P_1 Z = \left\| \begin{matrix} \omega_{11} & \cdots & \omega_{1r} & 0 & \cdots & 0 \\ \omega_{21} & \cdots & \omega_{2r} & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \omega_{r1} & \cdots & \omega_{rr} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{matrix} \right\|$$

We now bring $Z_1$ to the normal form of §3.02, say

$$PZ_1Q = PP_1ZQ = \left\| \begin{matrix} \zeta_1 & & & & & & \\ & \zeta_2 & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \zeta_r & & \\ & & & & & 0 & \\ & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & & 0 \end{matrix} \right\|$$

where $\zeta_1, \zeta_2, \cdots, \zeta_r$ are the invariant factors of $Z$ (or $Z_1$) and in doing so only the first $r$ rows and columns are involved so that

$$Qe_i = e_i,\ (i = r + 1, r + 2, \cdots, n). \tag{3}$$

Therefore, if $x_i = (PP_1)^{-1}e_i$, we have successively

$$PP_1ZQe_i = \zeta_i e_i,\ ZQe_i = \zeta_i(PP_1)^{-1}e_i = \zeta_i x_i,\ (i = 1, 2, \cdots, r)$$

and, if $Q = \| q_{ij} \|$,

$$Qe_i = q_{1i}e_1 + q_{2i}e_2 + \cdots + q_{ri}e_r,\ (i = 1, 2, \cdots, r)$$

and hence

$$\zeta_i x_i = ZQe_i = q_{1i}z_1 + q_{2i}z_2 + \cdots + q_{ri}z_r,\ (i = 1, 2, \cdots, r). \tag{4}$$

But from (3) and the fact that $| Q |$ is a constant different from 0, it follows that the determinant of the coefficients in (4) is also a constant different from 0, and hence these equations can be solved for the $z$'s in terms of the $x$'s giving, say

$$z_i = \Sigma b_{ji}\zeta_j x_j\ (i = 1, 2, \cdots, r)$$

where the $b$'s are scalar polynomials.

Returning now to the case $k \geq r$, we see that, since we can pass to the case $k = r$ by elementary transformations, the $\zeta$'s are still the invariant factors of $Z = \sum_1^k z_i S e_i$. They are therefore also invariants of the integral set independently of the basis chosen to represent it, and so we shall call them the *invariant factors* of the set.

We can now state the following theorem.

THEOREM 6. *If $\zeta_1, \zeta_2, \cdots, \zeta_r$ are the invariant factors of an integral set of vector polynomials, we can find a basis of the form*

$$\zeta_1 x_1, \zeta_2 x_2, \cdots, \zeta_r x_r$$

*where $x_1, x_2, \cdots, x_r$ define an elementary set.*

4.05 **Linear elementary bases.** We shall derive in this section a canonical form for a basis of an elementary linear set. If

$$z_1, z_2, \cdots, z_r, z_i = x_i\lambda + y_i, \tag{5}$$

is a basis of an elementary linear set, it is convenient, though not necessary, to associate with it the matrix

$$A\lambda - B = \sum_1^r g_i S z_i = \sum_1^r g_i S(x_i\lambda + y_i) \tag{6}$$

where $g_1, g_2, \cdots, g_r$ is a sequence of linearly independent constant vectors. When this is done, it should be noted that multiplying $A\lambda - B$ on the right by an elementary matrix $P$ corresponds to replacing (5) by the similar sequence $P'z_1, \cdots, P'z_r$. Multiplying on the left by $P$ has no immediate interpretation in terms of the sequence except when

$$Pg_i = \sum_{j=1}^r p_{ij}g_j \quad (i = 1, 2, \cdots, r)$$

in which case we can write

$$P(A\lambda - B) = \sum_{i=1}^{r} g_i S \sum_{j=1}^{r} p_{ji} z_j = \sum_{i=1}^{r} g_i S z_i'$$

and the set $(z_1', z_2', \cdots, z_r')$ is equivalent to (5); when $P$ is constant, the equivalence is strict.

Instead of restricting ourselves to the matrix (6), we shall only assume to begin with that $A\lambda - B$ is a linear matric polynomial of rank $r < n$. The nullspace $\mathfrak{N}$ of $A\lambda - B$ is then an elementary integral set, a normalized basis of which we shall take to be

$$(7) \qquad a_i(\lambda) = a_{i0}\lambda^{m_i} + a_{i1}\lambda^{m_i - 1} + \cdots + a_{im_i} \quad (i = 1, 2, \cdots, n - r).$$

From $(A\lambda - B)a_i = 0$ we have

$$Aa_{i0} = 0, Aa_{i1} = Ba_{i0}, \cdots, 0 = Ba_{im_i},$$

or, if we set $a_{it} = 0$ for $t < 0$ or $t > m_i$,

$$(8) \qquad Aa_{it} = Ba_{i,t-1} \quad (t = 0, 1, 2, \cdots, m_i + 1).$$

We shall now show that the vectors $a_{ij}$ $(i = 1, 2, \cdots, n - r; j = 0, 1, \cdots, m_i)$ are linearly independent. Assume that $a_{ij}$ are linearly independent for $(i = 1, 2, \cdots, p - 1; j = 0, 1, \cdots, m_i)$ and $(i = p; j = 0, 1, \cdots, q - 1)$ but that

$$(9) \qquad a_{pq} + \sum_{j=0}^{q-1} \alpha_{pj} a_{pj} + \sum_{i=1}^{p-1} \sum_{j=0}^{m_i} \alpha_{ij} a_{ij} = 0.$$

Let $s'$ be the greatest value of $j$ for which some $\alpha_{ij} \neq 0$ and let $s$ be the greater of $s'$ and $q$. If we set

$$(10) \qquad c_t = a_{p,q-s+t} + \sum_{j=0}^{q-1} \alpha_{pj} a_{p,j-s+t} + \sum_{i=1}^{p-1} \sum_{j=0}^{m_i} \alpha_{ij} a_{i,j-s+t},$$

then $c_{-1} = 0$, $c_s = 0$ and

$$c_{s-1} = a_{p,q-1} + \sum_{j=1}^{q-1} \alpha_{pj} a_{p,j-1} + \sum_{i=1}^{p-1} \sum_{j=1}^{m_i} \alpha_{ij} a_{i,j-1},$$

which is not 0 by hypothesis, except perhaps when $q = 0$ and every $\alpha_{ij}$ $(j \neq 0)$ in (9) is 0, which, however, is not possible since by Theorem 1 the leading coefficients $a_{i0}$ in (7) are linearly independent. Also from (8) it follows that $Ac_t = Bc_{t-1}$, and hence

$$(11) \qquad c(\lambda) = c_0\lambda^{s-1} + c_1\lambda^{s-2} + \cdots + c_{s-1}$$

is a null-vector of $A\lambda - B$ of degree less than $m_p$. But every such integral null-vector is linearly dependent on $a_1, a_2, \cdots, a_{p-1}$ with integral coefficients, say

$$c(\lambda) = \sum_1^{p-1} \gamma_i(\lambda) a_i(\lambda); \tag{12}$$

and this gives

$$c_{s-1} = c(0) = \sum_1^{p-1} \gamma_i(0) a_i(0) = \sum_1^{p-1} \gamma_i(0) a_{im_i},$$

which is impossible since $c_{s-1}$ is obtained from (9) by lowering the second subscript in each term and no such subscript greater than $m_i$ can occur in any $a_{ij}$. Hence the $a_{ij}$ are linearly independent.

In order to simplify the notation we shall now set

$$a_{ij} = Qe_j^i \ (i = 1, 2, \cdots, n - r; j = 0, 1, \cdots, m_i) \tag{13}$$

where $Q$ is a constant non-singular matrix and $e_j^i$ are fundamental units rearranged by setting, say, $e_k = e_j^i$ when $k = \sum_{q=1}^{i-1} (m_q + 1) + j + 1$; as before $e_j^i = 0$ for $j < 0$ and $j > m_i$. We shall denote the space defined by the $e_j^i$ by $\mathfrak{M}_1$ and the complementary space by $\mathfrak{M}_2$; since the bases of $\mathfrak{M}_1$ and $\mathfrak{M}_2$ can be chosen as sequences of fundamental units, they are reciprocal as well as complementary.

We return to the particular case in which $A\lambda - B$ is given by (6). Corresponding to (12) we define a new set of vectors $w$ by

$$w_i = Q'z_i \quad (i = 1, 2, \cdots, r) \tag{14}$$

and when this is done a normal basis of the nullspace $\mathfrak{N}_1$ of $(A\lambda - B)Q = \Sigma g_i S w_i$ is given by

$$b_i = e_0^i \lambda^{m_i} + e_1^i \lambda^{m_i - 1} + \cdots + e_{m_i}^i \quad (i = 1, 2, \cdots, n - r). \tag{15}$$

We have seen in §1.10 that $b_1, b_2, \cdots, b_{n-r}$ is the space reciprocal to $w_1, w_2, \cdots, w_r$. Now in $\mathfrak{M}_1$ the $\sum_1^{n-r} m_i$ vectors

$$f_j^i = e_{j-1}^i - \lambda e_j^i \quad (i = 1, 2, \cdots, n - r; j = 1, 2, \cdots, m_i) \tag{16}$$

are linearly independent; and they form the set reciprocal to (14) in $\mathfrak{M}_1$ since $Sf_j^i b_p = 0$ for all $i, j, p$ and the sum of the orders of the two sets is $\Sigma m_i + (n - r)$ which is the order of $\mathfrak{M}_1$. Hence the total set $(w_1, w_2, \cdots, w_r)$ reciprocal to $(b_1, b_2, \cdots, b_{n-r})$ is composed of (15) together with $\mathfrak{M}_2$. We shall call this form of basis a *canonical* basis of the set (13). We can now state the following theorems.

THEOREM 7. *A linear elementary set of order r has a basis of the form*

(17) $g_j^0$ $(j = 1, 2, \cdots, m)$, $g_j^i = a_{i,j-1} + \lambda a_{ij}$ $(i = 1, 2, \cdots, \nu; j = 1, 2, \cdots, m_i)$

*where the constant vectors $g_j^0$, $a_{ij}$ are linearly independent for all $j$ and $i$ and the integers $m_i$ are those degree invariants of the reciprocal set that are not* 0.

We shall call each set $g_j^i$ $(j = 1, 2, \cdots, m_i)$ a *chain* of *index* $m_i$, and define the integers $m_1, m_2, \cdots, m_\nu$ as the *Kronecker invariants* of the set. A basis of the form (17) will be called canonical.

THEOREM 8. *Two linear elementary sets are similar if, and only if, they have the same Kronecker invariants and the same order.*

It should be noted that, if $r$ is the order of the set, then

$$(18) \qquad m + \sum_1^\nu m_i = r,\ m + \sum_1^\nu (m_i + 1) \leq n,\ \nu \leq n - r.$$

If $r = n$, all the Kronecker invariants are 0 and there are no chains in the basis.

If $z_1, z_2, \cdots, z_r$ is a normal basis of an elementary linear set, the first $m$ being constant and the rest linear in $\lambda$, and $g_j^0$, $g_j^i$ is a canonical basis, the notation being that of Theorem 7, then clearly the set $g_j^0$ $(j = 1, 2, \cdots, m)$ is strictly equivalent to $(z_1, z_2, \cdots, z_m)$ and the remaining vectors have the form

$$(19) \qquad g_j^i = w_j^i + u_j^i + \lambda v_j^i$$

where $u_j^i$ and $v_j^i$ belong to $(z_1, z_2, \cdots, z_m)$ and the $w_j^i$ are constant linear combinations of $z_{m+1}, \cdots, z_r$. Since a canonical basis is also normal,

$$(20) \qquad g_j^0 \ (j = 1, 2, \cdots, m), \quad w_j^i \ (i = 1, 2, \cdots, \nu; j = 1, 2, \cdots, m_i)$$

is a normal basis strictly equivalent to $(z_1, z_2, \cdots, z_r)$. Now (19) may be written

$$(21) \qquad w_j^i = g_j^i - u_j^i - \lambda v_j^i = a_{i,j-1} - u_j^i + \lambda(a_{ij} - v_j^i) = z_j^i + b_j^i$$

where $b_j^i = v_{j-1}^i - u_j^i$ is a constant vector of the linear set $(g_1^0, g_2^0, \cdots, g_m^0)$ and

$$(22) \qquad z_j^i = a_{i,j-1} - v_{j-1}^i + \lambda(a_{ij} - v_j^i).$$

Here (22) together with the $g_j^0$ form a canonical basis which from (21) is strictly equivalent to (20) and therefore to $(z_1, z_2, \cdots, z_r)$. We therefore have the following theorem.

THEOREM 9. *Every normal basis of a linear elementary set is strictly equivalent to some canonical basis.*

4.06 **Singular linear polynomials.** Let $A\lambda + B$ be a matric polynomial of rank $r < n$. Its left and right grounds are linear integral sets of rank $r$, and by Theorems 3 and 7 we can find canonical bases in terms of which the vectors of the two grounds can be integrally expressed, say

$$(23) \qquad z_1, z_2, \cdots, z_r \quad \text{and} \quad w_1, w_2, \cdots, w_r,$$

respectively, where the first $\alpha$ $z$'s and $\beta$ $w$'s are constant and the rest linear in $\lambda$. When $A\lambda + B$ is expressed in terms of these bases, then, remembering that no second degree term can appear, we see that it has the form

$$(24)\quad \sum_{i=1}^{\alpha}\sum_{j=1}^{\beta}(h_{ij}\lambda + k_{ij})z_i S w_j + \sum_{i=1}^{r-\alpha}\sum_{j=1}^{\beta} k_{\alpha+i,j} z_{\alpha+i} S w_j + \sum_{i=1}^{\alpha}\sum_{j=1}^{r-\beta} k_{i,\beta+j} z_i S w_{\beta+j}.$$

The row vectors

$$(25')\quad p_i = \sum_{j=1}^{\beta}(h_{ij}\lambda + k_{ij})w_j + \sum_{j=1}^{r-\beta} k_{i,\beta+j} w_{\beta+j} \qquad (i = 1, 2, \cdots, \alpha)$$

$$(25'')\quad p_{\alpha+i} = \sum_{j=1}^{\beta} k_{\alpha+i,j} w_j \qquad (i = 1, 2, \cdots, r-\alpha)$$

form a set of $r$ linearly independent vectors and, since the set (25'') depends only on $\beta$ $w$'s, we must have $r - \alpha \leq \beta$. Setting $\gamma = \alpha + \beta - r$ we may replace $w_{\gamma+1}, \cdots, w_\beta$ by $p_{\alpha+1}, \cdots, p_r$ in (23) without destroying the canonical form of the basis. A similar change can be made independently in the $z$ basis by replacing $z_{\gamma+j}$ by $\sum_{1}^{\alpha} k_{i,\beta+\gamma} z_{\beta+j}$ $(j = 1, 2, \cdots, r-\beta = \alpha-\gamma)$.

When we assume that these changes have been made to begin with, we may take in place of (24)

$$(26)\quad A\lambda + B = \sum_{i=1}^{\alpha}\sum_{j=1}^{\beta}(h_{ij}\lambda + k_{ij})z_i S w_j + \sum_{i=1}^{r-\alpha} z_{\alpha+i} S w_{\gamma+i} + \sum_{j=1}^{r-\beta} z_{\gamma+j} S w_{\beta+j}.$$

Figure 1 shows schematically the effect of this change of basis. To begin with the coefficients in (24) may be arranged in a square array $AR$ of side $r$; the

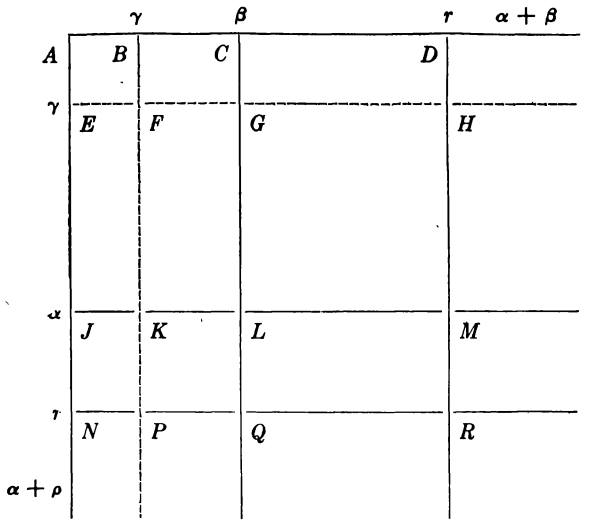



FIG. 1

first double sum corresponds to the rectangle $AL$, the second to $JQ$ and the third to $CM$, and the rectangle $LR$ contains only zeros. After the transformation which leads to (26), the only change in the scheme is that in $JQ$ the part $JP$ is now zero and the square $KQ$ has 1 in the main diagonal and zeros elsewhere, and $CM$ also takes a similar form.

If we set

$$z'_{\alpha+i} = z_{\alpha+i} + \sum_{j=1}^{\alpha} (h_{j,\gamma+i}\lambda + k_{j,\gamma+i})z_j \quad (i = 1, 2, \cdots, r-\alpha)$$

$$w'_{\beta+i} = w_{\beta+i} + \sum_{j=1}^{\gamma} (h_{\gamma+i,j}\lambda + k_{\gamma+i,j})w_j \quad (i = 1, 2, \cdots, r-\beta)$$

then $z_1, \cdots, z, z'_{\alpha+1}, \cdots, z'_r$ and $w_1, \cdots, w_\beta, w'_{\beta+1}, \cdots, w'_r$ are still elementary bases of the right and left grounds, and in terms of them (26) becomes

$$\sum_{i,j=1}^{\gamma} (h_{ij}\lambda + k_{ij})z_i S w_j + \sum_{i=1}^{r-\alpha} z'_{\alpha+i} S w_{\gamma+i} + \sum_{i=1}^{r-\beta} z_{\gamma+i} S w'_{\beta+i}.$$

The number of terms in these summations after summing for $j$ is $\gamma + (r-\alpha) + (r-\beta) = r$. Hence the rank of the square array $h_{ij}\lambda + k_{ij}$ $(i, j = 1, 2, \cdots, \gamma)$ is $\gamma$ and by a change of variable of the form $\lambda - \lambda_1 = \lambda'$, if necessary, we can secure that the array $k_{ij}$ is also of rank $\gamma$.

The transformation just employed disturbs the canonical form of the basis and we have now to devise a different transformation which will avoid this. Let us set in place of $w_1, w_2, \cdots, w_\gamma$

$$w'_j = w_j - \sum_{t=\gamma+1}^{\beta} p_{j,t-\gamma} w_t \qquad (j = 1, 2, \cdots, \gamma)$$

where the $p$'s are constants to be determined later, and for brevity set also

$$h_j = \sum_{i=1}^{\gamma} h_{ij}z_i, \; k_j = \sum_{i=1}^{\gamma} k_{ij}z_i \qquad (j = 1, 2, \cdots, \beta);$$

since the rank of $k_{ij}$ $(i = 1, 2, \cdots, \gamma)$ is $\gamma$, the vectors $k_1, k_2, \cdots, k_\gamma$ form a basis of $(z_1, z_2, \cdots, z_\gamma)$. After this change of basis the part of the first double sum (cf. (26)) which corresponds to $i = 1, 2, \cdots, \gamma; j = \gamma+1, \cdots, \beta$ is

$$\sum_{j=\gamma+1}^{\beta} \left[ h_j\lambda + k_j + \sum_{t=1}^{\gamma} p_{t,j-\gamma}(h_t\lambda + k_t) \right] S w_j. \tag{27}$$

Consider now a single chain of $z$'s of index $s$ which by a suitable change of notation we may suppose to be $z_{\alpha+1}, z_{\alpha+2}, \cdots, z_{\alpha+s}$; we shall seek to determine the $p$'s so that the corresponding part of (27) shall become

$$\sum_{j=\gamma+1}^{\gamma+s} (g_{j-1-\gamma} + \lambda g_{j-\gamma}) S w_j, \tag{28}$$

the $g$'s being vectors in the space $(z_1, z_2, \cdots, z_\gamma)$. Equating corresponding terms in (27) and (28) we have

$$g_0 = k_{\gamma+1} + \sum_t p_{t1}k_t,\ g_1 = h_{\gamma+1} + \sum_t p_{t1}h_t = k_{\gamma+2} + \sum_t p_{t2}k_t$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$g_{s-1} = h_{\gamma+s-1} + \sum_t p_{t,s-1}h_t = k_{\gamma+s} + \sum_t p_{ts}k_t,\ g_s = h_{\gamma+s} + \sum_t p_{ts}h_t.$$

Choosing $p_{t1}$ $(t = 1, 2, \cdots, \gamma)$ arbitrarily we define $g_0$ by the first equation; then the second defines $p_{t2}$ since the vector $h_{\gamma+1} + \sum_t p_{t1}h_t - k_{\gamma+2}$ can be expressed uniquely in terms of the basis $(k_1, k_2, \cdots, k_\gamma)$; and the remaining $p$'s are similarly determined in succession, while the last equation defines $g_s$.

If we now in our basis put in place of $z_{\alpha+i}$

$$z_{\alpha+i} = z_{\alpha+i} + g_{i-1} + \lambda g_i \qquad (i = 1, 2, \cdots, s)$$

and combine the corresponding part of (27) with $\sum_{i=1}^{s} z_{\alpha+i}Sw_{\gamma+i}$, the two together give $\sum_{i=1}^{s} z'_{\alpha+i}Sw_{\gamma+i}$ and the new basis is still canonical. We then treat all the $z$ chains in the same way and have finally in place of (26)

$$\sum_{i,j=1}^{\gamma}(h_{ij}\lambda + k_{ij})z_iSw'_j + \sum_{j=\gamma+1}^{\alpha}\sum_{j=1}^{\beta}(h_{ij}\lambda + k_{ij})z_iSw'_j + \sum_{i=1}^{r-\alpha} z'_{\alpha+i}Sw_{\gamma+i} + \sum_{j=1}^{r-\beta} z_{\gamma+j}Sw_{\beta+j}.$$

The changes in the bases used above have replaced the coordinates $h_{ij}$, $k_{ij}$ by 0 for the range $i = 1, 2, \cdots, \gamma$; $j = \gamma+1, \cdots, \beta$ and have left them wholly unaltered for $i = \gamma+1, \cdots, \alpha$; $j = 1, 2, \cdots, \gamma$. We can therefore interchange the rôles of the $z$'s and $w$'s and by modifying now the $w$-chains we can make these coordinates zero for the second range of subscripts without altering the zeros already obtained for the first range. Hence it is possible by a suitable choice of the original canonical bases to assume that (26) is replaced by

$$A\lambda + B = \sum_{i,j=1}^{\gamma}(h_{ij}\lambda + k_{ij})z_iSw_j + \sum_{i=\gamma+1}^{\alpha}\sum_{j=\gamma+1}^{\beta}(h_{ij}\lambda + k_{ij})z_iSw_j + \sum_{i=1}^{r-\alpha} z_{\alpha+i}Sw_{\gamma+i} + \sum_{j=1}^{r-\beta} z_{\gamma+j}Sw_{\beta+j}. \tag{29}$$

Here the first summation can be reduced to a canonical form without affecting the rest of the expression; we therefore neglect it in the meantime and deal only with the remaining terms. This is equivalent to putting $\gamma = 0$, and in making this change it is convenient to alter the notation so as to indicate the chains in the bases.

As in §4.05 let the chains of the $z$ and $w$ bases be

$$f_{j-1}^{i} + \lambda f_j^i \qquad (j = 1, 2, \cdots, s_i; i = 1, 2, \cdots, \nu_1)$$

and

$$g_{q-1}^{p} + \lambda g_q^p \qquad (q = 1, 2, \cdots, t_p; p = 1, 2, \cdots, \nu_2)$$

respectively, and denote the constant vectors of the respective bases by $z_q^p$ and $w_j^i$ where $i, j, p, q$ take the values indicated above since, when $\gamma = 0$, we have $\sum_1^{\nu_1} s_j = r - \alpha = \beta$, $\sum_1^{\nu_2} t_q = r - \beta = \alpha$. We have then to determine a canonical form for the matrix

$$(30) \quad \sum_{i,j,p,q} (h_{qj}^{pi}\lambda + k_{qj}^{pi}) z_q^p S w_j^i + \sum_{i,j} (f_{j-1}^i + \lambda f_j^i) S w_j^i + \sum_{p,q} z_q^p S(g_{q-1}^p + \lambda g_q^p),$$

and in doing so we shall show that the first summation can be eliminated by a proper choice of the bases of the chains.

It will simplify the notation if we consider first only two chains, one of index $s$ in the $z$-basis and the other of index $t$ in the $w$-basis and, omitting the superscripts, choose the notation so that these chains are $f_0 + \lambda f_1, \cdots, f_{s-1} + \lambda f_s$ and $g_0 + \lambda g_1, \cdots, g_{t-1} + \lambda g_t$. We now modify these by adding $a_{j-1} + \lambda a_j$ to $f_{j-1} + \lambda f_j$ and $b_{i-1} + \lambda b_i$ to $g_{i-1} + \lambda g_i$ choosing

$$a_j = \sum_{i=1}^{t} \alpha_{ij} z_i \qquad (j = 0, 1, \cdots, s), \qquad b_i = \sum_{j=1}^{s} \beta_{ij} w_j \qquad (i = 0, 1, \cdots, t)$$

in such a way as to eliminate the corresponding terms in the first summation of (30). To do this we must choose the $\alpha$'s and $\beta$'s so that

$$(31) \quad h_{ij} = \alpha_{ij} + \beta_{ij}, \qquad k_{ij} = \alpha_{i,j-1} + \beta_{i-1,j} \qquad (i = 1, 2, \cdots, t; j = 1, 2, \cdots, s).$$

For $j > 1$ this gives $\alpha_{i,j-1} = h_{i,j-1} - \beta_{i,j-1}$ and hence if $l_{ij} = k_{ij} - h_{i,j-1}$ we may write

$$(32') \quad k_{i1} = l_{i1} = \alpha_{i0} + \beta_{i-1,1}$$

$$(32'') \quad l_{ij} = \beta_{i-1,j} - \beta_{i,j-1} \qquad (i = 1, 2, \cdots, t; j = 2, 3, \cdots, s).$$

If we give $\alpha_{i0}$ $(i = 1, 2, \cdots, t)$ arbitrary values, (32′) defines $\beta_{i1}$ for $i = 0, 1, \cdots, t - 1$ and leaves $\beta_{t1}$ arbitrary; then $j = 2$ in (32″) gives $\beta_{i2}$ for $i = 0, 1, \cdots, t - 1$ and leaves $\beta_{t2}$ arbitrary, and so on: and when the $\beta$'s are found in this way, certain of them being arbitrary, the first equation of (31) gives the remaining $\alpha$'s.

Combining every $z$ chain in this way with each $w$ chain in turn, we finally eliminate all the terms in the quadruplê sum in (30), and (29) may therefore, by a proper choice of the two bases, be replaced by

(33)

$$A\lambda + B = \sum_{i,j=1}^{\gamma} (h_{ij}\lambda + k_{ij}) z_i S w_j + \sum_{i,j} (f_{j-1}^i + \lambda f_j^i) S w_j^i + \sum_{p,q} z_q^p S(g_{q-1}^p + \lambda g_q^p)$$

where no two of the linear sets

$$(34') \quad (z_1, z_2, \cdots, z_\gamma), \quad (f_0^i, f_1^i, \cdots, f_{s_i}^i, i = 1, 2, \cdots, \nu_1), \quad (z_1^p, z_2^p, \cdots, z_{t_p}^p, p = 1, 2, \cdots, \nu_2)$$

have any vector in common, and also no two of

$$(34'') \quad (w_1, w_2, \cdots, w_\gamma), \quad (w_1^i, w_2^i, \cdots, w_{s_i}^i, i = 1, 2, \cdots, \nu_1), \quad (g_0^p, g_1^p, \cdots, g_{t_p}^p, p = 1, 2, \cdots, \nu_2)$$

have any vector in common.

We shall now for the moment suppose that the order $n$ of the fundamental space is taken so large that we can introduce vectors $z_0^p$ $(p = 1, 2, \cdots, \nu_2)$ into the third set in (34') without causing the three spaces to overlap, and also $w_0^i$ $(i = 1, 2, \cdots, \nu_1)$ into the second set of (34''). As a matter of convenience we can then find two constant non-singular matrices $P$, $Q$ such that

$$Pz_i = e_i = Q'w_i, \qquad Pf_j^i = e_j^i = Q'w_j^i, \qquad Pz_q^p = e_q^{\nu_1 + p} = Q'g_q^p$$

where the range of the affices is as in (34) and where

$$e_j^i = e_k, \quad k = \gamma + \sum_{\alpha=1}^{i-1} (s_\alpha + 1) + j + 1, \quad e_q^{\nu_1+p} = e_k,$$

$$k = \gamma + \sum_{i=1}^{\nu_1} (s_i + 1) + \sum_{\alpha=1}^{p-1} (t_\alpha + 1) + q + 1$$

and, when this is done,

$$(35) \quad P(A\lambda + B)Q = \sum_{i,j=1}^{\gamma} (h_{ij}\lambda + k_{ij}) e_i S e_j + \sum_{i,j} (e_{j-1}^i + \lambda e_j^i) S e_j^i + \sum_{p,q} e_q^{\nu_1 + p} S(e_{q-1}^{\nu_1 + p} + \lambda e_q^{\nu_1+p}).$$

This matrix is composed of a number of blocks of terms arranged along the main diagonal, the remaining coordinates being 0. It must be carefully observed however that, owing to the introduction of the vectors $z_0^p$, $w_0^i$ and to the fact that a chain of index $s$ depends on $s + 1$ constant vectors, the total number

of rows and columns employed is greater than the rank $r$ by the total number of chains in the left and right grounds.

The first summation in (35) gives a block $|| h_{ij}\lambda + k_{ij} ||$ of $\gamma$ rows and columns. Each chain in the second and third summation gives a block of the respective forms

$$(36)\qquad \begin{array}{ccccccc} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{array} \qquad \begin{array}{ccccccc} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda. \end{array}$$

If we take $A\lambda + B\mu$ in place of $A\lambda + B$ and calculate the invariant factors these forms show that we obtain the invariant factors of $\sum_i (h_{ij}\lambda + k_{ij}\mu)e_iSe_j$ together with a number of 1's from the blocks of type (36), the number contributed by each being one less than the number of rows it contains, that is, the index of the corresponding chain. This gives the following theorem.

THEOREM 10. *Two matric polynomials $A\lambda + B\mu$ and $C\lambda + D\mu$ are strictly equivalent if, and only if, they have the same invariant factors and their respective right and left grounds have the same Kronecker invariants.*

That these conditions are necessary is obvious; that they are sufficient follows readily from the form (33) derived above. In the first place, since the Kronecker invariants are the same for both, the second and third summations in (33) have the same form for both and are therefore strictly equivalent by means of transformations which do not change the terms in the first summation. Secondly, the first summation in both yields the same invariant factors since the number of 1's due to the remaining terms depends only on the number of chains, which is the same for both; hence these summations are strictly equivalent and, because of the linear independence of the constant vectors involved, the equivalence is obtainable by transformations which do not affect the remaining terms.

When the first summation in (35) is in canonical form, we shall say that $A\lambda + B$ is in its canonical form. This is however not altogether satisfactory since the space necessary for this form may be of greater order than $n$. If $\nu$ is the greater of $\nu_1$ and $\nu_2$, (33) shows that the minimum order of the enveloping space is $\gamma + \Sigma s_i + \Sigma t_p + \nu$. A canonical form for this number of dimensions can be obtained as follows. Pair the blocks of the first and second types of (36) till all of one type are used up, taking the order of the constituents in, say, the order of (36); then in the composite block formed from such a pair discard

the first column and also the row which contains the first row of the second block. This gives a canonical form for such a pair, namely,

$$
\begin{array}{cc}
 & \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\lambda & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & \lambda & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\
0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & \lambda & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 1 & \lambda & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 1 & \lambda & \cdots & 0 & 0 \\
\cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & \lambda & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & \lambda.
\end{pmatrix} \\
(37) &
\end{array}
$$

If the number of chains in the left and right grounds is not the same, there will of course be blocks of one of the types (36) left unpaired.