SUBJECT

"ON SINGULAR PENCILS OF MATRICES"

being a thesis presented to the University of St.Andrews in application for the degree of Doctor of Philosophy by

Walter Ledermann

(a) DECLARATION

I declare that the work in my thesis is original, that the thesis has been composed by me and that it has not been accepted for any other degree.

(b) STATEMENT OF STUDY WITH DETAILS ABOUT ADMISSION

In 1933 I passed the "States Examination" in Mathematics and Physics at the University of Berlin. When I came to St.Andrews in January 1934 the Senatus aggreed that I might be enrolled as a research student on account of my German Diploma. The period of my research extended for nine terms (including two long vacations). With the permission of the Senatus I carried on my work at the University of Edinburgh during the Candlemas Term 1935. In addition to my research work I attended various Advanced Courses at Edinburgh and at St.Andrews.

(c) CERTIFICATE

I certify that Walter Ledermann has spent nine terms (including two long vacations) in research work and that the work upon which he was engaged has been completed. He has fulfilled the conditions of Ordinance No.16 and is qualified to submit the accompanying thesis in application for the Ph.D. degree, in the University of St Andrews.

H.W. Turnbull

Professor of Mathematics,

University of St Andrews

PREFACE

This thesis is a study of Singular Matrix Pencils under various aspects. In part (I) a new derivation of the Canonical Form of singular matrix pencils is gi-ven. This suggests investigation of the transformations of a pencil into itself (part (II)). Finally, part (III) deals with the canonical form of singular pencils of special types, namely those whose members are induced (or invariant) matrices.

PART I

REDUCTION OF SINGULAR PENCILS OF MATRICES

(Extracted from the *Proc.* of the *Edinburgh Mathematical Society*, Series 2 - Vol. 4 - (1934))

Reduction of Singular Pencils of Matrices[1]

By W. LEDERMANN.

(*Received* 14*th October* 1934. *Read* 2*nd November* 1934.)

§ 1. *Introduction.*

Let $\rho A + \sigma B = [\rho a_{\mu\nu} + \sigma b_{\mu\nu}]$ be a pencil of type $m \times m'$, *i.e.* with m rows and m' columns, where A and B are matrices with constant elements which are not mere scalar multiples of each other; and ρ and σ are homogeneous parameters.

The pencil $\rho A_1 + \sigma B_1$ of the same type is said to be *equivalent* to $\rho A + \sigma B$ if two non-singular constant square matrices P and Q of degree m and m' respectively can be found of such a kind as to yield an equation

$$(1) \qquad P(\rho A + \sigma B) Q = \rho A_1 + \sigma B_1; \quad |P| \neq 0, \ |Q| \neq 0.$$

Hence the totality of pencils of type $m \times m'$ may be divided up into different classes such that all members of a class are equivalent to one another, while no pencils belonging to different classes can be transformed into each other by an equation (1). The problem which now arises, viz. to carry out this classification, was first solved by Weierstrass and Kronecker in classical papers, and has since been treated by many authors.[2]

They have distinguished a certain "*canonical*" pencil in every class such that any pencil is equivalent to one of these canonical pencils.

Weierstrass dealt only with the case in which $m = m'$ and the determinant of $\rho A + \sigma B$ does not vanish identically. The general case which includes rectangular and singular pencils has been treated by Kronecker. According to Kronecker the general canonical form is

$$(2) \qquad \text{diag}\,(\Lambda_{p_1}, \Lambda_{p_2}, \ldots, \Lambda_{p_n}, \Lambda'_{q_1}, \Lambda'_{q_2}, \ldots, \Lambda'_{q_l}, M)$$

---

[1] This paper is intended as a continuation of Prof. Turnbull's paper, pages 67 to 76 above. I should like to express my special thanks to Prof. Turnbull for suggesting this investigation to me, and to thank both him and Dr Aitken for their helpful criticism.

[2] *Cf.* Turnbull and Aitken, *Canonical Matrices* (1928), p. 125 *ff*, where references may be found.

where Λ_p is a pencil of type $(p + 1) \times p$, thus

$$(3) \qquad \Lambda_1 = \begin{bmatrix} \rho \\ \sigma \end{bmatrix}, \ \Lambda_2 = \begin{bmatrix} \rho & \cdot \\ \sigma & \rho \\ \cdot & \sigma \end{bmatrix}, \ldots\ldots, \qquad \Lambda_p = \begin{bmatrix} \rho & \cdot & \cdot & \cdot & \cdot \\ \sigma & \rho & \cdot & \cdot & \cdot \\ \cdot & \sigma & \rho & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \sigma & \rho \\ \cdot & \cdot & \cdot & \cdot & \sigma \end{bmatrix}.$$

In (2) Λ'_p is the transposed matrix of Λ_p, and M is a non-singular pencil which may be reduced either to Weierstrass's classical canonical shape, the knowledge of which we shall assume, or to a rational form.[1]

Kronecker deduced the canonical form (2) under two conditions. In the first place he excluded *degenerate* pencils: *i.e.* although the pencil $\rho A + \sigma B$ is singular it must not be equivalent to a pencil $\rho A_1 + \sigma B_1$ some rows or columns of which are zero. In particular, no non-zero vector $u = [u_1, u_2, \ldots, u_m]$ can be found for which

$$uA = uB = 0.$$

For then we could construct a non-singular square matrix U of degree m whose first row is u. The pencil

$$U(\rho A + \sigma B) = \rho A_1 + \sigma B,$$

would be degenerate, its first row being zero.

It is easy to see that this assumption is not an essential restriction and we shall therefore adopt it following Kronecker.

But there is a second hypothesis which was made by Kronecker and most of the other authors[2] which from one point of view seems to be a loss of generality. They postulated that in $\rho A + \sigma B$ the rank of B should be as great as the rank of $\rho A + \sigma B$ (identically in ρ and σ).

It is always possible to fulfil this condition by introducing new variables ρ', σ' instead of ρ, σ, where

$$\begin{aligned} \rho' &= \alpha_{11}\rho + \alpha_{12}\sigma, \\ \sigma' &= \alpha_{21}\rho + \alpha_{22}\sigma, \end{aligned} \qquad \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} \neq 0.$$

This may be described as *changing the basis A, B of the pencil.* This process, however, can in general not be effected by an equivalent

[1] *Cf.* Turnbull and Aitken, *Canonical Matrices*, Chapter IX.

[2] Bromwich, however, deals with the general case (*Proc. London Math. Soc.* (1), 32 (1900)).

transformation (1) so that we lose some classes of pencils if we admit transformations of basis as well as equivalent transformations.

This applies also to the non-singular case of a square pencil $\rho A + \sigma B$ the determinant of which does not vanish identically. It has mostly been assumed that B is non-singular so that the determinant $|\rho A + \sigma B|$ has no root $\rho = 0$, $\sigma \neq 0$ or, putting $\lambda = \sigma/\rho$, that the determinant $|A + \lambda B|$ has no *infinite elementary divisors.*

In what follows we shall give a new proof for the fact that *every pencil can be reduced to the form*

$$(4) \quad \operatorname{diag}(\Lambda_{p_1}, \Lambda_{p_2}, \ldots, \Lambda_{p_n}, N_{r_1}, N_{r_2}, \ldots, N_{r_l}, \Lambda'_{q_1}, \Lambda'_{q_2}, \ldots, \Lambda'_{q_k}, M)$$

Λ_p being the same as defined in (3) and Λ'_p being its transposed. Here M is a pencil $\rho A_1 + \sigma B_1$ in which $|B_1| \neq 0$ so that the Weierstrassian method may be applied. The pencils N_r which do not occur in Kronecker's form (2) correspond to the infinite elementary divisors; thus

$$(5) \qquad N_r = \begin{bmatrix} \rho & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma & \rho & \cdot & \cdot & \cdot & \cdot \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdot & \cdot & \cdot & \cdot & \sigma & \rho \end{bmatrix} = \rho I_r + \sigma H_r$$

the determinant of N_r being ρ^r. In (5) I_r is the unit matrix of degree r and

$$(6) \qquad H_r = \begin{bmatrix} \cdot & \cdot & \cdots & \cdot \\ 1 & \cdot & \cdots & \cdot \\ \cdot & 1 & \cdots & \cdot \\ \cdots & \cdots & \cdots & \cdots \\ \cdot & \cdot & \cdots 1 & \cdot \end{bmatrix}.$$

There is no loss of generality in assuming that in $\rho A + \sigma B$ the number of rows is at least as great as the number of columns, *i.e.* $m \geqq m'$. If we had originally $m < m'$, we should consider the transposed pencil $\rho A' + \sigma B'$. We can transform this pencil into (4) and hence $\rho A + \sigma B$ into

$$\operatorname{diag}(\Lambda'_{p_1}, \ldots, \Lambda'_{p_n}, N'_{r_1}, \ldots, N'_{r_l}, \Lambda_{q_1}, \ldots, \Lambda_{q_k}, M'),$$

involving N'_r instead of N_r. But as is well known, N and N' are equivalent (they are, in fact, similar), *e.g.*

$$\begin{bmatrix} \rho & \cdot & \cdot \\ \sigma & \rho & \cdot \\ \cdot & \sigma & \rho \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & 1 \\ \cdot & 1 & \cdot \\ 1 & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \rho & \sigma & \cdot \\ \cdot & \rho & \sigma \\ \cdot & \cdot & \rho \end{bmatrix} \begin{bmatrix} \cdot & \cdot & 1 \\ \cdot & 1 & \cdot \\ 1 & \cdot & \cdot \end{bmatrix}.$$

Our proof will partly be based on the

LEMMA:

The matrix equation for Z,

$$Z = P + QZK, \tag{7}$$

where P and Q are given constant matrices admits of one and only one solution if a power of K vanishes (or if all latent roots of K are zero).

Proof:

Let $K^k = 0$. Then

$$Z_0 = \sum_{r=0}^{k-1} Q^r P K^r$$

is a solution of (7) as is easily verified. In order to prove that there is but one solution we show that the homogeneous equation

$$Y = QYK \tag{7'}$$

has only the trivial solution $Y = 0$. Let Y_0 be a solution of (7′), thus

$$Y_0 = QY_0K.$$

By iterating this equation we get

$$Y_0 = QY_0K = Q^2 Y_0 K^2 = \ldots = Q^{k-1} Y_0 K^{k-1} = Q^k Y_0 K^k = 0,$$

since $K^k = 0$.

§ 2. *Special Basis for a System of Vectors.*

Consider a system of k row-vectors of degree m:

$$z_1, z_2, \ldots, z_k. \tag{1}$$

If a row-vector z of the same type can be expressed as a linear aggregate of the vectors (1), we write:

$$z \subset (z_1, z_2, \ldots, z_k).$$

It will be convenient to introduce a matrix Z the rows of which are the vectors (1). Thus

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix} \tag{2}$$

so that Z is of type $k \times m$. The vectors (1) need not be linearly independent of one another. Let l be their rank (and the rank of Z). We may then find l basis vectors $z_{k_1}, z_{k_2}, \ldots, z_{k_l}$ out of the system (1) which are linearly independent themselves and which allow every z_k of (1) to be represented as a linear aggregate of the basis vectors. The most natural way to construct such a basis is the following: We go through the sequence (1) beginning with z_1 cancelling every vector that is linearly dependent on its predecessors. In particular every zero-vector has to be dropped. The remaining vectors may be called $z_{k_1}, z_{k_2}, \ldots, z_{k_l}$. This basis is uniquely determined by the process and may be named a "special basis." Every z_{k_λ} is a member of the sequence (1) and we have

$$k_1 < k_2 < \ldots < k_l.$$

We put

$$\overline{Z} = \begin{bmatrix} z_{k_1} \\ z_{k_2} \\ \vdots \\ z_{k_l} \end{bmatrix}. \tag{3}$$

E.g. Consider the set of vectors $z_1, z_2, z_3 = \alpha z_1 + \beta z_2, z_4, z_5 = \gamma z_1 + \delta z_4$ z_1, z_2, z_4 being independent of one another. Then we have $z_{k_1} = z_1$, $z_{k_2} = z_2$, $z_{k_3} = z_4$.

§ 3. *Rough Reduction of the Pencil $\rho A + \sigma B$.*

I. DEFINITION. The k *linearly independent* vectors $x_1, x_2, \ldots, x_k$ form an *A-stair* if they satisfy the conditions

$$\begin{aligned} &x_1 B \subset (0), \ (i.e.\ x_1 B = 0) \\ &x_2 B \subset (x_1 A), \\ &x_3 B \subset (x_1 A, x_2 A), \\ &x_4 B \subset (x_1 A, x_2 A, x_3 A), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &x_k B \subset (x_1 A, x_2 A, \ldots, x_{k-1} A). \end{aligned} \tag{1}$$

In the notation of § 2 (2), we may write this as:

$$XB = M{\cdot}XA, \tag{2}$$

where M is a square matrix of degree k in which only the elements below the diagonal can be non-zero. The number k is, of course, less than or equal to m, since there are only m linearly independent vectors x of degree m. Let us suppose that $k < m$ and that *the stair cannot be continued.*

We may add further rows to X to make a non-singular square matrix of degree m, thus

$$\begin{bmatrix} X \\ Y \end{bmatrix}.$$

Let the rows of Y be $y_1, y_{2}, \ldots, y_{m-k}$. The vectors

(3) $$x_1 A, \; x_2 A, \; \ldots, \; x_k A$$

need not be linearly independent. Let their special basis be

(4) $$x_{k_1} A, \; x_{k_2} A, \; \ldots, \; x_{k_l} A$$

which is represented by the matrix

$$\begin{bmatrix} x_{k_1} A \\ x_{k_2} A \\ \vdots \\ x_{k_l} A \end{bmatrix} = \bar{X} A,$$

the rows of $(\bar{X}A)$ being independent. We shall now prove that *the rows of* $\begin{bmatrix} \bar{X}A \\ YB \end{bmatrix}$ *are independent.* Supposing this were not true, we should have a relation

(5) $$(\alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_{m-k} y_{m-k}) B = (\beta_1 x_{k_1} + \beta_2 x_{k_2} + \ldots + \beta_l x_{k_l}) A.$$

The α cannot all vanish for we should then get

$$(\beta_1 x_{k_1} + \beta_2 x_{k_2} + \ldots + \beta_l x_{k_l}) A = 0$$

which is impossible because the vectors (4) are independent.

Hence

$$y = \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_{m-k} y_{m-k}$$

is non-zero and independent of $x_1, x_2, \ldots, x_k$ since the rows of the non-singular matrix $\begin{bmatrix} X \\ Y \end{bmatrix}$ are independent.

From (5) it now follows that

$$yB = (x_{k_1} A, \; x_{k_2} A, \; \ldots, \; x_{k_l} A);$$

or since every x_{k_ν} is a certain x_μ

$$yB \subset (x_1 A, \; x_2 A, \; \ldots, \; x_k A)$$

which would prolong our stair by another step in contradiction to our hypothesis. Hence (5) is impossible. We may therefore add further rows to $\begin{bmatrix} \bar{X}A \\ YB \end{bmatrix}$ to form a non-singular square matrix of degree n,

$$(6) \qquad \begin{bmatrix} \bar{X}A \\ YB \\ Z \end{bmatrix}$$

whose rows form a basis for all vectors of degree m'.

Let

$$k_\lambda = g.$$

According to the properties of our special basis each of the vectors

$$x_1 A,\ x_2 A,\ \ldots,\ x_{g-1} A$$

can be expressed by $x_{k_1} A,\ x_{k_2} A,\ \ldots,\ x_{k_{\lambda-1}} A$. Instead of

$$x_g B \subset (x_1 A,\ x_2 A,\ \ldots,\ x_{g-1} A)$$

(by (1)) we may therefore write

$$x_{k_\lambda} B \subset (x_{k_1} A,\ x_{k_2} A,\ \ldots,\ x_{k_{\lambda-1}} A)$$

or in matrix notation

$$(7) \qquad \bar{X}B = K \cdot \bar{X}A$$

where K (like M in (2)) has non-zero elements only below the main diagonal. As is known, such a matrix has only the latent root zero and a certain power of it must vanish.

Consider the matrix YA. As its rows are vectors of degree m' they must be expressible by the rows of the matrix (6); thus

$$(8) \qquad YA = P\bar{X}A + QYB + RZ.$$

It is obvious that XA and XB can be expressed by the rows of $\bar{X}A$. Let

$$(9) \qquad XA = F \cdot \bar{X}A \quad \text{and} \quad XB = G \cdot \bar{X}A \quad \text{by (1)}.$$

If in $\begin{bmatrix} X \\ Y \end{bmatrix}$ we add a certain aggregate of $x_1,\ x_2,\ \ldots,\ x_k$ or of $x_{k_1},\ x_{k_2},\ \ldots,\ x_{k_l}$ to every row of Y the matrix will still be non-singular. We may for example replace Y by $Y_1 = Y - \Xi\bar{X}$ where Ξ is an arbitrary matrix of type $(m-k) \times l$ which we shall choose in a suitable way. If we carry out this substitution in (8), we get

$$Y_1 A = (P - \Xi)\bar{X}A + Q(Y_1 + \Xi\bar{X})B + RZ$$

and by (7)

$$Y_1 A = (P - \Xi + Q\Xi K)\bar{X}A + QY_1 B + RZ.$$

According to the lemma of § 1 we can choose Ξ so as to make

$$P - \Xi + Q \Xi K$$

vanish. Hence

(10) $$Y_1 A = Q Y_1 B + RZ.$$

If we now multiply the original pencil by $\begin{bmatrix} X \\ Y_1 \end{bmatrix}$, we get by (9) and (10)

$$\begin{bmatrix} X \\ Y_1 \end{bmatrix} (\rho A + \sigma B) = \begin{bmatrix} \rho XA + \sigma XB \\ \rho Y_1 A + \sigma Y_1 B \end{bmatrix} =$$

$$\begin{bmatrix} \rho F + \sigma G & 0 & 0 \\ 0 & \rho Q + \sigma I & \rho R \end{bmatrix} \begin{bmatrix} \overline{X} A \\ Y_1 B \\ Z \end{bmatrix}.$$

The last matrix is non-singular, because

$$\begin{bmatrix} \overline{X} A \\ Y_1 B \\ Z \end{bmatrix} = \begin{bmatrix} \overline{X} A \\ YB - \Xi XB \\ Z \end{bmatrix} = \begin{bmatrix} \overline{X} A \\ YB - \Xi K \overline{X} A \\ Z \end{bmatrix} \text{ by (7)}$$

$$= \begin{bmatrix} I & \cdot & \cdot \\ -\Xi K & I & \cdot \\ \cdot & \cdot & I \end{bmatrix} \begin{bmatrix} \overline{X} A \\ YB \\ Z \end{bmatrix}.$$

Hence the pencil

$$\rho A_1 + \sigma B_1 = \begin{bmatrix} \rho F + \sigma G & \cdot & \cdot \\ \cdot & \rho Q + \sigma I, & \rho R \end{bmatrix}$$

is equivalent to the original pencil. But $\rho A_1 + \sigma B_1$ splits up into two pencils with fewer rows and columns unless $k = m$ (p. 93). Therefore if $k < m$, the proof is completed by induction.

II. We shall now suppose that $k = m$, *i.e.* the longest A-stair contains m independent vectors $x_1, x_2, \ldots, x_m$. We may assume that the original pencil has this property. According to (2) we have

(2) $$XB = MXA,$$

where now X is a non-singular square matrix of degree m and M is a matrix with zero latent roots only.

We have to distinguish two cases.

(a) In $\rho A + \sigma B$ the matrix A has no row dependence: *i.e.* there is no vector $y \neq 0$ for which $yA = 0$. Since we had assumed $m \geqq m'$ it follows $m = m'$ and $|A| \neq 0$. The reduction of $\rho A + \sigma B$ can easily be performed; multiply by X:

$$X(\rho A + \sigma B) = \rho XA + \sigma XB = (\rho I + \sigma M) XA,$$

by (2), where X and XA are non-singular. We may therefore continue by reducing $\rho I + \sigma M$. Since M has only the latent root 0, the Weierstrassian form of M will be

$$PMP^{-1} = \operatorname{diag}(H_{r_1}, H_{r_2}, \ldots. H_{r_l}),\ r_1 + r_2 + \ldots. + r_l = m = m'$$

where

$$H_r = \begin{bmatrix} . & . & . & . \\ 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \end{bmatrix}_r .$$

Hence

$$P(\rho I + \sigma M)P^{-1} = \operatorname{diag}(\rho I_{r_1} + \sigma H_{r_1}, \rho I_{r_2} + \sigma H_{r_2}, \ldots., \rho I_{r_p} + \sigma H_{r_l})$$
$$= \operatorname{diag}(N_{r_1}, N_{r_2}, \ldots., N_{r_l})$$

which proves the theorem.

(*b*) We have now to deal with the more difficult case when a vector $y \neq 0$ exists for which $yA = 0$. It is then possible to construct a "B-stair" in the same way as in (*a*) only with A and B interchanged. Every other step remains unaltered: We construct a stair whose length[1] may be l. If l be less than m, we should again be able to split up the pencil and the proof would be concluded by induction. We shall therefore suppose that not only the A-stair but also the B-stair exhausts the whole m-dimensional vector-space. Writing these conditions down in full, we have

$$(11)\quad (\alpha)\ \begin{array}{l} x_1 B = 0 \\ x_2 B \subset (x_1 A) \\ x_3 B \subset (x_1 A, x_2 A) \\ \ldots\ldots\ldots\ldots\ldots \\ x_m B \subset (x_1 A, x_2 A, \ldots., x_{m-1} A) \end{array} \qquad (\beta)\ \begin{array}{l} y_1 A = 0 \\ y_2 A \subset (y_1 B) \\ y_3 A \subset (y_1 B, y_2 B) \\ \ldots\ldots\ldots\ldots\ldots \\ y_m A \subset (y_1 B, y_2 B, \ldots., y_{m-1} B), \end{array}$$

where $x_1, x_2, \ldots., x_m$ and $y_1, y_2, \ldots., y_m$ are two sets of m linearly independent vectors of degree m. Pencils $\rho A + \sigma B$ with the properties (1) require a more elaborate study which we are going to explain in § 4.

§ 4. *Reduction by means of Vector Chains.*

Let $\rho A + \sigma B$ be a pencil which fulfils the condition (11) of § 3, *i.e.* we assume that at least one B-stair and one A-stair exists, each of length m. But it is easy to see that every non-zero vector z that annihilates B can be extended to a stair of m elements unless the

[1] By saying the stair is of length l, we mean that it consists of l vectors and cannot be continued by another vector.

pencil splits up into two pieces. For if a stair beginning with z should break down at the k^{th} step, *i.e.* if the stair be of length $k\ (k < m)$, we could split up the pencil as shown in § 3.

From § 3 (11), we see that for every vector u we can find a vector $\bar{u}$ such that

$$(1) \qquad uA = \bar{u}B,$$

for u must be a linear aggregate of $y_1, y_2, \ldots, y_m$ whence the existence of $\bar{u}$ is evident from § 3 (11) β. It is not self-evident that the unknown components of the vector $\bar{u}$ can be calculated from the non-homogeneous equation (1) because the coefficients of the unknowns do not form a non-singular matrix. The vector $\bar{u}$, however, is not uniquely determined.

Let $v_0 B = 0 \ (v_0 \neq 0)$. We may then determine other vectors $v_1, v_2, \ldots, v_{p_1}, \ldots$, which form the following "*vector chain.*" (*cf.* Turnbull, page 72 of this volume.)

$$(2) \qquad 0 = v_0 B, \quad v_0 A = v_1 B, \quad v_1 A = v_2 B, \ \ldots, \ v_{p_1 - 1} A = v_{p_1} B, \\ v_{p_1} A = v_{p_1 + 1} B, \ \ldots .$$

We can continue the chain as long as we want, but the vectors occurring in it will not be linearly independent. Let $v_{p_1} A$ be the first vector in (2) to be linearly dependent on its predecessors $v_0 A, v_1 A, \ldots, v_{p_1 - 1} A$. We then have the relation

$$(3) \qquad \left(\sum_{\nu=0}^{p_1} a_\nu v_{p_1 - \nu}\right) A = 0, \text{ where } a_0 \neq 0.$$

It is convenient to put

$$(4) \qquad v_{-k} = 0, \qquad k = 1, 2, 3, \ldots,$$

making the equation $v_{\nu-1} A = v_\nu B$ valid also for zero and negative integersν. The number p_1, *i.e.* the number of consecutive linearly independent vectors in (2) starting with $v_0 A$ is called the length of the chain. The length is always positive, otherwise we should have $v_0 B = v_0 A = 0$ and the pencil $\rho A + \sigma B$ would be degenerate (§ 1). Let p_1 be as small as possible. We derive another chain from (2) by putting

$$(5) \qquad u_k^{(1)} = \sum_{\nu=0}^{p_1} a_\nu v_{k-\nu} \qquad (k \leqq p_1).$$

In fact, the $u_k^{(1)}$ form a chain, for by (2)

$$u_k^{(1)} B = \left(\sum_{\nu=0}^{p_1} a_\nu v_{k-\nu}\right) B = \left(\sum_{\nu=0}^{p_1} a_\nu v_{k-1-\nu}\right) A = u_{k-1}^{(1)} A.$$

In particular $u_0^{(1)} B = 0$ by (4) and

$$u_{p_1}^{(1)} A = \Big(\sum_{\nu=0}^{p_1} a_\nu v_{p_1-\nu} \Big) A = 0 \text{ by (3).}$$

We have therefore constructed the chain

$$(6) \quad 0 = u_0^{(1)} B,\ u_0^{(1)} A = u_1^{(1)} B,\ u_1^{(1)} A = u_2^{(1)} B,\ \ldots,\ u_{p_1-1}^{(1)} A = u_{p_1}^{(1)} B,\ u_{p_1}^{(1)} A = 0$$

The vectors $u_0^{(1)} A,\ u_1^{(1)} A,\ \ldots,\ u_{p_1-1}^{(1)} A$ must be independent, otherwise we could build up a chain of length less than p_1 which would be contradictory.

If there is a vector $u_0^{(2)} \neq 0$ which annihilates B and which is independent of the first chain, *i.e.* of the vectors $u_0^{(1)}, u_1^{(1)}, \ldots, u_{p_1}^{(1)}$ we form another chain like (6) the length p_2 of which shall be taken as small as possible. Naturally $p_1 \leqq p_2$. We then proceed to a third chain provided that its first or "leading" vector $u_0^{(3)}$ is independent of all vectors of the first and second chain its length p_3 being minimal. In this way we get a whole system of chains

$$(7) \quad \begin{array}{l} 0 = u_0^{(1)} B,\ u_0^{(1)} A = u_1^{(1)} B,\ u_1^{(1)} A = u_2^{(1)} B, \ldots,\ u_{p_1-1}^{(1)} A = u_{p_1}^{(1)} B,\ u_{p_1}^{(1)} A = 0 \\ 0 = u_0^{(2)} B,\ u_0^{(2)} A = u_1^{(2)} B,\ u_2^{(2)} A = u_2^{(2)} B, \ldots,\ u_{p_2-1}^{(2)} A = u_{p_2}^{(2)} B,\ u_{p_2}^{(2)} A = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0 = u_0^{(n)} B,\ u_0^{(n)} A = u_1^{(n)} B,\ u_1^{(n)} A = u_2^{(n)} B, \ldots,\ u_{p_n-1}^{(n)} A = u_{p_n}^{(n)} B,\ u_{p_n}^{(n)} A = 0 \end{array}$$

As we have shown, this system possesses the following properties :

(*a*) The lengths are increasing

$$(8) \qquad p_1 \leqq p_2 \leqq \ldots\ldots \leqq p_n.$$

(*b*) The first vector of every chain is independent of all vectors of the preceding chains.

(*c*) Each length is as small as possible, *i.e.* there is no chain independent of the first chain whose length is less than p_2, nor does a chain exist whose first vector is independent of the first and second chains and the length of which is less than p_3, etc.

(*d*) We have exhausted all chains, *i.e.* we cannot find any vector $u_0^{(n+1)}$ for which $u_0^{(n+1)} B = 0$ unless $u_0^{(n+1)}$ is a linear aggregate of the previous chains.

We shall now prove that the vectors

(9) $u_0^{(1)} A, u_1^{(1)} A, \ldots, u_{p_1-1}^{(1)} A, u_0^{(2)} A, u_1^{(2)} A, \ldots, u_{p_2-1}^{(2)} A, \ldots, u_0^{(n)} A, u_1^{(n)} A, \ldots, u_{p_n}^{(n)-1} A$ are independent of one another. If this were not so, we should have a relation

$$(10) \qquad \Big(\sum_{\tau=1}^{n} \sum_{\mu_\tau = 0}^{q_\tau} \beta_{\mu_\tau}^{(\tau)} u_{q_\tau - \mu_\tau}^{(\tau)} \Big) A = 0$$

where

(11) $$q_\tau \leqq p_\tau - 1,$$

and $u^{(\tau)}_{q_\tau}$ is the last element of the τth chain that really enters the relation (10) with a non-zero coefficient $\beta^{(\tau)}_0 \neq 0$.

If the τth chain does not occur at all in (10), we put $q_\tau = 0$ and $\beta^{(\tau)}_0 = 0$. Let q_g be the maximum of $q_1, q_2, \ldots, q_n$; if several q are equally great, we take g as great as possible so that

(12) $$q_g \geqq q_k \quad (k = 1, 2, \ldots, g); \quad q_g > q_\lambda \quad (\lambda = g+1, \ldots, n).$$

We now construct the chain

(13) $$v_k = \sum_{\tau=1}^{n} \sum_{\mu_\tau = 0}^{q_\tau} \beta^{(\tau)}_{\mu_\tau} \; u^{(\tau)}_{k+q_\tau - q_g - \mu_\tau}.$$

In fact, the vectors $v_0, v_1, \ldots, v_{q_g}$ form a chain. For

$$v_k B = (\sum_{\tau=1}^{n} \sum_{\mu_\tau=0}^{q_\tau} \beta^{(\tau)}_{\mu_\tau} u_{k+q_\tau - q_g - \mu_\tau}) B = (\sum_{\tau=1}^{n} \sum_{\mu_\tau=0}^{q_\tau} \beta^{(\tau)}_{\mu_\tau} u_{k-1+q_\tau - q_g - \mu_\tau}) A$$

$= v_{k-1} A$ because according to the chain properties (7) we have $u^{(\tau)}_\nu B = u^{(\tau)}_{\nu-1} A$ for every $\nu \leqq p_\tau$.

In particular we get $v_0 B = v_{-1} A = 0$ and $v_{q_g} A = 0$ by (10). Also v_0 reduces to

$$v_0 = \beta^{(1)}_0 u^{(1)}_{q_1 - q_g} + \beta^{(2)}_0 u^{(2)}_{q_2 - q_g} + \ldots\ldots + \beta^{(g)}_0 u^{(g)}_0 \text{ (by (4) and } \beta^{(g)}_0 \neq 0).$$

The suffixes of the u are either 0 or negative since $q_g \geqq q_\tau$ ($\tau = 1, 2, \ldots h$). All terms behind the g^{th} term could be dropped because $q_g > q_\lambda$ for $\lambda > g$. v_0 is independent of the first, second, .., $(g-1)^{\text{th}}$ chain. For, otherwise $u^{(g)}_0$ would be dependent upon its predecessors in contradiction to (b). It is therefore permissible to start the g^{th} chain with v_0 instead of $u^{(g)}_0$. But the length of the v-chain is $q_g \leqq p_g - 1$ or less, viz. if the vectors $v_0 A, v_1 A, \ldots, v_{q_g - 1} A$ be linearly dependent. In any case the length of this modified g^{th} chain would be smaller than p_g which contradicts (c). Hence the vectors (9) must be independent of each other.

We shall now show that also the vectors

(14) $$u^{(1)}_0, u^{(1)}_1, \ldots, u^{(1)}_{p_1}; \; u^{(2)}_0, u^{(2)}_1, \ldots, u^{(2)}_{p_2}; \; \ldots; \; u^{(n)}_0, u^{(n)}_1, \ldots, u^{(n)}_{p_n}$$

are linearly independent.

If there were a relation between them, it could be written :

(15) $$\gamma_1 u^{(1)}_0 + \gamma_2 u^{(2)}_0 + \ldots + \gamma_n u^{(n)}_0 + \sum_{\tau=1}^{n} \sum_{\mu_\tau=1}^{p_\tau} \delta^{(\tau)}_{\mu_\tau} u^{(\tau)}_{\mu_\tau} = 0.$$

The $\delta^{(\tau)}_{\mu_\tau}$ cannot all vanish. For then the "leading" vectors $u_0^{(1)}, u_0^{(2)}, \ldots, u_0^{(n)}$ would be dependent in contradiction to (b). Multiplying (15) by B we get

$$\Big(\sum_{\tau=1}^{n}\ \sum_{\mu_\tau=1}^{p_\tau} \delta^{(\tau)}_{\mu_\tau} u^{(\tau)}_{\mu_\tau}\Big) B = 0$$

since $u_0^{(\tau)} B = 0$; applying the chain properties (7) we have

$$\Big(\sum_{\tau=1}^{n}\ \sum_{\mu_\tau=0}^{p_\tau-1} \delta^{(\tau)}_{\mu_\tau+1} u^{(\tau)}_{\mu_\tau}\Big) A = 0$$

which is incompatible with the vectors (9) being independent.

Hence the vectors (14) are independent.

What are the connections between the vector chains and the reduction of the pencil $\rho A + \sigma B$? Consider one of the chains (7):

$$0=u_0^{(\tau)}B,\ u_0^{(\tau)}A=u_1^{(\tau)}B,\ u_1^{(\tau)}A=u_2^{(\tau)}B, \ldots\ldots,\ u^{(\tau)}_{p_\tau-1}A=u^{(\tau)}_{p_\tau}B,\ u^{(\tau)}_{p_\tau}A=0.$$

Let

$$(16)\qquad U_\tau=\begin{bmatrix} u_0^{(\tau)}\\ u_1^{(\tau)}\\ \vdots\\ u^{(\tau)}_{p_\tau}\end{bmatrix} \text{ and } \overline{U}_\tau=\begin{bmatrix} u_0^{(\tau)}\\ u_1^{(\tau)}\\ \vdots\\ u^{(\tau)}_{p_\tau-1}\end{bmatrix} \qquad (\tau=1,2,\ldots,n).$$

It follows by (7) that

$$U_\tau(\rho A+\sigma B)=\begin{bmatrix} \rho u_0^{(\tau)}A+\sigma u_0^{(\tau)}B\\ \rho u_1^{(\tau)}A+\sigma u_1^{(\tau)}B\\ \cdots\cdots\cdots\cdots\\ \rho u^{(\tau)}_{p_\tau}A+\sigma u^{(\tau)}_{p_\tau}B\end{bmatrix}=\begin{bmatrix} \rho u_0^{(\tau)}A\\ \rho u_1^{(\tau)}A+\sigma u_0^{(\tau)}A\\ \cdots\cdots\cdots\cdots\\ \sigma u^{(\tau)}_{p_\tau-1}A\end{bmatrix}$$

and

$$(17)\qquad U_\tau(\rho A+\sigma B)=\begin{bmatrix} \rho & \cdot & \cdot & \cdot & \cdot\\ \sigma & \rho & \cdot & \cdot & \cdot\\ \cdot & \sigma & \rho & \cdot & \cdot\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdot & \cdot & \cdot & \cdot & \rho\\ \cdot & \cdot & \cdot & \cdot & \sigma\end{bmatrix}\overline{U}_\tau A=\Lambda_{p_\tau}\overline{U}_\tau A,$$

where Λ_{p_τ} has been defined in § 1 (3).

Hence

$$\begin{bmatrix} U_1\\ U_2\\ \vdots\\ U_n\end{bmatrix}(\rho A+\sigma B)=\begin{bmatrix} U_1(\rho A+\sigma B)\\ U_2(\rho A+\sigma B)\\ \vdots\\ U_n(\rho A+\sigma B)\end{bmatrix}=\begin{bmatrix} \Lambda_{p_1} & & & \\ & \Lambda_{p_2} & & \\ & & \ddots & \\ & & & \Lambda_{p_n}\end{bmatrix}\begin{bmatrix} \overline{U}_1 A\\ \overline{U}_2 A\\ \vdots\\ \overline{U}_n A\end{bmatrix}$$

or

(18) $$U(\rho A + \sigma B) = \Lambda \cdot \overline{U}A,$$

where

$$\Lambda = \text{diag}\,(\Lambda_1, \Lambda_2, \ldots, \Lambda_n)$$

and

(19) $$U = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix} \text{ and } \overline{U} = \begin{bmatrix} \overline{U}_1 \\ \overline{U}_2 \\ \vdots \\ \overline{U}_n \end{bmatrix}.$$

Obviously, the equations (7) can be interpreted as a vector A-stair in the sense explained in § 2. It contains $k=(p_1+1)+(p_2+1)+\ldots+(p_n+1)$ vectors the independency of which we have proved.

We shall show that $k=m$. If $k<m$, it must be possible to continue the stair by another vector z such that

(20) $$zB \subset (u_0^{(1)}A, u_1^{(1)}A, \ldots, u_{p_n-1}^{(n)}A, u_{p_n}^{(n)}A)$$

z being independent of all u. By (7) we may write instead of (20)

$$zB \subset (u_1^{(1)}B, u_2^{(1)}B, \ldots, u_{p_n}^{(n)}B)$$

or in full

$$(z - (\epsilon_1^{(1)}u_1^{(1)} + \epsilon_2^{(1)}u_2^{(1)} + \ldots + \epsilon_{p_n}^{(n)}u_{p_n}^{(n)}))\,B = 0$$

$\epsilon_{\nu_\tau}^{(\tau)}$ being certain coefficients. Here we should have obtained a vector which is independent of the u and yet annihilates B in contradiction to condition d). Hence k must be m and U has m rows and is therefore square and non-singular.

Finally, we shall show that also $\overline{U}A$ is square (of degree m). If it were not so, we could add further rows to make a non-singular square matrix $\begin{bmatrix} \overline{U}A \\ Z \end{bmatrix}$.

From (18) we should then get

$$U(\rho A + \sigma B) = [\Lambda, 0] \begin{bmatrix} \overline{U}A \\ Z \end{bmatrix}.$$

Hence $[\Lambda, 0]$ would be equivalent to $\rho A + \sigma B$ but it contains null rows and columns which we had excluded. The matrix Z must therefore be illusory and (18) may be written as

$$U(\rho A + \sigma B)(\overline{U}A)^{-1} = \text{diag}\,(\Lambda_1, \Lambda_2, \ldots, \Lambda_n).$$

This completes the proof.

In his paper Professor Turnbull has shown how the minimal vector chains are connected with Kronecker's minimal relations between the rows of the pencil $\rho A + \sigma B$. In particular, it has been pointed out that the lengths of the vector chains (7) are identical with Kronecker's *Minimalgradzahlen.*

PART II

THE AUTOMORPHIC TRANSFORMATIONS OF A SINGULAR MATRIX PENCIL.

§1.

In the theory of Canonical Matrices two matrix pencils

$$\Gamma_1 = \rho \mathcal{A}_1 + \sigma \mathcal{B}_1 \quad \text{and} \quad \Gamma_2 = \rho \mathcal{A}_2 + \sigma \mathcal{B}_2$$

each with m rows and n columns, are said to be <u>equiv-</u> <u>-alent</u> if two constant non-singular matrices P and Q of degrees m and n exist such that

(1) $$\mathcal{P} \Gamma_1 \mathcal{Q} = \Gamma_2 \qquad ; \quad |P| \neq 0, \quad |Q| \neq 0$$

The fundamental result is that two pencils can be reduced to the same canonical form, if and only if, they are equivalent. Suppose now that Γ_1 and Γ_2 are equivalent; then it is possible to solve (1) for P and Q, and the natural question arises, what is the most general solution of this equation. It is easily seen that this problem is equivalent to finding the most general pair of matrices P ,Q which transforms a given pencil $\Gamma = \rho \mathcal{A} + \sigma \mathcal{B}$ into itself,i.e.,which satisfies the equation

(2) $$\mathcal{P} \Gamma \mathcal{Q} = \Gamma \quad ,$$

or, comparing coefficients of ρ and σ

(2a) $$\mathcal{P} \mathcal{A} \mathcal{Q} = \mathcal{A} \quad ;$$

(2b) $$\mathcal{P} \mathcal{B} \mathcal{Q} = \mathcal{B} \quad .$$

In this case we shall say that the pair of matrices $(\mathcal{P}, \mathcal{Q})$ is an automorphic transformation of Γ and we propose to determine all such automorphic transformations; in particular, we shall express the number of linearly independent ones among them in terms of the invariants of Γ ,i.e.,in terms of the Invariant Factors and the Kronecker Indices of Γ .

Let us now make some convenient assumptions regarding $\Gamma = \rho A + \sigma B$ without restricting the generality of our investigations.

First we shall assume that A is of the same rank as Γ itself, a condition which can always be fulfilled by a linear transformation of the variables ρ, σ . For, let ρ_0, σ_0 be such that

$$A_1 = \rho_0 A + \sigma_0 B$$

is of maximun rank, i.e., of the same rank as Γ itself, and put

$$B_1 = \rho_1 A + \sigma_1 B$$

where ρ_1 and σ_1 are only subject to the condition

(3) $$\begin{vmatrix} \rho_0 & \rho_1 \\ \sigma_0 & \sigma_1 \end{vmatrix} \neq 0 .$$

If we now introduce new variables ρ', σ' by the trans-formation

$$\rho = \rho_0 \rho' + \rho_1 \sigma'$$
$$\sigma = \sigma_0 \rho' + \sigma_1 \sigma'$$

we see that Γ can be written as

$$\Gamma = \rho' A_1 + \sigma' B_1$$

and that it now has the property required. Again, the automorphic transformations remain the same, since the equations

(4a) $$\mathcal{P} A_1 \mathcal{Q} = A_1$$ and

(4b) $$\mathcal{P} B_1 \mathcal{Q} = B_1$$

are equivalent to (2a) and(2b) on account of (3). After this preliminary remark we shall replace the two homogeneous variables ρ, σ by <u>one</u> variable, λ, and write the pencil in the form

$$\Gamma = \lambda A + B$$

where A is of the same rank as Γ.

Next, Γ may be replaced by any pencil Γ_o which is equivalent to Γ. For, let

$$\Gamma_o = S \Gamma T \qquad ; \; |S| \neq 0 \, , |T| \neq 0 \, .$$

Then, if (P,Q) is an automorphic transformation of Γ, (SPS^{-1},T^{-1}QT) is an automorphic transformation of Γ_o and vice versa. Thus a (1-1) correspondence is established between the automorphic tranformations of Γ and Γ_o. In particular, we may assume that Γ_o is the canonical form[1)] of Γ which we write in the form

(5) $$\Gamma_o = \Gamma_1 \dotplus \Gamma_2 \dotplus \Gamma_3$$

WHERE

(5,1) $$\Gamma_1 = \sum_{i=1}^{p} (\alpha_i \Lambda_{m_i}) \qquad (0 < m_1 < m_2 < \cdots < m_p)$$

(5,2) $$\Gamma_2 = \lambda I_k + \mathcal{M}_k$$

(5,3) $$\Gamma_3 = \sum_{j=1}^{q} (\beta_j \Lambda_{n_j}) \qquad (0 < n_1 < n_2 < \cdots < n_q)$$

using the "direct sum" of matrices which is defined as follows :

$$\Gamma_1 \dotplus \Gamma_2 \dotplus \cdots \dotplus \Gamma_r = \sum_{i=1}^{r} \Gamma_i = \mathrm{diag}(\Gamma_1, \Gamma_2, .., \Gamma_r) = \begin{bmatrix} \Gamma_1 & & \\ & \Gamma_2 & \\ & & \ddots \\ & & & \Gamma_r \end{bmatrix} ;$$

and when $\Gamma_1 = \Gamma_2 = \cdots = \Gamma_r = \Gamma$, we write

$$\Gamma \dotplus \Gamma \dotplus \cdots \dotplus \Gamma = (r\,\Gamma) \, .$$

In (5,1) and (5,3) Λ_{m_i} is the typical singular submatrix corresponding to a row vector of minimal degree m_i annihilating Γ, e.g.,

---

1) See Turnbull and Aitken " Canonical Matrices" (Glasgow 1932) and W.Ledermann Proc.Edin.Math.Soc.(2) Vol 4 (1934). [= part I of this Thesis.]

$$\Lambda_1 = \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \; ; \; \Lambda_2 = \begin{bmatrix} \lambda & \cdot \\ 1 & \lambda \\ \cdot & 1 \end{bmatrix}, \; \Lambda_3 = \begin{bmatrix} \lambda & \cdot & \cdot \\ 1 & \lambda & \cdot \\ \cdot & 1 & \lambda \\ \cdot & \cdot & 1 \end{bmatrix}; \cdots$$

and in general

$$(6) \qquad \Lambda_s = \begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & 1 & & \\ & & & \lambda \\ & & & 1 \end{bmatrix}_{s+1,\, s}$$

We remark that Λ_s has s+1 rows and s columns. The pencil Γ_2 in (5,2) is the " nonsingular core of Γ " of type $k \times k$,say, and there is no loss of generality in assuming that the matrix coefficient of λ in Γ is the unit matrix; for since that coefficient is of the same rank as Γ_2 , it must be nonsingular and may be removed as a factor.

The Kronecker Minimal Indices are exhibited in the canonical form,viz:

α_1 times m_1 , α_2 times m_2 , ... for the first set,

β_1 times n_1 , β_2 times n_2 , ...for the second set,

these numbers together with the Invariant Factors of are the invariants of Γ , they are the same for all pencils equivalent to Γ . In what follows we shall assume that Γ is already in canonical form, i.e., $\Gamma_0 = \Gamma$.

Since $|Q| \neq 0$, we may put $R = Q^{-1}$ and write (2) as

$$(7) \qquad P\Gamma = \Gamma R \quad , \; |P| \neq 0 \, , \; |R| \neq 0$$

or by (5)

$$(8) \qquad P(\Gamma_1 \dotplus \Gamma_2 \dotplus \Gamma_3) = (\Gamma_1 \dotplus \Gamma_2 \dotplus \Gamma_3)R \, .$$

We now partition P and R in accordance with the three isolated submatrices of Γ, i.e., we put

$$\mathcal{P} = [\mathcal{P}_{ij}] \quad ; \quad R = [R_{ij}] \qquad (i, j = 1, 2, 3)$$

Hence (8) resolves into nine partial equations

$$\mathcal{P}_{ij}\,\Gamma_j = \Gamma_i\,R_{ij}$$

or in full

$$(9)\quad \begin{cases} (i)\ \mathcal{P}_{11}\Gamma_1 = \Gamma_1 R_{11} & (iv)\ \mathcal{P}_{12}\Gamma_2 = \Gamma_1 R_{12} & (vii)\ \mathcal{P}_{13}\Gamma_3 = \Gamma_1 R_{13} \\ (ii)\ \mathcal{P}_{21}\Gamma_1 = \Gamma_2 R_{21} & (v)\ \mathcal{P}_{22}\Gamma_2 = \Gamma_2 R_{22} & (viii)\ \mathcal{P}_{23}\Gamma_3 = \Gamma_2 R_{23} \\ (iii)\ \mathcal{P}_{31}\Gamma_1 = \Gamma_3 R_{31} & (vi)\ \mathcal{P}_{32}\Gamma_2 = \Gamma_3 R_{32} & (ix)\ \mathcal{P}_{33}\Gamma_3 = \Gamma_3 R_{33}\,, \end{cases}$$

and it is the object of the following pages to give a complete solution to these equations, the m^2+n^2 elements of P and R being regarded as the unknowns. Only equation (V) seems to have received attention in the literature. For, substituting for Γ_2 from (5,2) and suppressing unnecessary indices we can write this equation as

$$\mathcal{P}(\lambda I + \mathcal{M}) = (\lambda I + \mathcal{M})R$$

whence
$$\mathcal{P} = R$$
and
$$\mathcal{P}\mathcal{M} = \mathcal{M}R$$
Hence
$$\mathcal{P}\mathcal{M} = \mathcal{M}\mathcal{P}$$

The solution of (v), therefore, involves the finding of the most general matrix P that commutes with a given matrix M.

This problem was first solved by Frobenius and has since been treated by several authors[1]:

---

[1] Frobenius, " Ueber die mit einer Matrix vertauschbaren Matrizen" Berl.Sitzb.(1910) where other references may be found, also D.E.Rutherford, Proc. Amsterdam Vol 35 (1932)

It was found that the number of linearly independent solutions of (V) is

(10) $$t_5 = \tau = e_1 + 3e_2 + 5e_3 + 7e_4 + \cdots$$

where e_ν is the degree in λ of the ν^{th} Invariant Factor of the matrix $\lambda I_k + \mathcal{M}_k$. In what follows we shall obtain similar results for the remaining eight equations (9) to which we shall refer later simply by Roman numerals. The total number $\underline{t}$ of parameters in the general solution of (7) is equal to the sum of the subtotals $t_1, t_2, \ldots, t_9$ giving the numbers of parameters in the solutions of those nine equations (9). It will be found that three of these numbers are zero, i.e., that the corresponding equations have only the trivial soltution in which all unknowns vanish, and the final result will be

$$t = \sum_{s=1}^{9} t_s = \tau + \sum_{i \geq j} \alpha_i \alpha_j (m_i - m_j + 1) + \sum_{j \geq i} \beta_i \beta_j (n_j - n_i + 1) + k\Big(\sum_i \alpha_i + \sum_j \beta_j\Big) + \sum_{i,j} \alpha_i \beta_j (m_i + n_j).$$

§2.

Before solving the commutantal equations §1(9), we shall make some remarks on the typical singular submatrix (§1,(6)):

(1) $$\Lambda_s = \begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & 1 & \ddots & \\ & & & \lambda \\ & & & 1 \end{bmatrix}_{s+1,s} = \lambda \begin{bmatrix} I_s \\ \cdot \end{bmatrix} + \begin{bmatrix} \cdot \\ I_s \end{bmatrix},$$

where I_s is the unit matrix of degree s, and the dot below or above indicates a row of zeros.

First, we wish to determine row vectors and column vectors that annihilate Λ_s. Let

$$x = \{ \xi_1, \ \xi_2, \ \ldots, \ \xi_s \}$$

be a column vector with s elements and

$$y = [\eta_0, \eta_1, \ldots, \eta_s]$$

a row vector with s+1 elements. We then prove the following lemma :

LEMMA I.

The equation

(2) $$\Lambda_s x = 0$$

admits only of the equation $x = 0$, i.e., the columns of Λ_s are linearly independent ; the general solution of

(3) $$y \Lambda_s = 0$$

is $$y = \phi(\lambda)[1, -\lambda, (-\lambda)^2, \ldots, (-\lambda)^s]$$

where $\phi(\lambda)$ is an arbitrary function of λ and

$$u_s = [1, -\lambda, (-\lambda)^2, \ldots, (-\lambda)^s]$$

is the " vector of apolarity" of degree s.

Proof : The rank of Λ_s is s , because on cancelling the first row of Λ_s we obtain a minor which is identically equal to unity. In the set of homogeneous equations (2) , the number of unknowns therefore equals the rank, and by the fundamental theorem of linear equations the system has only the trivial solution $x = 0$. In the set of equations (3) , however, the number of unknowns exceeds the rank by unity and the most general solution is a scalar multiple of any particular solution which may be taken to be the vector of apolarity u_s , since

(4) $$u_s \Lambda_s = 0$$

as is easily verified.

LEMMA II

If P is a constant matrix with s+1 rows, then the equation

$$u_s P = 0$$

is impossible unless

$$P = 0$$

The proof follows immediately on writing out (5) in full and equating the coefficients of $1, \lambda, \lambda^2, \dots, \lambda^s$ to zero.

---

§3.

We now turn to the discussion of equation §1,(9).

consider the simple case

(1) $$P\Lambda_s = \Lambda_{s'} R$$

which is solved by the following theorem:

THEOREM I

When $s' < s$, equation (1) has only the trivial solution

$$P = 0 \ , \quad R = 0 \quad ;$$

when $s' \geq s$ and $0 \geq s' - s = d$, say, then the general solution is

(3) $$P = [\phi_{i-j}]_{s'+1,\, s+1} \qquad \begin{pmatrix} i = 0, 1, \dots, s' \\ j = 0, 1, \dots, s \end{pmatrix}$$

and

$$R = [\phi_{i-j}]_{s',\, s} \qquad \begin{pmatrix} i = 1, 2, \dots, s' \\ j = 1, 2, \dots, s \end{pmatrix}$$

where $\phi_0, \phi_1, \dots, \phi_d$ are d+1 arbitrary constants and

$$\phi_{i-j} = 0 \begin{cases} \text{for } i-j > d \\ \text{for } i-j < 0. \end{cases}$$

E.g., in the case $s' = 5, s = 3, d = 2$, we have

$$\mathcal{P} = \begin{bmatrix} \phi_0 & \cdot & \cdot & \cdot \\ \phi_1 & \phi_0 & \cdot & \cdot \\ \phi_2 & \phi_1 & \phi_0 & \cdot \\ \cdot & \phi_2 & \phi_1 & \phi_0 \\ \cdot & \cdot & \phi_2 & \phi_1 \\ \cdot & \cdot & \cdot & \phi_2 \end{bmatrix} \quad ; \quad R = \begin{bmatrix} \phi_0 & \cdot & \cdot \\ \phi_1 & \phi_0 & \cdot \\ \phi_2 & \phi_1 & \phi_0 \\ \cdot & \phi_2 & \phi_1 \\ \cdot & \cdot & \phi_2 \end{bmatrix}$$

<u>Proof</u> : On premultiplying (1) by $u_{s'}$ we obtain by § 2,(4)

$$(u_{s'}\mathcal{P})\Lambda_s = 0$$

i.e., the vector $(u_{s'}\mathcal{P})$ annihilates Λ_s . Hence by lemma I (§2) we have (5)

(5) $$u_{s'}\mathcal{P} = \phi(\lambda)\,[1, -\lambda, (-\lambda)^2, \ldots, (-\lambda)^s].$$

The elements of the vector on the left hand side are polynomials of degree not higher than s'. Hence compar-ing the first elements of either side we see that $\phi(\lambda)$ is also such a polynomial.

Now, when s' is less than s , equation (5) is obvi-ously impossible unless $\phi(\lambda)$ is equal to zero and hence

(6) $$u_{s'}\mathcal{P} = 0$$

which, by lemma II (§ 2) implies

$$\mathcal{P} = 0 \ .$$

Equation (1) then becomes $0 = \Lambda_s R$ which entails

$$R = 0$$

since the columns of Λ_s are linearly independent. This proves the first part of the theorem.

Next, when s' ≥ s, (5) can evidently be solved and $\phi(\lambda)$ will generally be a polynomial of degree d = s'-s,

$$\phi(\lambda) = \phi_0 - \phi_1\lambda + \phi_2\lambda^2 - \cdots + \phi_d(-\lambda)^d \quad \text{, say.}$$

Putting

$$\mathcal{P} = [p_{ij}] \qquad (i = 0, 1, \ldots, s'; j = 0, 1, \ldots, s)$$

and comparing the j-th elements of either side of (5) we obtain

$$p_{oj} + p_{1j}(-\lambda) + p_{2j}(-\lambda)^2 + \cdots + p_{s'j}(-\lambda)^{s'} = \phi_o(-\lambda)^j + \phi_1(-\lambda)^{j+1} + \phi_2(-\lambda)^{j+2} + \cdots + \phi_d(-\lambda)^{j+d}$$

$$(j = 0, 1, 2, \ldots, s)$$

Hence

$$p_{oj} = p_{1j} = \cdots = p_{j-1,j} = 0$$

$$p_{jj} = \phi_o \ ; \ p_{j+1,j} = \phi_1 \ ; \ \cdots ; \ p_{j+d,d} = \phi_d$$

$$p_{j+d+1,j} = \cdots = p_{s',j} = 0$$

or

$$p_{ij} = \begin{cases} 0 \text{ for } i-j < 0 \\ \phi_{i-j} \text{ for } 0 \leq i-j \leq d \\ 0 \text{ for } i-j > d \end{cases}$$

which proves the statement (3) regarding P. In order to

*/ determine R, we substitute § 2,(1) in (1):

$$\mathcal{P}\left\{\lambda \begin{bmatrix} I_{s'} \\ \cdot \end{bmatrix} + \begin{bmatrix} \cdot \\ I_{s'} \end{bmatrix}\right\} = \left\{\lambda \begin{bmatrix} I_s \\ \cdot \end{bmatrix} + \begin{bmatrix} \cdot \\ I_s \end{bmatrix}\right\} R .$$

Comparing the constant terms we get

$$\mathcal{P} \begin{bmatrix} \cdot \\ I_{s'} \end{bmatrix} = \begin{bmatrix} \cdot \\ I_s \end{bmatrix} R$$

which on premultiplication by $[\cdot \ I_s]$ becomes

$$[\cdot \ I_s] \, \mathcal{P} \begin{bmatrix} \cdot \\ I_s \end{bmatrix} = R .$$

The matrix R is thus expressed in terms of P, and it is readily seen that R is obtained by cancelling the first row and column of P. This proves equation(4) .It is easily verified that (4) and (5) are also sufficient conditions that P and R should satisfy (1).

Corollary:

When $s' = s$, the general solution of

$$\mathcal{P}\Lambda_s = \Lambda_s R$$

is

$$\mathcal{P} = \phi_o I_{s+1}$$

$$R = \phi_o I_s$$

involving one parameter ϕ_0. The solution is nonsingular, if and only if $\phi_0 \neq 0$.

We now come to the equation

$$\mathcal{P}(\alpha\Lambda_s) = (\alpha'\Lambda_{s'})R \tag{7}$$

where $(\alpha\Lambda_s) = \Lambda_s \dotplus \Lambda_s \dotplus \cdots \dotplus \Lambda_s$ (α times repeated). Writing P and R as

$$\mathcal{P} = [\mathcal{P}_{ij}] \quad ; \quad R = [R_{ij}] \quad \begin{pmatrix} i = 1, 2, \ldots, \alpha' \\ j = 1, 2, \ldots, \alpha \end{pmatrix}$$

we see that (7) resolves into $\alpha\alpha'$ matrix equations:

$$\mathcal{P}_{ij}\Lambda_s = \Lambda_{s'}R_{ij} \tag{8}$$

of the type which we have just considered. Hence, if $s' < s$ (8) and therefore (7) is impossible , and if $s' \geq s$, each equation (8) has $s' - s + 1$ linearly independent solutions, and the number of parameters in the general solution (7) is consequently

$$\alpha\alpha'(s' - s + 1).$$

In particular, when $s' = s$ and $\alpha' = \alpha$, $\mathcal{P}_{ij}$ and R_{ij} must be of the form

$$\mathcal{P}_{ij} = \phi_{ij} I_{s+1} \quad ; \quad R_{ij} = \phi_{ij} I_s \qquad (i, j = 1, 2, \ldots, \alpha)$$

where ϕ_{ij} are α^2 constants (Corollary to theorem I).

Introducing a matrix

$$\Phi = [\phi_{ij}]$$

we can write the result as

$$\mathcal{P} = [\phi_{ij} I_{s+1}] = \Phi \times I_{s+1}$$

$$R = [\phi_{ij} I_s] = \Phi \times I_s$$

using the familiar notation for Zehfuss (or Kronecker) matrices. The solutions P and Q are nonsingular , if and only if Φ is nonsingular.

This enables us at last to obtain the complete solu-tion of(i)(§1,(9)):

THEOREM II

The general solution of

(9) $$\mathcal{P}\Big[\sum_{i=1}^{p}(\alpha_i \Lambda_{m_i})\Big] = \Big[\sum_{i=1}^{p}(\alpha_i \Lambda_{m_i})\Big] R$$

is of the form

$$\mathcal{P} = \begin{bmatrix} \mathcal{P}_{11} & \cdot & \cdot & \cdot \\ \mathcal{P}_{21} & \mathcal{P}_{22} & \cdot & \\ \cdots & \cdots & \cdots & \\ \mathcal{P}_{p1} & \mathcal{P}_{p2} & \cdots & \mathcal{P}_{pp} \end{bmatrix} \qquad R = \begin{bmatrix} R_{11} & \cdot & \cdot & \cdot \\ R_{21} & R_{22} & \cdot & \\ \cdots & \cdots & & \\ R_{p1} & R_{p2} & \cdots & R_{pp} \end{bmatrix}$$

involving in all

(10) $$t_1 = \sum_{i \geq j} \alpha_i \alpha_j (m_i - m_j + 1)$$

arbitrary parameters.

The proof is obvious from the preceding results.For, again,we partition P and R in a suitable way so as to resolve(9)into the equations

$$\mathcal{P}_{ij}(\alpha_j \Lambda_{m_j}) = (\alpha_i \Lambda_{m_i}) R_{ij}$$

which have already been discussed.When i < j, we have $m_i < m_j$ (§1,(5,1))and therefore $\mathcal{P}_{ij} = 0$; $R_{ij} = 0$.Further, $\mathcal{P}_{ii}$ and R_{ii} are of the form $\mathcal{P}_{ii} = \phi_i \times I_{m_i+1}$; $R_{ii} = \phi_i \times I_{m_i}$ and P and R are non-singular if and only if $\phi_1, \phi_2, \ldots, \phi_p$ are non-singu-lar.

The solution of(IX),(§1,(9)),viz.,

$$\mathcal{P}\Big[\sum_{j=1}^{q}(\beta_j \Lambda'_{n_j})\Big] = \Big[\sum_{j=1}^{q}(\beta_j \Lambda'_{n_j})\Big] R$$

does not involve any new difficulties,since the trans-position of this equation leads us back to the pre-vious case.The number of parameters in the general solution of(IX)is therefore

(11) $$t_q = \sum_{j \geq i} \beta_i \beta_j (n_j - n_i + 1)$$

We shall now consider(IV), viz.,

$$\mathcal{P}\Gamma_2 = \Gamma_1 R \qquad \text{or}$$

(12) $$\mathcal{P}\Gamma_2 = \left[\sum_{i=1}^{p}(\alpha_i \Lambda_{m_i})\right] R$$

where $$\Gamma_2 = \lambda I_k + \mathcal{M}_k$$

is a nonsingular pencil of type k×k. As before, (12) resolves into a number of equations of the kind

(13) $$\mathcal{P}_i \Gamma_2 = \Lambda_{s_i} R_i \qquad (i = 1, 2, \ldots, \sum_j \alpha_j)$$

where s_i stands for m_1 (α_1 times repeated), m_2 (α_2 times repeated) etc.. It is easy to see that (13) is impossible because premultiplying it by u_{s_i} we get

$$(u_{s_i} \mathcal{P}_i)\,\Gamma_2 = 0$$

Since Γ_2 is a non-singular pencil, it follows that

$$u_{s_i} \mathcal{P}_i = 0$$

and hence by lemma II(§1)

$$\mathcal{P}_i = 0 \quad .$$

Equation(13) now becomes

$$0 = \Lambda_{s_i} R_i$$

which entails

$$R_i = 0$$

the columns of Λ_{s_i} being independent. The solution of (IV) therefore contributes no parameters, i.e. .

(14) $$t_4 = 0$$

We get a different result, however, when Γ_2 and Γ_1 are inter--changed, as is the case in (II), viz.:

$$\mathcal{P}\Gamma_1 = \Gamma_2 R$$

or

(15) $$\mathcal{P}\left[\sum_i (\alpha_i \Lambda_{m_i})\right] = (\lambda I_k + \mathcal{M}_k) R \,,$$

which may be resolved into $\sum_j \alpha_j$ equations of this kind:

(16) $\mathcal{P}_i \Lambda_{s_i} = \mathcal{P}_i \left\{ \lambda \begin{bmatrix} I_{s_i} \\ \cdot \end{bmatrix} + \begin{bmatrix} \cdot \\ I_{s_i} \end{bmatrix} \right\} = (\lambda I_k + \mathcal{M}_k) R_i \quad (i = 1, 2, \ldots, \sum_j \alpha_j)$,

where s_i takes the values $m_1, m_2, \ldots, m_p$, m_i being re--peated α_i times. Hence comparing coefficients of λ we get:

(16a) $$\mathcal{P} \begin{bmatrix} I_{s_i} \\ \cdot \end{bmatrix} = R$$

(16b) $$\mathcal{P} \begin{bmatrix} \cdot \\ I_{s_i} \end{bmatrix} = \mathcal{M}_k R$$

where, for simplicity, we have dropped the suffixes of the matrices P and R. Let

$$\mathcal{P} = [p_0, p_1, \ldots, p_{s_i}]$$

$$R = [r_1, r_2, \ldots, r_{s_i}]$$

introducing column vectors $p_0, p_1, \ldots, r_1, r_2, \ldots$ for the columns of P and R each having k rows. It is easy to see that postmultiplication by $\begin{bmatrix} I_{s_i} \\ \cdot \end{bmatrix}$ or $\begin{bmatrix} \cdot \\ I_{s_i} \end{bmatrix}$ has the effect of cancelling the last or the first row. Hence (16a) and (16b) become

$$[p_0, p_1, \ldots, p_{s_i - 1}] = [r_1, r_2, \ldots, r_{s_i}]$$

$$[p_1, p_2, \ldots, p_{s_i}] = [\mathcal{M}_k r_1, \mathcal{M}_k r_2, \ldots, \mathcal{M}_k r_{s_i}]$$

and by eliminating $r_1, r_2, \ldots, r_{s_i}$

$$[p_1, p_2, \ldots, p_{s_i}] = [\mathcal{M}_k p_0, \mathcal{M}_k p_1, \ldots, \mathcal{M}_k p_{s_i - 1}] \quad \text{whence}$$

$$p_1 = \mathcal{M}_k p_0 \; ; \; p_2 = \mathcal{M}_k p_1 = \mathcal{M}_k^2 p_2 \; ; \ldots \; ; \; p_\nu = \mathcal{M}_k^\nu p_0 \; ; \ldots \quad (\nu = 0, 1, \ldots, s_i)$$

$$r_1 = p_0 \; ; \; r_2 = \mathcal{M}_k p_0 \; ; \ldots ; \; r_\mu = \mathcal{M}_k^{\mu - 1} p_0 \; ; \ldots \quad (\mu = 1, 2, \ldots, s_i)$$

where the k elements of the vector p_0 remain arbitrary.

we have therefore:

THEOREM(III)

The general solution of the equation

(16) $$P_i \Lambda_{s_i} = (\lambda I_k + \mathcal{M}_k) R_i$$

is of the form

$$P_i = [p_0, \mathcal{M}_k p_0, \mathcal{M}_k^2 p_0, \ldots, \mathcal{M}_k^{s_i} p_0]$$

$$R_i = [p_0, \mathcal{M}_k p_0, \mathcal{M}_k^2 p_0, \ldots, \mathcal{M}_k^{s_i-1} p_0]$$

involving k arbitrary parameters,namely the elements of p_0.

Since(15),or(II),is equivalent to $\sum_i \alpha_i$ equations of type(16),we have at once

(17) $$t_2 = k \sum_i \alpha_i$$

As before,we see that(VI) (§1,(9)),i.e.,

$$\mathcal{P} \Gamma_2 = \Gamma_3 R$$

and (VIII),i.e.

$$\mathcal{P} \Gamma_3 = \Gamma_2 R$$

are merely different froms of(II)and(IV)to which they are reduced by transposition. Γ_3 is then replaced by a pencil of the same type as Γ_1 (§1,(5,1)and(5,3))and the order of the factors is reversed while Γ_2 and Γ_2' play the same rôle since the only property we have used,was that Γ_2 was a non-singular pencil of type k×k.

By analogy,we obtain therefore

(18) $$t_6 = k \sum_{j=1}^{q} \beta_j \cdot$$ and

(19) $$t_8 = 0.$$

We shall now show that (VII),i.e.

$$\mathcal{P} \Gamma_3 = \Gamma_1 R$$,or

(20) $$\mathcal{P}[\sum_j (\beta_j \Lambda'_{n_j})] = [\sum_i (\alpha_i \Lambda_{m_i})] R$$

is satisfied only in the trivial case $\mathcal{P} = 0$; $R = 0$ .For

(20) splits up into a number of equations

(21) $$P_{ij}\Lambda'_{n_j} = \Lambda_{m_i} R_{ij}$$

which after premultiplication by u_{m_i} yield

$$(u_{m_i} P_{ij})\Lambda'_{n_j} = 0$$

whence

$$u_{m_i} P_{ij} = 0$$

the <u>rows</u> of Λ'_{n_j} being linearly independent (§2, lemma II).

Hence as before, it follows that

$$P_{ij} = 0 \quad \text{and} \quad R_{ij} = 0 \ .$$

We therefore have

(22) $$t_7 = 0$$

It remains to solve (III) which we write as

(23) $$\mathcal{P}\left[\sum_{i=1}^{p}(\alpha_i \Lambda_{m_i})\right] = \left[\sum_{j=1}^{q}(\beta_j \Lambda'_{n_j})\right] R$$

and which reduces to $(\sum_i \alpha_i)\cdot(\sum_j \beta_j)$ partial equations

(24) $$P_{ij}\Lambda_{m_i} = \Lambda'_{n_j} R_{ij}$$

each occurring $\alpha_i \beta_j$ times. Substituting for Λ_{m_i} and Λ'_{n_j} from §2, (1) we get

$$P_{ij}\left\{\lambda\begin{bmatrix} I_{m_i} \\ \cdot \end{bmatrix} + \begin{bmatrix} \cdot \\ I_{m_i} \end{bmatrix}\right\} = \left\{\lambda\,[I_{n_j}\ \cdot] + [\cdot\ I_{n_j}]\right\} R_{ij}, \text{ i.e.,}$$

(25) $$P_{ij}\begin{bmatrix} I_{m_i} \\ \cdot \end{bmatrix} = [I_{n_j}\ \cdot]\, R_{ij}$$

(26) $$P_{ij}\begin{bmatrix} \cdot \\ I_{m_i} \end{bmatrix} = [\cdot\ I_{n_j}]\, R_{ij}$$

Let

$$P_{ij} = [p_{\mu\nu}] \qquad \begin{pmatrix} \mu = 1, 2, \ldots, n_j \\ \nu = 1, 2, \ldots, m_i+1 \end{pmatrix}$$

$$R_{ij} = [r_{\rho\sigma}] \qquad \begin{pmatrix} \rho = 1, 2, \ldots, n_j+1 \\ \sigma = 1, 2, \ldots, m_i \end{pmatrix}$$

Then (25) and (26) become

(27) $$p_{\mu\nu} = r_{\mu\nu} \qquad \begin{pmatrix} \mu = 1, 2, \ldots, n_j \\ \nu = 1, 2, \ldots, m_i \end{pmatrix}$$

(28) $$p_{\mu,\nu+1} = r_{\mu+1,\nu} \qquad \begin{pmatrix} \mu = 1, 2, \ldots, n_j \\ \nu = 1, 2, \ldots, m_i \end{pmatrix}$$

From these two equations we infer

$$p_{\mu,\nu+1} = p_{\mu+1,\nu}$$

or, replacing ν by $\nu-1$ and iterating the equation :

$$p_{\mu\nu} = p_{\mu+1,\nu-1} = p_{\mu+2,\nu-2} = \cdots = p_{\mu+h,\nu-h} = \cdots = \phi_{\mu+\nu-2},$$

say, i.e. the value of $p_{\mu\nu}$ depends only on the sum of the suffixes, and similarly

$$r_{\mu,\nu} = r_{\mu+1,\nu-1} = r_{\mu+2,\nu-2} = \cdots = p_{\mu\nu} = \phi_{\mu+\nu-2}$$

We have therefore proved the following theorem:

THEOREM IV:

The general solution of $\mathcal{P}_{ij}\Lambda_{m_i} = \Lambda'_{n_j} R_{ij}$ is of the form

$$\mathcal{P}_{ij} = \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{m_i} \\ \phi_1 & \phi_2 & \cdots & \phi_{m_i+1} \\ & & & \\ \phi_{n_j-1} & \phi_{n_j} & \cdots & \phi_{m_i+n_j-1} \end{bmatrix}_{n_j, m_i+1} ; \quad R = \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{n_j-1} \\ \phi_{-1} & \phi_2 & \cdots & \phi_{n_j} \\ & & & \\ \phi_{n_j} & \phi_{n_j+1} & \cdots & \phi_{m_i+n_j-1} \end{bmatrix}_{n_j+1, m_i}.$$

involving m_i+n_j parameters $\phi_0, \phi_1, \ldots, \phi_{m_i+n_j-1}$.

The number of parameters that occur in the general solution of (III) is therefore

(29) $$t_3 = \sum_{i,j} \alpha_i \beta_j (m_i + n_j) \qquad (i = 1, 2, \ldots, p\ ;\ j = 1, 2, \ldots, q)$$

We now add up the subtotals $t_1, t_2, t_3, \ldots, t_8, t_9$ as given in § 1, (10) and § 3, (10), (11), (14), (17), (18), (19), (22), and (29), and we find that the number of linearly inde-pendent solutions of

(30) $$\mathcal{P}\Gamma = \Gamma R$$

is given by

$$t = \sum_{h=1}^{9} t_h = \tau + k\sum_i \alpha_i + k\sum_j \beta_j + \sum_{i \geq j} \alpha_i \alpha_j (m_i - m_j + 1) + \sum_{j \geq i} \beta_i \beta_j (n_j - n_i + 1) + \sum_{i,j} \alpha_i \beta_j (m_i + n_j),$$
$$(i = 1, 2, \ldots, p\ ;\ j = 1, 2, \ldots, q).$$

Moreover, a method has been given for actually ob-taining all matrices P and R satisfying (30). Since

$$t_4 = t_7 = t_8 = 0 ,$$

it follows that in the scheme §1,(9)

$$\mathcal{P}_{12}=0\,,\ R_{12}=0\ ;\ \mathcal{P}_{13}=0\,,\ R_{13}=0\ ;\ \mathcal{P}_{23}=0\,,\ R_{23}=0\,,$$

so that the general solution is of the form

$$\mathcal{P}=\begin{bmatrix}\mathcal{P}_{11} & \cdot & \cdot\\ \mathcal{P}_{21} & \mathcal{P}_{22} & \cdot\\ \mathcal{P}_{31} & \mathcal{P}_{32} & \mathcal{P}_{33}\end{bmatrix}\qquad R=\begin{bmatrix}R_{11} & \cdot & \cdot\\ R_{21} & R_{22} & \cdot\\ R_{31} & R_{32} & R_{33}\end{bmatrix}.$$

Further, P and R are non-singular if and only if the matrices $\mathcal{P}_{11}$, $\mathcal{P}_{22}$, $\mathcal{P}_{33}$ and R_{11}, R_{22}, R_{33} are non-sin-gular, and we have already found the necessary and sufficient conditions that those matrices should be non-singular (see p. 11.).

PART III

ON SINGULAR PENCILS OF ZEHFUSS, COMPOUND, AND SCHÄFLIAN MATRICES.

ON SINGULAR PENCILS OF ZEHFUSS, COMPOUND, AND INDUCED MATRICES.

§1.

Introduction.

In this paper the canonical form of matrix pencils will be disscussed which are based on a pair of direct products (Zehfuss matrices), compound, or Schäflian (induced) matrices derived from given pencils whose canonical forms are known.

When all the pencils concerned are non singular (i.e. when their determinants do not vanish identically), the problem is equivalent to finding the elementary divisors of the pencil. This has been solved by A.C. Aitken (ref.1), D.E. Littlewood (ref.2), and W.E. Roth (ref.3). In the singular case, however, the so-called minimal indices or Kronecker Invariants have to be determined in addition to the elementary divisors (ref.4, Chapter IX). The answer to this question forms the subject of the following investigation.

The method employed is that of the principle of vector chains which was first used in this connection by H.W. Turnbull (ref.5).

Let $\rho A+\sigma B$ be a pencils of type $m\times n$ i.e. with m rows and n columns. It is then possible to find two non-singular constant matrices P of degree m and R of degree n such that

(1a) $$P(\rho A+\sigma B)R=\begin{bmatrix} O_{ef} & O \\ O & \rho A_1+\sigma B_1 \end{bmatrix}$$

there being $e\geq o$ zero rows and $f\geq o$ zero columns and

(1b) $$(\rho A_1+\sigma B_1)=\mathrm{diag}(L_{n_1},L_{n_2}\dots L_{n_a},M_{m_1},M_{m_2}\dots M_{m_b},N_{p_1},N_{p_2}\dots N_{p_c},Z)$$

where the symbols on the right hand side have the following meaning

---

ref.6, p.94.

(2a) $$L_i = \rho F_i + \sigma G_i = \begin{bmatrix} \rho & \sigma & \cdot & \cdot & \cdot \\ \cdot & \rho & \sigma & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \rho & \sigma \end{bmatrix}_{i,\,i+1}$$

is a pencil of i rows and $i+1$ columns, e.g.

$$L_4 = \begin{bmatrix} \rho & \sigma & \cdot & \cdot & \cdot \\ \cdot & \rho & \sigma & \cdot & \cdot \\ \cdot & \cdot & \rho & \sigma & \cdot \\ \cdot & \cdot & \cdot & \rho & \sigma \end{bmatrix}$$

and:

(2b) $$M_j = \rho G'_j + \sigma F'_j = \begin{bmatrix} \sigma & \cdot & \cdot & \cdot \\ \rho & \sigma & \cdot & \cdot \\ \cdot & \rho & & \\ & & \rho & \sigma \\ \cdot & \cdot & \cdot & \rho \end{bmatrix}_{j+1,\,j}$$

has $j+1$ rows and j columns. N_k is defined as

(3) $$N_k = \rho I_k + \sigma U_k = \begin{bmatrix} \rho & \sigma & & \\ \cdot & \rho & \sigma & \\ & & \rho & \sigma \\ & & \cdot & \rho \end{bmatrix}_{k,\,k}$$

where I_k and U_k are the unit matrix and the auxiliary matrix of degree k.* N_k represents the "infinite" latent roots while L_i and M_j correspond to the linear relations between rows and between columns of the pencil. Finally,

$$Z = \rho C + \sigma D$$

is a non-singular pencil for which $|D| \neq 0$. It will be noticed that this canonical form is slightly different from that given loc. cit. But it is easy to see that the two forms are equivalent; for apart from rearranging the submatrices we have

$$L_i = \Lambda'_i \text{ (loc. cit. §1, (3))}$$

$$M_j = J_{j+1} \Lambda_j J_j \text{ , where}$$

(4) $$J_j = \begin{bmatrix} \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & 1 & \cdot \\ & & & \\ 1 & \cdot & \cdot & \cdot \end{bmatrix}_{j,\,j}$$

and N_p is replaced by $N'_p = J_p N_p J_p$.

In order to state the result (1) more conveniently, we introduce the "direct sum" of matrices, viz.

* Ref. 4, p 62

$$A \dotplus B = \begin{bmatrix} A & \cdot \\ \cdot & B \end{bmatrix}$$

and in general

$$A_1 \dotplus A_2 \dotplus \ldots \dotplus A_h = \dot{\sum_{i=1}^{h}} A_i = \mathrm{diag}(A_1, A_2, \ldots, A_h) \quad .$$

We shall also use the abbreviation

$$A \dotplus A = 2A \;, A \dotplus A \dotplus A = 3A \;, \text{etc.}$$

Further, for the sake of symmetry we introduce the two symbols

$$L_o \text{ and } M_o$$

to the following effect: if A is a matrix with $r(>o)$ rows and s $(>o)$ columns, we define

(5a) $\quad L_o \dotplus A = [\cdot \; A]$; $2L_o \dotplus A = [\cdot \; \cdot \; A]$, etc.

(5b) $\quad M_o \dotplus A = \begin{bmatrix} \cdot \\ A \end{bmatrix}$; $2M_o \dotplus A = \begin{bmatrix} \cdot \\ \cdot \\ A \end{bmatrix}$, etc.

i.e. the term L_o $[M_o]$ means that A has to be augmented by a zero column [row]. L_o and M_o are, of course, not proper matrices.

In this notation we can write (1a) and (1b) as

(6a) $\quad P(\rho A+\sigma B)R = eL_o \dotplus fM_o \dotplus (\rho A_1+\sigma B_1)$

(6b) $\quad \rho A_1+\sigma B_1 = \dot{\sum_i} L_{n_i} \dotplus \dot{\sum_j} M_{m_j} \dotplus \dot{\sum_k} N_{p_k} \dotplus Z$

or more briefly

(7) $\quad \rho A+\sigma B \sim \dot{\sum_i} L_{n_i} \dotplus \dot{\sum_j} M_{m_j} \dotplus \dot{\sum_k} N_{p_k} \dotplus Z$

if we include the zero indices, if any, among the n_i and m_j .

According to the general theory, Z may be transformed into the classical Weierstrassian form, viz.

$$Z = \dot{\sum_r} W_{h_r}(\alpha_r) \dotplus \dot{\sum_\ell} Q_{g_\ell}$$

where

(8) $$W_h(\alpha) = (\alpha\rho+\sigma)I_h+\rho U'_h = \begin{bmatrix} \alpha\rho+\sigma & & & \\ \rho & \alpha\rho+\sigma & & \\ & \ddots & \ddots & \\ & & \rho & \alpha\rho+\sigma \end{bmatrix}_{h,h}$$

and

(9) $$Q_g = W_g(o) = \sigma I_g+\rho U'_g = \begin{bmatrix} \sigma & & \\ \rho & \sigma & \\ & \ddots & \ddots \\ & \rho & \sigma \end{bmatrix}_{g,g}$$

The quantities $\alpha_1, \alpha_2, \cdots$ are non-zero, but not necessarily distinct. Substituting in(7) we obtain

(10) $$\rho A+\sigma B \sim \sum_i L_{n_i} \dot{+} \sum_j M_{m_j} \dot{+} \sum_k N_{p_k} \dot{+} \sum_h Q_{g_h} \dot{+} \sum_r W_{h_r}(\alpha_r)$$

where n_i and m_j are the two sets of minimal indices of the pencil $\rho A+\sigma B$ referring to the columns and rows respectively. They are the minimal degrees in ρ and σ of the column and row vectors which annihilate the pencil, and they may be positive or zero. It is no loss of generality, however, to assume that no zeros occur among them i.e. that e=f=o in(1a)[1]. But later on, when dealing whith composite pencils we shall see that zero values for n_i and m_j cannot be avoided.

It should be noted that direct summation of matrices is commutative if we do not distinguish between equivalent pencils[2]; thus

$$\rho(C_1 \dot{+} C_2)+\sigma(D_1 \dot{+} D_2) \sim \rho(C_2 \dot{+} C_1)+\sigma(D_2 \dot{+} D_1)$$

i.e. those two pencils have the same canonical form.

§2.

Vector Chains and Canonical Form of a Pencil.

In two papers by H.W.Turnbull[3] and the author[3] it has

---

[1] ref 6, p.93
[2] ref 2, §3
[3] ref 5 and ref 6.

been shown how a pencil can be brought in to its canonical shape by successively forming certain vector chains. We shall here use an extension of this principle which has also been hinted at by Prof.Turnbull.Let

$$\rho A+\sigma B$$

be a pencil with m rows and n columns and consider the equation

(Ia) $$(\rho A+\sigma B)R= \bar{R}\cdot L_{i-1}$$

where L_{i-1} is defined in §1,(2), R is of type $n\times i$ and $\bar{R}$ is of type $m\times(i-1)$. Throughout ~~in~~ this section the columns of R and $\bar{R}$ will be supposed to be linearly independent. Put

(1) $R = [r_1, r_2,\dots,r_i]$ and $\bar{R} = [\bar{r}_1,\bar{r}_2,\dots,\bar{r}_{i-1}]$
where the r's and the $\bar{r}$'s are column vectors of n and m elements respectively. Writing (Ia) in full, we have

(2) $$(\rho A+\sigma B)\cdot[r_1,r_2,\dots,r_i] =[\bar{r}_1,\bar{r}_2,\dots,\bar{r}_{i-1}]\begin{bmatrix}\rho & \sigma & & & \\ \cdot & \rho & \sigma & & \\ & & & & \\ \cdot & & & \rho & \sigma\end{bmatrix}_{i-1,i}$$

Hence

$$[\rho Ar_1+\sigma Br_1,\rho Ar_2+\sigma Br_2,\dots\rho Ar_i+\sigma Br_i] = [\rho\bar{r}_1,\sigma\bar{r}_1+\rho\bar{r}_2,\dots,\sigma\bar{r}_{i-1}]$$

~~and~~ that is

(3) $Ar_1=\bar{r}_1$; $Ar_2= \bar{r}_2$; .... ;$Ar_{i-1}=\bar{r}_{i-1}$; $Ar_i=0$

(4) $Br_1=0$; $Br_2=\bar{r}_1$;...................;$Br_{i-1}=\bar{r}_{i-2}$; $Br_i =\bar{r}_{i-1}$

From (3) and (4) we derive the vector chain*

(Ib) $0=Br_1$; $Ar_1 =Br_2$; $Ar_2 =Br_3$;....;$Ar_{i-1} =Br_i$;$Ar_i =0$, of "type L"

---

* In the two papers cited at the beginning of this paragraph row vectors are used instead of column vectors which we here prefer merely for technical reasons. The fact that either row or column vectors have to be distinguished, is certainly a disantvantage which is, however shared by most of the other theories.

On the other hand, suppose we have found i vectors $r_1, r_2, \ldots, r_i$ satisfying (Ib); we can then define the vectors $\bar{r}_1, \bar{r}_2, \ldots \bar{r}_{i-1}$ by (3) or (4) and get back to (Ia). Hence (Ia) and (Ib) are fully equivalent. In the same manner we can deal with the other submatrices $L_j, M_k, Q_h,$ and $W_h(\alpha)$ that occur in the general canonical form §1(10).

Let

(IIa) $$(\rho A+\sigma B)R = \bar{R}M_j \quad ,$$

where R must now be of type $(n\times j)$ and $\bar{R}$ of type $(m\times(j-1))$

Put

(5) $$R=[r_1, r_2, \ldots r_j] \text{ and } \bar{R}=[\bar{r}_0, \bar{r}_1, \ldots \bar{r}_j]$$

Substituting (5) and §1, (2) in (IIa) we get

$$(\rho A+\sigma B)\cdot[r_1, r_2, \ldots, r_j] = [\bar{r}_0, \bar{r}_1, \ldots, \bar{r}_j]\begin{bmatrix} \sigma & \cdot & & \\ \rho & \sigma & & \\ \cdot & \rho & & \\ & & \rho & \sigma \\ & & \cdot & \rho \end{bmatrix}_{j-1, j}.$$

Hence

(6) $$Ar_1=\bar{r}_1;\ Ar_2=\bar{r}_2; \ldots\ldots; Ar_{j-1}=\bar{r}_{j-1}; Ar_j=\bar{r}_j$$

(7) $$Br_1=\bar{r}_0; Br_2=\bar{r}_1; \ldots\ldots; Br_{j-1}=\bar{r}_{j-2}; Br_j=\bar{r}_{j-1}$$

whence we derive the vector chain

(IIb) $$0\neq Br_1; Ar_1=Br_2; \ldots\ldots; Ar_{j-1}=Br_j;\ Ar_j\neq 0$$, of "type M".

Next, consider

(IIIa) $$(\rho A+\sigma B)\cdot R = \bar{R}\cdot N_k$$

where R is of type $(n\times k)$ and $\bar{R}$ of type $(m\times k)$. Put

(8) $$R=[r_1, r_2, \ldots, r_k] \text{ and } \bar{R}=[\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_k]$$

and substitute (8) and §1, (3) in (IIIa)

$$(\rho A+\sigma B)[r_1, r_2, \ldots, r_k] = [\bar{r}_1, \bar{r}_2, \ldots \bar{r}_k]\begin{bmatrix} \rho & \sigma & \cdot & \cdot & \cdot \\ \cdot & \rho & \sigma & \cdot & \\ & & & & \\ \cdot & & & \rho & \sigma \\ & & & & \rho \end{bmatrix}_{k,k}$$

Hence

(9) $Ar_1=\bar{r}_1;\ Ar_2=\bar{r}_2;\ \ldots\ldots;Ar_{k-1}=\bar{r}_{k-1};\ Ar_k=\bar{r}_k$.

(1o) $Br_1=o;\ Br_2=\bar{r}_1;\ \ldots\ldots;Br_{k-1}=\bar{r}_{k-2};\ Br_k=\bar{r}_{k-1}$.
Equation(9)and(1o)yield the chain

(IIIb) $o=Br_1;\ Ar_1=Br_2;\ \ldots;Ar_{k-1}=Br_k;\ Ar_k\neq o$, of "type N".
Finally,let

(IVa) $$(\rho A+\sigma B)\cdot R = \bar{R}\cdot Q_g$$

where r is of type(n×g) and $\bar{R}$ of type (m×g).Put

(11) $$R=[r_1,r_2,\ldots r_g],\bar{R}=[\bar{r}_1,\bar{r}_2,\ldots\bar{r}_g]$$

On substituting(11)and §1(9)in(IVa)we obtain

$$(\rho A+\sigma B)\,[r_1,r_2,\ldots r_g]=[\bar{r}_1,\bar{r}_2,\ldots\bar{r}_g]\begin{bmatrix}\sigma & \cdot & \cdot & \\ \rho & \sigma & \cdot & \\ \cdot & \rho & \sigma & \\ & & & \\ & & \rho & \sigma\end{bmatrix}_{g,g}$$

Hence comparing coefficients of ρ and σ we get

(12) $Ar_1=\bar{r}_2;\ Ar_2=\bar{r}_3;\ \ldots\ldots;Ar_{g-1}=\bar{r}_g;\ Ar_g=o$

(13) $Br_1=\bar{r}_1;\ Br_2=\bar{r}_2;\ \ldots\ldots;Br_{g-1}=\bar{r}_{g-1};Br_g=\bar{r}_g$.

this gives rise to the vector chain

(IVb) $o\neq Br_1;\ Ar_1=Br_2;\ Ar_2=Br_3;\ \ldots;\ Ar_{g-1}=Br_g;\ Ar_g=o$, of "type Q"

It will facilitate the work if we write the chains(Ib) (IIb),(IIIb),and(IVb) in the following standard form which enables us to deal with those types more uniformly.Let

(14) $\bar{r}_o=Br_1;\ \bar{r}_1=Ar_1=Br_2;\ \ldots;\bar{r}_{p-1}=Ar_{p-1}=Br_p;\bar{r}_p=Ar_p$

According as r_o and r_p are zero or not the vector chain(14)is of one of those four types,viz.of

$$
(15)\qquad
\begin{array}{lll}
\text{type } L \text{ if } \bar{r}_o = o;\ \bar{r}_p = o & & (Ib) \\
\text{type } N \text{ if } \bar{r}_o = o;\ \bar{r}_p \neq o & & (IIIb) \\
\text{type } Q \text{ if } \bar{r}_o \neq o;\ \bar{r}_p = o & & (IVb) \\
\text{type } M \text{ if } \bar{r}_o \neq o;\ \bar{r}_p \neq o & & (IIb)
\end{array}
$$

They correspond to the submatrices L_{p-1}, M_p, N_p, Q_p.The number p is called the <u>length</u> of the chain(4). R always has the columns $r_1,r_2,\dots,r_p$ while the columns of $\bar{R}$ are the non zero vectors out of the set $\bar{r}_o,\bar{r}_1,\dots,\bar{r}_{p-1},\bar{r}_p$.

We add a simple example in order to show the principal idea of our method.Supposing we have determined three pairs of matrices $R_1,\bar{R}_1;R_2,\bar{R}_2;R_3,\bar{R}_3$ which satisfy(Ia),(IIa), and (IVa)resp.We can then comprehend these equations in

$$(\rho A+\sigma B)\,[R_1,R_2,R_3]=[\bar{R}_1,\bar{R}_2,\bar{R}_3](L_i\dotplus M_j\dotplus Q_g)$$

Now, if the numbers of columns in R_1,R_2,R_3 and $\bar{R}_1,\bar{R}_2,\bar{R}_3$ are such that the matrices R_1,R_2,R_3 and $\bar{R}_1,\bar{R}_2,\bar{R}_3$ are <u>square</u> and <u>non-singular</u>, then we have proved that the canonical form of $\rho A+\sigma B$ is $L_i\dotplus M_j\dotplus Q_g$ Our plan will therefore be to construct as many matrices $R_1,R_2,\dots;\ \bar{R}_1,\bar{R}_2,\dots$ as are necessary to build up two non-singular matrices $[R_1,R_2,\dots]$ and $[\bar{R}_1,\bar{R}_2,\dots]$.

§3.

Preliminary Remarks about Direct Products.

If $A = [a_{ij}]$ is a matrix of type $m\times n$ and $B = [b_{kl}]$ a matrix of type $p\times q$, then their direct product

$$(1) \qquad A\times B = \begin{bmatrix} a_{11}B \dots a_{1n}B \\ a_{m1}B \dots a_{mn}B \end{bmatrix}$$

is a matrix of type $(mp\times nq)$. It can be defined for any two matrices. Direct multiplication obeys the associative and the distributive law, viz.

$$(A\times B)\times C = A\times(B\times C) = A\times B\times C$$

$$A\times(B+C) = (A\times B) + (A\times C)$$

$$(A+B)\times C = (A\times C) + (B\times C)$$

as can easily be verified.

There is also a distributive law connecting <u>direct addition</u> and direct multiplication, viz.

$$(2) \qquad (A_1 \dotplus A_2) \times B = (A_1\times B) \dotplus (A_2\times B)$$

Again,

$$(3) \qquad (A\times B)' = (A'\times B')$$

Both these rules readily follow from the definition(1). It should be noted that no simple relation like(2) exists when the <u>second</u> factor is a direct sum.

The most important property of direct multiplication is the multiplicative law

$$(4) \qquad (R\times S)\cdot(A\times B) = ((R\cdot A)\times(S\cdot B)) \quad ;$$

if A and B are square matrices of non-singular determinants, so is $A\times B$. This follows from(4) by putting

$R = A^{-1}$ and $S = B^{-1}$ .The right hand side then becomes a unit matrix and neither of the factors on the left can be of zero determinant.

Consider the transformation of two sets of variables:

(5a)
$$\rho_i = \sum_{j=1}^{n} a_{ij}\, \xi_j \qquad (i = 1, 2, \ldots, m)$$
$$\sigma_k = \sum_{h=1}^{q} b_{kh}\, \eta_h \qquad (k = 1, 2, \ldots, p)$$

or in vector form

(5b)
$$r = A\,x$$
$$s = B\,y$$

introducing column vectors

$$r = \{\rho_1, \rho_2, \ldots, \rho_m\} \qquad x = \{\xi_1, \xi_2, \ldots, \xi_n\}$$
$$s = \{\sigma_1, \sigma_2, \ldots, \sigma_p\} \qquad y = \{\eta_1, \eta_2, \ldots, \eta_q\}$$

On account of(5a) the m·p products $\rho_i\,\sigma_k$ are linear functions of the np products $\xi_j \cdot \eta_h$ the matrix of the coefficients being $A \times B$;this follows at once from (4),because

$$(r \times s) = (Ax \times By) = (A \times B)(x \times y)$$

and the elements of r×s and of(x×y) obviously are just the products $\rho_i\,\sigma_k$ and $\xi_j\;\eta_h$

Direct multiplication is not commutative,but we shall obtain a substitute for this property through

THEOREM I

THEOREM I

The two products $A\times B$ and $B\times A$ are related to each other by an identity

(6) $$Q(A\times B)P^{-1} = (B\times A)$$

where P and Q are permutation matrices which depend only on the types of A and B and not on their elements.

Proof: Apart from the order the vectors $(x\times y)$ and $(y\times x)$ contain the same elements, viz. the nq products $\xi_j \eta_h$. We can therefore find a permutation matrix P of degree nq such that

(7) $$(y\times x) = P(x\times y) \quad \text{and similarly}$$

$$(s\times r) = Q(r\times s)$$

where Q is a permutation matrix of degree mr. Evidently ~~these matrices~~ P and Q do not depend on the elements of x,y,u,v but only on the numbers m,n,p,q. By(5a) and(4) we have

$$(r\times s) = (A\times B)(x\times y)$$

$$(s\times r) = (B\times A)(y\times x)$$

On pre-multiplying the first equation by Q and substituting(7) we get

$$(s\times r) = Q(A\times B)(x\times y) = (B\times A)P(x\times y) \ .$$

Since there is obviously no linear relation between the elements of $(x\times y)$ we obtain

$$Q(A\times B) = (B\times A)P$$

or

$$Q(A\times B)P^{-1}=(B\times A) \ , \quad \text{q. e. d.}$$

Let $r_1, r_2, \ldots, r_m$ be the columns of the non-singular matrix

$$(8) \qquad R = [r_1, r_2, \ldots, r_m] = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1m} \\ & & & \\ r_{m1} & r_{m2} & \ldots & r_{mm} \end{bmatrix}$$

i.e. suppose that the m vectors

$$r_j = \{r_{1j}, r_{2j}, \ldots r_{mj}\} \quad (j=1,2,\ldots,m)$$

are linearly independent. Similarly, let $s_1, s_2, \ldots, s_m$ be the columns of a non-singular matrix

$$(9) \qquad S = \begin{bmatrix} s_{11} & s_{12} & \ldots & s_{1n} \\ & & & \\ s_{n1} & s_{n2} & \ldots & s_{nn} \end{bmatrix} = [s_1, s_2, \ldots, s_n]$$

where

$$s_k = \{s_{1k}, s_{2k}, \ldots, s_{nk}\} \quad (k=1,2,\ldots n)$$

We shall then prove the

LEMMA I

If r_j $(j=1,2,\ldots,m)$ are m linearly independent column vectors of degree m, and s_k $(k=1\ 2,\ldots n)$ n linearly independent column vectors of degree n, then the mn vectors

$$(10) \qquad r_j \times s_k$$

of degree mn are linearly independent.

<u>Proof</u>: Supposing we had a linear relation

$$\sum_{j,k} c_{jk}(r_j \times s_k) = o$$

or written in components

$$\sum_{j,k} c_{jk}\, r_{ij}\, s_{hk} = o \quad (i=1,2.,m;h=1,2,..,n)$$

which is equivalent to the matrix equation

$$RCS' = o \quad ;$$

if we put

$$C = [c_{jk}] \quad (j=1,2,..,m;k=1,2,..,n).$$

Since R and S are non-singular, it follows that

$$C = o$$

i.e. the vectors(lo) are linearly independent.

We shall require another lemma which goes a little further than the preceding one:

<u>LEMMA II</u>

Let r_j $(j=1,2,...,m)$ and s_k $(k=1,2,...,n)$ be defined as in lemma I and let $E^{(k)}$ $(k=1,2,..,n)$ be a set of n non-singular matrices of degree m; then the mn vectors

$$(E^{(k)} r_j \times s_k)$$

of degree mn are linearly independent.

<u>Proof:</u> Since R in(8) is non-singular, we can write down the matrix equation

$$(12) \qquad E^{(k)}R = R(R^{-1}E^{(k)}R) = RA^{(k)}$$

where

$$(13) \qquad A^{(k)} = R^{-1}E^{(k)}R = (a_{hl}^{(k)})$$

is non-singular. By picking out the j^{th} column on either side of(12)we get by(8)and(13)

$$E^{(k)} r_j = \sum_h r_h a_{hj}^{(k)}$$

and after direct post multiplication by s_k

$$(E^{(k)} r_j \times s_k) = \sum_h a_{hj}^{(k)} \; (r_h \times s_k) \quad (j=1,2,.m; \; k=1,2..n)$$

Supposing these vectors were linearly dependent we could find mn constants z_{jk} not all zero such that

$$\sum_{j,k} z_{jk}(E^{(k)} r_j \times s_k) = \sum_{h,j,k} z_{jk} a_{hj}^{(k)} (r_h \times s_k) = o$$

According to lemma I it now follows that

$$\sum_j a_{hj}^{(k)} z_{jk} = o \quad \begin{pmatrix} h=1,2,..,m \\ k=1,2,..,n \end{pmatrix}$$

or

(14) $$A^{(k)} z_k = o \quad (k=1,2,..,m)$$

if we put

$$z_k = \{ z_{1k}, z_{2k}, \dots z_{mk} \}$$

But since $A^{(k)}$ is non-singular,we infer from(14)that

$$z_k = o \quad (k=1,2,\dots,m)$$

i.e.the vectors(11)are linearly independent.

Now,consider the pencils:

(15) $$H = \rho A + \sigma B \quad \text{of type } m \times n$$

and

(16) $$K = \rho C + \sigma D \quad \text{of type } p \times q$$

From these we derive a new pencil

(17) $$(H|K) = \rho(A\times C)+\sigma(B\times D) = \rho X+\sigma Y$$ *)

of type mp×nq.In the subsequent section we shall determine the minimal indices and the elementary divisors of(17) ~~if~~ when those of(15)and(16)are known.Here we will first establish some simple rules regarding (H|K)which will facilitate the investigation:

(18) $$(H|K)' = (H'|K')$$

(19) $$(H|K) \sim (K|H)$$

For by theorem I the products(A×C)and(B×D)can simultaneously be transformed into (C×A)and (D×B),i.e.we have

$$Q(A\times B)P^{-1} = (B\times A)$$

$$Q(C\times D)P^{-1} = (D\times C)$$

with the same matrices Q and P.On multiplying by ρ and σ and adding we get(19).

If $\bar{H}$ is equivalent to H and if $\bar{K}$ is equivalent to K,then

(20) $$(\bar{H}|\bar{K}) \sim (H|K)$$

For let $\bar{H} = P_1(\rho A+\sigma B)Q_1 = \rho P_1AQ_1+\sigma P_1BQ_1$

and $$\bar{K} = P_2(\rho C+\sigma D)Q_2 = \rho P_2CQ_2+\sigma P_2DQ_2$$

then

$$(P_1\times P_2)[\rho(A\times C)+\sigma(B\times D)]\cdot(Q_1\times Q_2) = \rho(P_1AQ_1\times P_2CQ_2)+\sigma(P_1BQ_1\times P_2DQ_2)$$

---

*) This is,of course,not equal to(H×K) which would be in general of the second degree with respect to ρ and σ.

or ~~by(15)~~ *in the notation (17)*:

$$(P_1 \times P_2)(H|K)(Q_1 \times Q_2) = (\bar{H}|\bar{K})$$

Next, let

$$H = H_1 \dotplus H_2 = \rho(A_1 \dotplus A_2) + \sigma(B_1 \dotplus B_2) \quad .$$

By(2), we obtain

(21) $$(H_1 \dotplus H_2 | K) = (H_1|K) \dotplus (H_2|K)$$

If, on the other hand

$$K = K_1 \dotplus K_2 \quad ,$$

we have

$$(H|K_1 \dotplus K_2) \sim (K_1 \dotplus K_2|H)$$

by(19) and hence

(22) $$(H|K_1 \dotplus K_2) \sim (H|K_1) \dotplus (H|K_2) \quad .$$

(21)and(22) t~~h~~ogether yield the useful formula

$$(H_1 \dotplus H_2|K_1 \dotplus K_2) \sim (H_1|K_1) \dotplus (H_1|K_2) \dotplus (H_2|K_1) \dotplus (H_2|K_2)$$

or, more generally*

(23) $$\left(\sum_i H_i \,\Big|\, \sum_j K_j\right) \sim \sum_{i,j} (H_i|K_j)$$

When investigating the invariants(i.e. minimal indices and elementary divisors)of the pencil

$$(H|K)$$

we shall first of all replace H and K by their canonical forms(§1,1o)which is permissible by(2o). Secondly, since H and K will then appear as a direct sum, we only need to determine the invariants of the different terms

---

* This formula has implicitly been used before by various writers, e.g.
Williamson, Bull. Amer. Math. Soc. 37, p.386 *(1931)*
Rutherford, Proc, Akad. Wetensch. Amsterdam, 36, p435 *(1933)*
Roth, ref.3, p.463.

(24) $\qquad (H_r | K_s)$

by(23).Now,in the canonical form§1,(1o) there occur five different kinds of pencils,viz.

(25) $\qquad$ L,M,N,Q,andW,

if we for brevity leave out the indices referring to the degree and write W instead of $W(\alpha)$or $W(\beta)$etc.. Hence twenty-five different pairs(24)seem to be possible since each pencil(25)must be combined with itself and all the others.But on account of(19)their number at once reduces to fifteen which may be arranged in the following scheme:

	1.(L|L)	2.(L|M)	3.(L|N)	4.(L|Q)	5.(L|W)
		6.(M|M)	7.(M|N)	8.(M|Q)	9.(M|W)
(26)			1o.(N|N)	11.(N|Q)	12,(N|W)
				13.(Q|Q)	14.(Q|W)
					15.(W|W)

Some of these cases will readily be eliminated on account of symmetry in the formulae or similar arguments;only the last three cases have been already considered*;they cover the cases of non-singular pencils. For the remaining ones we shall obtain explicit solutions in the next section.

§4.

Special Pencils of Direct Products.

We shall first quote Aitken's and Roth's results in our notation:

---

ref.1, 2, 3.

THEOREM II (case15)
If $m \geq n$, $\alpha \neq o$, $\beta \neq o$, we have

$$(W(\alpha) | W(\beta)) \sim W_{m+n-1}(\alpha\beta) \dotplus W_{m+n-3}(\alpha\beta) \dotplus W_{m+n-5}(\alpha\beta) \dotplus \cdots \dotplus W_{m-n+1}(\alpha\beta);$$

if $m \leq n$, we have to interchange m and n in the above results.

THEOREM III(case 14)

$$(Q_m | W_n(\alpha)) \sim Q_m \dotplus Q_m \dotplus \cdots\cdots \dotplus Q_m = nQ$$

for $\alpha \neq o$.

and

THEOREM IV (case 13)

$$(Q_m | Q_n) \sim 2Q_1 \dotplus 2Q_2 \dotplus \cdots \dotplus 2Q_{n-1} \dotplus (m-n+1)\, Q_n$$

for $m \geq n$ .If $m \leq n$,we have to interchange m and n in the result.

An equation between or an equivalence of two pencils is always an identity in ρ and σ, we are therefore allowed e.g. to interchange ρ and σ. If we do this, the pencil

$$Q_k = \sigma I_k + \rho U'_k \qquad (\S 1,9)$$

becomes

$$N'_k = \rho I_k + \sigma U'_k \qquad (\S 1,4)$$

which is equivalent to

$$N_k = \rho I_k + \sigma U_k$$, for $N'_k = J_k N_k J_k$

(see p.2). By the same substitution

$$(Q_m | Q_n) = \sigma(I_m \times I_n) + \rho(U'_m \times U'_n)$$

is transformed into

$$(N'_m | N'_n) = \rho(I_m \times I_n) + \sigma(U'_m \times U'_n)$$

, which is equivalent to

$$(N_m|N_n) = \rho(I_m\times I_n)+\sigma(U_m\times U_n) \qquad \text{by}\S 3\ ,2o$$

Hence from theorem IV we can at once derive the

COROLLARY (case 1o).

$$(N_m|N_n)\sim 2N_1\dot{+}2N_2\dot{+}\dots\dot{+}2N_{n-1}\dot{+}(m-n+1)N_m$$

for $m\geq n$. If $m\leq n$, the indices m and n have to be interchanged.

Before entering into the discussion of new cases we shall explain our method by a simple example:
Find the canonical form of the pencil

(1) $\qquad (L_2|M_3)\sim\rho(A\times C)+\sigma(B\times D) = \rho X+\sigma Y$

where

(1a) $\qquad L_2\sim\rho A+\sigma B \qquad ; \ M_3\sim\rho C+\sigma B$

and

(1b) $\qquad X = (A\times C) \qquad ; \ Y = (B\times D)$.

L_2 is of type 2×3 and M_3 is of type 4×3 (§1,(2)and(3))
Hence $(L_2|M_3)$ is of type 8×9 .We write(1a) as

(2) $\qquad (\rho A+\sigma B)R = \bar{R}L_2$

(3) $\qquad (\rho C+\sigma D)S = \bar{S}M_3$

where $R,\bar{R},S,\bar{S}$ are non-singular matrices;let

(4) a) $R=[r_1,r_2,r_3]$ $\qquad$ b) $\bar{R}=[\bar{r}_1,\bar{r}_2]$

(5) a) $S=[s_1,s_2,s_3]$ $\qquad$ b) $\bar{S}=[\bar{s}_0,\bar{s}_1,\bar{s}_2,\bar{s}_3]$

The vectors $r_i,\bar{r}_k,s_j,\bar{s}_h$, are therefore of degrees 3 , 2,3,and4 resp.According to§2 equations(2)and(3)are equivalent to vector chains of type(Ib)and(IIb)which we write in the standard form §2,(14):

(6) $\bar r_0 = Br_1$; $\bar r_1 = Ar_1 = Br_2$; $\bar r_2 = Ar_2 = Br_3$; $\bar r_3 = Ar_3$

(7) $\bar s_0 = Ds_1$; $\bar s_1 = Cs_1 = Ds_2$; $\bar s_2 = Cs_2 = Ds_3$; $\bar s_3 = Cs_3$

where

(8) $\bar r_0 = 0$; $\bar r_3 = 0$; $\bar s_0 \neq 0$; $\bar s_3 \neq 0$.

We now introduce 9 vectors of degree 9

(9) $$(i,j) = (r_i \times s_j) \quad (i=1,2,3;\, j=1,2,3)$$

and 16 vectors of degree 8

(10) $$(\overline{h,k}) = (\bar r_h \times \bar s_k) \quad (h=0,1,2,3;\, k=0,1,2,3)$$

The sets (9)and(10) can be arranged in two arrays:

(11) a) $$\begin{array}{ccc} (11) & (12) & (13) \\ (21) & (22) & (23) \\ (31) & (32) & (33) \end{array} \qquad \text{b)} \quad \begin{array}{cccc} (\overline{00}) & (\overline{01}) & (\overline{02}) & (\overline{03}) \\ \hline (\overline{10}) & (\overline{11}) & (\overline{12}) & (\overline{13}) \\ (\overline{20}) & (\overline{21}) & (\overline{22}) & (\overline{23}) \\ \hline (\overline{30}) & (\overline{31}) & (\overline{32}) & (\overline{33}) \end{array}$$

Since R and S are non-singular matrices,the 9 vectors of(11a)are linearly independent(§3,lemma I).In scheme b)the vectors in the first and in the last row are zero on account of (8);the remaining 8 vectors are linearly independent according to the same lemma because $\bar R$ and $\bar S$ are non-singular.Next,we number the diagonals of those two schemes by attaching to them the differences j-i which are constant for all elements(i,j) that lie on the same diagonal. From each diagonal we ~~is connected with~~ construct a matrix whose columns are the vectors which lie on this diagonal,provided they are not zero. Thus

$$\text{(12) a)}\quad \begin{aligned} T_{-2} &= [(31)] \\ T_{-1} &= [(21),(32)] \\ T_o &= [(11),(22),(33)] \\ T_1 &= [(12),(23)] \\ T_2 &= [(13)] \end{aligned} \qquad\qquad \text{b)}\quad \begin{aligned} \overline{T}_{-2} &= [(\overline{2o})] \\ \overline{T}_{-1} &= [(\overline{1o}),(\overline{21})] \\ \overline{T}_o &= [(\overline{11}),(\overline{22})] \\ \overline{T}_1 &= [(\overline{12}),(\overline{23})] \\ \overline{T}_2 &= [(\overline{13})] \end{aligned}$$

Now by(6)and(7)and by §3,(4)we have

$$(\overline{r}_h\times\overline{s}_k)=(Ar_h\times Cs_k)=(r_h\times s_k)(A\times C) \qquad (h=1,2,3;k=1,2,3)$$

and

$$(\overline{r}_h\times\overline{s}_k)=(Ar_h\times Cs_k)= (Br_{h+1}\times Ds_{k+1})=(r_{h+1}\times s_{k+1})(B\times D)$$

(for $h=o,1,2;k=o,1,2$) ;or

$$(13) \qquad (\overline{h,k}) = X(h,k) \qquad (h=1,2,3;k=1,2,3)$$

$$(14) \qquad (\overline{h,k}) =Y(h+1,k+1) \qquad (h=o,1,2;k=o,1,2)$$

These two equations enable us to set up vector chains for the pencil $\rho X+\sigma Y$.

(15a) $(\overline{oo})=Y(11);(\overline{11})=X(11)=Y(22);(\overline{22})=X(22)=Y(33);(\overline{33})=X(33)$
(15b) $(\overline{o1})=Y(12);(\overline{12})=X(12)=Y(23);(\overline{23})=X(23)$
(15c) $(\overline{o2})=Y(13);(\overline{13})=X(13)$
(15d) $(\overline{1o})=Y(21);(\overline{21})=X(21)=Y(32);(\overline{32})=X(32)$
(15e) $(\overline{2o})=Y(31);(\overline{31})=X(31)$

In order to see to what submatrices these chains belong,we have to investigate their initial and final links.In(15a)we have by(8)

$$(\overline{oo})=o \qquad (\overline{33})=o$$

Hence(15a)is of type(L) (§2,(15)) and corresponds to the submatrix L_2.In fact,we can establish the equation

$$(16a) \qquad (\rho X+\sigma Y)T_o= \overline{T}_oL_2$$

The chain(15b)starts with a zero vector and finishes with a non-zero one.It is of type(~~IIIb~~ N)and equivalent to:

$$(16b)\qquad (\rho X+\sigma Y)T_2 = \bar{T}_2 N_1$$

In the same way we obtain the equations

$$(16))\qquad (\rho X+\sigma Y)T_2 = \bar{T}_2 M_1$$

$$(16d)\qquad (\rho X+\sigma Y)T_{-1} = \bar{T}_{-1} Q_2$$

$$(16e)\qquad (\rho X+\sigma Y)T_{-2} = \bar{T}_{-2} Q_1$$

The classification of the chains(15)can be illustrated by means of the scheme(11b)which is divided into three areas by the two horizontal lines;the middle part is occupied by ~~xxxxx~~ non-zero vectors while the top and the bottom is filled up by zeros. When produced both ends of the diagonal with suffix zero enter the zero area;the corresponding chain(15a) therefore starts and finishes with a zero-vector. The diagonals 1 and 2 ~~overlapp~~ enter the zero area only at the top,-1 and -2 only at the bottom.This is characteristic for sub - matrices N and Q resp,

Equations(16)can be combined in:

$$(17)\qquad (\rho X+\sigma Y)T = \bar{T}(L_2 \dotplus N_2 \dotplus N_1 \dotplus Q_2 \dotplus Q_1)$$

where

$$T=[T_0, T_1, T_2, T_{-1} T_{-2}] \text{ and } \bar{T}=[\bar{T}_0, \bar{T}_1, \bar{T}_2, \bar{T}_{-1} \bar{T}_{-2}]$$

T has nine columns,viz.the columns of $T_0, T_1, \ldots$ taken together;they are linearly independent because they are just the vectors arranged in(11a).Hence T is a square matrix of non-zero determinant;the same can be shown for $\bar{T}$ and we have therefore proved that the canonical form of $\rho X+\sigma Y$ is

$$L_2 \dotplus N_2 \dotplus N_1 \dotplus Q_2 \dotplus Q_1 \quad .$$

~~In~~ Turning from the example to the general case we have two standard chains(§2,14) of length m and n resp.

(18) $\bar{r}_0=Br_1;\bar{r}_1=Ar_1=Br_2;\bar{r}_2=Ar_2=Br_3;\ldots\bar{r}_{m-1}=Ar_{m-1}=Br_m;\bar{r}_m=Ar_m$

(19) $\bar{s}_0=Ds_1;\bar{s}_1=Cs_1=Ds_2;\bar{s}_2=Cs_2=Ds_3;\ldots\bar{s}_{n-1}=Cs_{n-1}=Ds_n;\bar{s}_n=Cs_n$

As in the example we define the vectors

(20) $$(i,j)=(r_i\times s_j)\qquad(i=1,2,..,m;j=1,2,..,n)$$

and

(21) $$(\overline{h,k})=(\bar{r}_h\times\bar{s}_k)\qquad(h=0,1,..,m;k=0,1,..,n)$$

and arrange them in two arrays

(22) a) $$\begin{matrix}(11) & (12) & \ldots & (1n)\\ (21) & (22) & \ldots & (2n)\\ \\ (m1) & (m2) & & (mn)\end{matrix}$$ b) $$\begin{matrix}(\overline{00}) & (\overline{01}) & \ldots & (\overline{0n})\\ (\overline{10}) & (\overline{11}) & \ldots & (\overline{1n})\\ \\ (\overline{m0}) & (\overline{m1}) & \ldots & (\overline{mn})\end{matrix}.$$

Again, put

$$X=(A\times C),\qquad Y=(B\times D).$$

As in (13) and (14) we have then by (18) and (19):

(23) $$(\overline{h,k})=X(h,k)\qquad(h=1,2,..,m;k=1,2,..,n)$$

(24) $$(\overline{h,k})=Y(h+1,k+1)\quad(h=0,1,..,m-1;k=0,1,..,n-1)$$

which allows us to establish vector chains correspon-ding to the diagonals of the schemes(22).E.g.if m<n, the chains referring to the diago-nals -1,0,and 1 are:

$$(\overline{10})=Y(21);(\overline{21})=X(21)=Y(32);..;(\overline{m-1,m-2})=X(m-1,m-2)=Y(m,m-1);$$
$$(\overline{m,m-1})=X(m,m-1)$$

$$(\overline{oo})=Y(11);(\overline{11})=X(11)=Y(22);..;(\overline{m-1,m-1})=X(m-1,m-1)=Y(mm);$$
$$(\overline{m,m})=X(m,m)$$

$$(\overline{o1})=Y(12);(\overline{12})=X(12)=Y(23);..;(\overline{m-1,m})=X(m-1,m)=Y(m,m+1);$$
$$(\overline{m,m+1})=X(m,m+1)$$

An investigation of the initial and final terms will show us what submatrices these chains represent; this can best be done by examining the corresponding diagonals -1, o, and 1 in the scheme (22b).

Let us now consider the case

$$(L_{m-1} \backslash L_{n-1}) \quad .$$

In analogy to (2) and (3) we put

$$(\rho A+\sigma B)R = \overline{R}L_{m-1}$$

$$(\rho C+\sigma D)S = \overline{S}L_{n-1}$$

where (cf. §2,1)

(25) a) $R=[r_1,r_2,..,r_m]$ b) $\overline{R}=[\overline{r}_1,\overline{r}_2,..,\overline{r}_{m-1}]$

(26) a) $S=[s_1,s_2,..,s_n]$ b) $\overline{S}=[\overline{s}_1,\overline{s}_2,..,\overline{s}_{n-1}]$

~~ARE NONSINGULA~~x are non-singular matrices whence the column vectors $r_i,\overline{r}_k,s_j,\overline{s}_h$ are of degrees m, m-1, n, and n-1 respectively. In the chains (18) and (19) we have by §2(15)

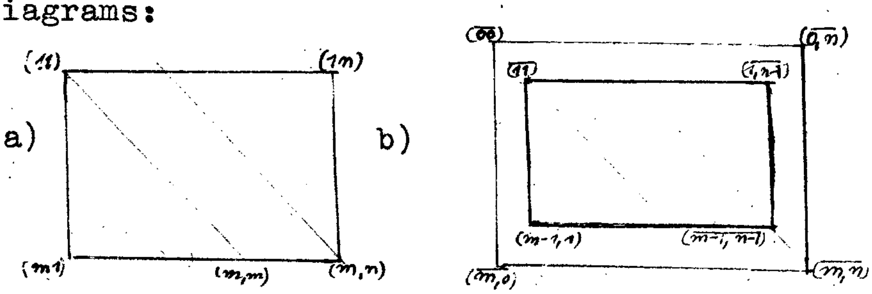
(27) $\overline{r}_o= o;\ \overline{r}_m= o;\ \overline{s}_o= o;\ \overline{s}_n= o.$

Consequently, the first and last row and column of the scheme (22b) consist of zero vectors while the remaining (m-1)(n-1) vectors of degree (m-1)(n-1)

$$(\overline{h,k}) \qquad (h=1,2,..,m-1;k=1,2,..,n-1)$$

are linearly independent by §3, lemma I. Suppose that $m \leq n$. The two schemes (22) are then sufficiently described by the diagrams:

(28) a) b)



From the left diagram we can find out the lengths of the n+m-1 diagonals. In the right hand diagram the non-zero vectors $(\overline{h,k})$ fill up the inner rectangle while the space between the two rectangles is (~~occupied by~~ a border of) zero vectors. From (28a) we see that the diagonals

(29) $-(m-1), -(m-2), \ldots, -1, 0, 1, 2, \ldots, (n-m)(n-m+1), (n-m+2) \ldots, (n-1)$

are of the respective lengths

$1, \quad 2, \ldots, m-1, m, m, m, \ldots, m, \quad m-1, \quad m-2, \quad, 1$

On the other hand, diagram (28b) shows us that every diagonal enters the zero area at either of its ends i.e. every chain starts and finishes with a zero vector and is therefore of type(Ib), (§2 (15)). As in (12) we introduce matrices T_p and $\overline{T}_p$ whose columns are the vectors which lie on the p^{th} diagonal of diagram a) and b). In the latter case we have to leave out the zeros of the diagonal, if any, e.g.:

$T_{-1} = [(10)(21)\ldots(m\,m-1)] \qquad \overline{T}_{-1} = [(\overline{21})(\overline{32})\ldots(\overline{m-1,m-2})]$

The corresponding chain is of length m-1 (by (29)) and is therefore equivalent to

$$(30) \qquad (\rho X+\sigma Y)T_{-1} = \overline{T}_{-1} L_{m-2} \qquad (§2, \text{end})$$

Similar equations hold for all indices. Finally, we remark that

$$T=\left[T_{-m+1}T_{-m+2}\cdots T_o\cdots T_{n-1}\right]$$

and

$$\overline{T}=\left[\overline{T}_{-m+2}\overline{T}_{-m+3}\cdots \overline{T}_o\cdots \overline{T}_{n-2}\right]$$

are non-singular matrices their columns being the vectors arranged in (22a) and b). We can therefore sum up equations(3o) and those related to other diagonals by

(31) $\quad (\rho X+\sigma Y)T=\overline{T}((n-m+1)L_{m-1}\dot{+}2\sum_{i=0}^{m-2}L_i)$.

Special attention is to be drawn to the fact that T contains two partial matrices more than $\overline{T}$ viz. T_{-m+1} and T_{n-1} because in diagram(28b) the corresponding diagonals -m+1 and n-1 consist only of zero vectors. In fact, by(27) we have

$$(\rho X+\sigma Y)(m,1)= o$$

$$(\rho X+\sigma Y)(1,n)= o$$

This is accounted for by two zero columns in the canonical form(§1,end) which we write as L_o (§1,5).(31) may be enunciated as

<u>THEOREM V</u> (case 1):

If $m\leq n$, we have

$$(L_m|L_n)\sim(n-m+1)L_m\dot{+}2\sum_{i=0}^{m-1}L_i \quad ,$$

Since

$$(L_m|L_n)\sim(L_n|L_m) \qquad (§3,19)$$

the assumption $m\leq n$ is no loss of generality. If $m\geq n$, we have to interchange m and n.

Again, by transposition we obtain(§3,18):

(32) $\quad (L'_m|L'_n) \sim (n-m+1)L'_m \dot{+} 2\sum_{i=0}^{m-1} L'_i$

Evidently,

(33) $\quad L'_i = J_{i+1}M_iJ_i$

(see p.2,equat.4).Hence $L'_i \sim M_i$ and(32)yields the

<u>COROLLARY</u> (case 6)

$$(M_m|M_n) \sim (n-m+1)M_m \dot{+} 2\sum_{i=0}^{m-1} M_i$$

for $m \leq n$.

Next,consider the case

$$(L_{m-1}|M_n) = \rho X + \sigma Y$$

They(18)and(19)still hold if we put

$$(\rho A + \sigma B)R = \bar{R}L_{m-1}$$

$$(\rho C + \sigma D)S = \bar{S}M_n$$

where

(34) a) $R = [r_1, r_2, \ldots, r_m]$;b) $\bar{R} = [\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_{m-1}]$

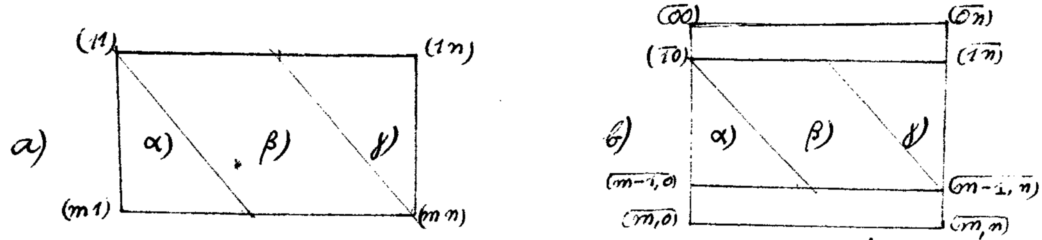
(35) a) $S = [s_1, s_2, \ldots, s_n]$;b) $\bar{S} = [\bar{s}_0, \bar{s}_1, \ldots, \bar{s}_n]$

Instead of(27)we have now(§2,15)

(36) $\quad \bar{r}_0 = 0;\ \bar{r}_m = 0;\ \bar{s}_0 \neq 0;\ \bar{s}_n \neq 0$

Consequently,the first and the last row of the scheme (22b)consist of zero vectors while the remaining vectors are a complete set* of vectors of degree$(m-1)(n+1)$. Supposing $m-1 \leq n$ we can represent(22a)and b)by the diagrams:

*i.e.they are linearly independent and their number equals their degree so that a matrix whose columns (rows)they are will be of non-zero determinant.

(37) 

The top and bottom story of b) are occupied by zeros. We have to distinguish three categories of diagonals:

α) the diagonals

$$-m+1,-m+2,\dots,-1$$

are of lengths

$$1\ ,\ 2\ ,\dots,m-1\quad .$$

the corresponding diagonals in b) enter the zero area at the bottom. The chains are therefore of type $\mathcal{Q}$ (§2, (15)) and give rise to equations

(38α) $$(\rho X+\sigma Y)T_x=\overline{T}_x Q_{p+m}\quad (x=-1,-2,\dots -m+1)$$

where the columns of $\overline{T}_x$ and $\overline{T}_x$ are the vectors that lie on the x^{th} diagonal of the schemes (37a) and b).

β) the diagonals

$$o,1,2,\dots,n-m-1,n-m$$

are all of length m; in b) they ~~overlap~~ enter the zero border twice and the chains are consequently of type $\mathcal{L}$. Hence

(38β) $$(\rho X+\sigma Y)T_y=\overline{T}_y L_{m-1}\quad (y=o,1,\dots n-m)$$

γ) in the third category the diagonals

$$n-m+1,n-m+2,\dots,n-1,n$$

are of lengths

$$m-1\ ,\ m-2\ ,\dots,\ 2\ ,1\quad .$$

In b) they ~~overlap~~ enter the border at the top; the corresponding chains are therefore of type ~~III~~ $\mathcal{N}$:

(38γ) $$(\rho X+\sigma Y)T_z=\overline{T}_z N_{n-p}\quad (z=n-m+1,n-m+2,\dots,n)$$

We can summarize the three equations (38) in:

(39) $$(\rho X+\sigma Y)T = \bar{T}\left(\sum_{i=1}^{m-1} Q_i \dot{+} (n-m+1)L_{m-1} \dot{+} \sum_{i=1}^{m-1} N_i\right)$$

where T and $\bar{T}$ consist of the columns of T_p and $\bar{T}_p$ taken together; as before we see that they are non-singular matrices whose columns are the complete sets of vectors arranged in diagram (37a) and in the non-zero area of diagram (37b). It can easily be shown that the formula still holds when $m-1=n$ in which case area β) does not appear. Replacing $m-1$ by m we have

THEOREM VI (case 2)

(40) $$(L_m|M_n) \sim (n-m)L_m \dot{+} \sum_{i=1}^{m} Q_i \dot{+} \sum_{i=1}^{m} N_i$$

for $m \leq n$ and

$$(L_m|M_n) \sim (m-n)M_n \dot{+} \sum_{i=1}^{n} Q_i \dot{+} \sum_{i=1}^{n} N_i$$

for $m \geq n$.

The second part of the theorem easily follows from the first one; for by transposition we obtain from (40):

$$(L_m'|M_n') \sim (n-m)L_m' \dot{+} \sum_{i=1}^{m} Q_i' \dot{+} \sum_{i=1}^{m} N_i'$$

By (33) we have

(41) $$L_m' \sim M_m \quad ; \quad M_n' \sim L_n$$

and similarly

(42) $$Q_i' = J_i Q_i J_i \quad ; \text{hence } Q_i' \sim Q_i$$

(43) $$N_i' = J_i N_i J_i \quad ; \text{hence } N_i' \sim N_i$$

Substituting this and using §3, (19) we get

$$(L_n|M_m) \sim (n-m)M_m \dot{+} \sum_{i=1}^{m} Q_i \dot{+} \sum_{i=1}^{m} N_i \quad (m \leq n)$$

which is identical with the second statement of the above theorem when m and n are interchanged.

The case m=n is especially interesting; for although both L_n and M_n are singular pencils, the composite pencil $(L_n|M_n)$ is <u>non-singular</u> of degree n(n+1) its determinants being equal to

$$\rho^{\frac{1}{2}n(n+1)}\ \sigma^{\frac{1}{2}n(n+1)}\ .$$

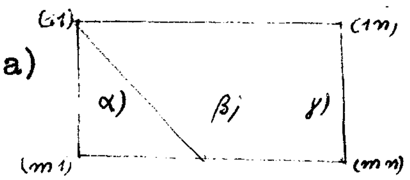
We now proceede to the discussion of the pencil

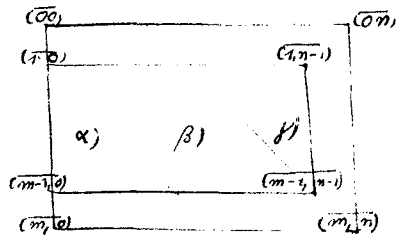
$$(L_{m-1}|Q_n) = \rho X+\sigma Y$$

Instead of (36) we have here:

(44) $\overline{r}_0 = o;\ \overline{r}_m = o;\ \overline{s}_o \neq o;\ \overline{s}_n = o$ (§2,15)

First let m≤n. The caracteristic diagrams which represent (22a) and b) are in this case:

(45) a) 

b) 

where the space between the two rectangles in b) is occupied by zero-vectors in virtue of (44). Again, we have three categories of diagonals ~~containing~~ which may be arranged in the following table containing the number of each diagonal, its length, and the submatrix which it represents according to §2,15.

	α)	β)	γ)
diagonals	-1, -2, .., -m+1	o , 1 , .., n-m	n-m+1, n-m+2, .., n-1
lengths	m-1, m-2, .., 1	m , m , .., m	m-1, m-2 , .., 1
submatrices	$Q_{m-1}, Q_{m-2}, .., Q_1$	$L_{m-1}, L_{m-1}, .., L_{m-1}$	$L_{m-2}, L_{m-3}, .., L_o$

Hence we can establish three kinds of equations:

(46α) $(\rho X+\sigma Y)T_x=\overline{T}_x Q_{m+p}$ (x=-1,-2,..-m+1)

(46β) $(\rho X+\sigma Y)T_y=\overline{T}_y L_{m-1}$ (y= o, 1,..,n-m)

(46γ) $(\rho X+\sigma Y)T_z=\overline{T}_z L_{n-p-1}$ (z = n-m+1, ..., n-1) or

(47) $$(\rho X+\sigma Y)T = \overline{T}\left(\sum_{i=1}^{m-1} Q_i \dotplus (n-m+1)L_{m-1} \dotplus \sum_{i=0}^{m-2} L_i\right) \quad (m \leq n)$$

where T and $\overline{T}$ are non-singular matrices whose columns are those of T_x, T_y, T_z taken together, i.e. the vectors

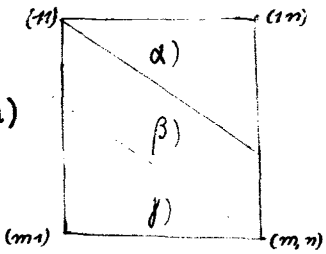
$$(i,j) \qquad (i=1,2,..,m;\ j=1,2,..,n)$$

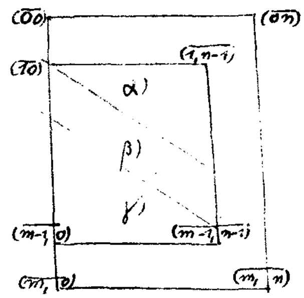
and

$$(\overline{h,k}) \qquad (h=1,2,..m-1;\ k=o,1,..n-1)$$

resp. which again form two complete sets of vectors.

if $m \geq n$, the characteristic rectangles assume the shapes

(48) a) 

b) 

and the table of diagonals becomes

diagonals	o , 1 , 2 ,..,n-1	-1,-2 ..,-m+n	-m+n-1, ,-m+1
lengths	n ,n-1,n-2,.., 1	n, n,.., n	n-1 ,.., 1
submatrices	$L_{n-1}, L_{n-2}, L_{n-3}, \ldots, L_o$	$Q_m, Q_m, \ldots, Q_m$	$Q_{n-1}, \ldots, Q_1$

In the same manner we now get

(49) $$(\rho X+\sigma Y)T = \overline{T}\left(\sum_{i=0}^{n-1} L_i \dotplus (m-n)Q_n \dotplus \sum_{i=1}^{n-1} Q_i\right) \qquad (m \geq n)$$

It is easy to see that (47) and (49) yield identical results when $m=n$:

<u>THEOREM VII(</u>case 4)

$$(L_{m-1}|Q_n) \sim (n-m)L_{m-1} \dotplus \sum_{i=1}^{m-1} Q_i \dotplus \sum_{i=0}^{m-1} L_i \qquad \text{for } m \leq n$$

$$(L_{m-1}|Q_n) \sim (m-n)Q_n \dotplus \sum_{i=1}^{n-1} Q_i \dotplus \sum_{i=0}^{n-1} L_i \qquad \text{for } m \geq n$$

By transposition and by(41)and(42)we obtain

COROLLARY I (case 8)

$$(M_{m-1}|Q_n) \sim (n-m)M_{m-1} \dotplus \sum_{i=1}^{m-1} Q_i \dotplus \sum_{i=0}^{m-1} M_i \qquad \text{for } m \leq n$$

$$(M_{m-1}|Q_n) \sim (m-n)\ Q_n \dotplus \sum_{i=1}^{n-1} Q_i \dotplus \sum_{i=0}^{n-1} M_i \qquad \text{for } m \geq n$$

Further corollaries can be obtained by interchanging ρ and σ in the above results.This process replaces Q_i by N_i and vice versa,while L_i and M_i are transformed into

$$J_i L_i J_{i+1} \quad \text{and} \quad J_{i+1} M_i J_i$$

resp.and therefore remain equivalent to themselves

COROLLARY II (case 3)

$$(L_{m-1}|N_n) \sim (n-m)L_{m-1} \dotplus \sum_{i=1}^{m-1} N_i \dotplus \sum_{i=0}^{m-1} L_i \qquad \text{for } m \leq n$$

$$(L_{m-1}|N_n) \sim (n-m)N_n \dotplus \sum_{i=1}^{n-1} N_i \dotplus \sum_{i=0}^{n-1} L_i \qquad \text{for } m \geq n$$

and

COROLLARY III (case 7)

$$(M_{m-1}|N_n) \sim (n-m)M_{m-1} \dotplus \sum_{i=1}^{m-1} N_i \dotplus \sum_{i=0}^{m-1} M_i \qquad \text{for } m \leq n$$

$$(M_{m-1}|N_n) \sim (m-n)N_n \dotplus \sum_{i=1}^{n-1} N_i \dotplus \sum_{i=0}^{m-1} M_i \qquad \text{for } m \geq n$$

Next,consider the pencil

$$(N_m|Q_n) = \rho X + \sigma Y \quad .$$

We have now to put

$$(\rho A + \sigma B)R = \bar{R}\ N_m$$

$$(\rho C + \sigma D)S = \bar{S}\ Q_n$$

where

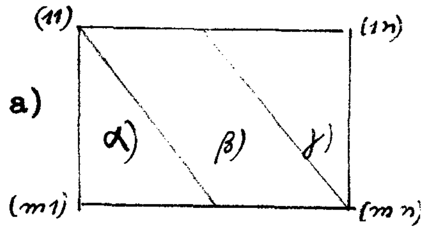
(5o) a) $R = [r_1, r_2, \ldots, r_m]$ b) $\bar{R} = [\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_m]$

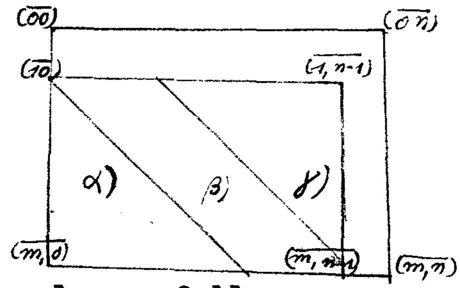
(51) a) $S = [s_1, s_2, \ldots, s_n]$ b) $\bar{S} = [\bar{s}_0, \bar{s}_1, \ldots, \bar{s}_{n-1}]$

are non-singular matrices and

(52) $r_o = o;\ r_m \neq o;\ s_o \neq o;\ s_n = o$

supposing that $m \leq n$ we can illustrate the different vector chains by the diagrams

(53) a) 

b) 

from which we tabulate the diagonals as follows

diagonals	$-1, -2, .., -m+1$	$o, 1, .., n-m-1$	$n-m, n-m+1, .., n-1$
lengths	$m-1, m-2, .., 1$	$m, m, .., m$	$m, m-1, .., 1$
submatrices	$M_{m-1}, M_{m-2}, .., M_1$	$N_m, N_m, .., N_m$	$L_{m-1}, L_{m-2}, .., L_o$

for the diagonals of α) do not ~~overlap~~ enter the zero area at all, those of β) ~~overlap~~ once viz. at the top and those of γ) twice. Correspon- ding to the three kinds of diagonals we have three sets of equations

(54α) $(\rho X + \sigma Y)T_x = \overline{T}_x M_{m+p} \quad (x=-1, -2, .., -m+1)$

(54β) $(\rho X + \sigma Y)T_y = \overline{T}_y N_m \quad (y = o, 1, .., n-m-1)$

(54γ) $(\rho X + \sigma Y)T_z = \overline{T}_z L_{n-z-1} \quad (z=n-m, n-m+1, .., n-1)$

Let T be the matrix whose columns are those of T_x, T_y, T_z taken together; it will then evidently posses as columns the mn vectors (i,j) arranged in (53a); they form a complete set of vectors and hence T is square and non-singular. If we define $\overline{T}$ in a similar way, we have by (54) α) β) γ) :

$$(55)\quad (\rho X+\sigma Y)T = \overline{T}\Big(\sum_{i=1}^{m-1} M_i \dotplus (n-m)N_m \dotplus \sum_{i=0}^{m-1} L_i\Big) = \overline{T}(\rho X_1+\sigma Y_1)$$

But $\overline{T}$ is not a square matrix, in that it contains only nm-1 columns, viz. all non-zero vectors of the scheme (53b) except $(\overline{m,o}) = (\overline{r}_m \times \overline{s}_o) \neq o$ which lies on the diagonal with suffix -m. This diagonal has not been considered since it does not occur in diagram a). The matrix

$$T^{*} = [(\overline{m,o})\ \overline{T}]$$

is, however, square and non-singular and we shall write the right hand side of (55) as

$$[(\overline{m,o})\ \overline{T}] \begin{bmatrix} \cdot \\ \rho X_1+\sigma Y_1 \end{bmatrix}$$

i.e. the canonical form is to be augmented by a zero ~~column~~ row which we denote by M_o (§1,5b). This is analogous to the term L_o which is due to the fact that a non-zero diagonal with index n-1 occurs in a) but not in b). We have therefore proved

<u>THEOREM VIII</u> (case 11)

$$(N_m|Q_n) \sim \sum_{i=0}^{m-1} M_i \dotplus (n-m)N_m \dotplus \sum_{i=0}^{m-1} L_i \qquad \text{for } m \leq n$$

$$(N_m|Q_n) \sim \sum_{i=0}^{n-1} M_i \dotplus (m-n)Q_n \dotplus \sum_{i=0}^{n-1} L_i \qquad \text{for } m \geq n$$

The second result easily follows from the first one by interchanging ρ and σ and m and n (cf. p.32). Although composed out of two non-singular pencils N_m and Q_n the pencil $(N_m|Q_n)$ is always singular.

For the remaining cases our method has to be modified owing to the appearence of the submatrix $W_m(\alpha), \alpha \neq o$.

We shall first consider the pencil

$$(W_m(\alpha)|L_n) = \rho X+\sigma Y \ .$$

By §1,(8)we have

(56) $$W_m(\alpha) = \rho(\alpha I_m+U'_m)+\sigma I_m \qquad (\alpha\neq o)$$

Put

$$\rho C+\sigma D = L_n$$

or

(58) $$(\rho C+\sigma D)S = \bar{S}\cdot L_n$$

where

(59) a) $S =[s_o,s_1,..,s_n]$ b) $\bar{S} =[\bar{s}_1,\bar{s}_2,..,\bar{s}_n]$

are square matrices of non-zero determinant;they are, in fact,unit matrices of degree n+1 and n. We have then by ~~(17)~~ §3,17:

(6o) $$X = (E\times C) \qquad Y = (I\times D)$$

where

(61) $$E = \alpha I_m+U'_m \qquad (\alpha\neq o)$$

is a non-singular matrix;for brevity,we shall write I instead of I_m.According to§2,equ.(58)above is equivalent to a vector chain of type(L)viz.

(62) $o=Ds_o$; $\bar{s}_1=Cs_o=Ds_1$; $\bar{s}_2=Cs_1=Ds_2$;...; $\bar{s}_n=Cs_{n-1}=Ds_n$; $Cs_n=o$.

Next,let

$$r_1,r_2,...,r_m$$

be any set of m linearly independent column vectors of degree m.By §3,lemma II,the m(n+1) vectors

(63) $$(i,j) = (E^j r_i \times s_j) \qquad (i=1,2,..,m;\, j=o,1,..,n)$$

each of which has m(n+1) elements, are linearly independent and so are the vectors

(64) $$(\overline{h,k}) = (E^k r_h \times \overline{s}_k) \qquad (h=1,2,..,m;\, k=1,2,.., n)$$

which are of degree mn. We arrange these two complete sets in two schemes

(65) $$a) \quad \begin{array}{cccc} (1o) & (11) & \ldots & (1n) \\ (2o) & (21) & \ldots & (2n) \\ & & & \\ (mo) & (m1) & \ldots & (mn) \end{array} \qquad b) \quad \begin{array}{cccc} (\overline{11}) & (\overline{12}) & \ldots & (\overline{1n}) \\ (\overline{21}) & (\overline{22}) & \ldots & (\overline{2n}) \\ & & & \\ (\overline{m1}) & (\overline{m2}) & \ldots & (\overline{mn}) \end{array}$$

By (60) and (62) we have (since $Ds_o = 0$; $Cs_n = 0$):

(66) $$Y(i,o)=o \;;\quad X(i,n)=o \quad (i=1,2,..,m)$$

(67) $$X(i,j-1)=(E^j r_i \times Cs_{j-1})=(E^j r_i \times Ds_j)=Y(i,j)=(\overline{i,j})$$

for i=1,2,..,m; j=1,2,..,n. These relations can be summarized in the vector chains:

(68) $$o=Y(i,o);\,(\overline{i,1})=X(i,o)=Y(i,1);\ldots;(\overline{i,n})=X(i,n-1)=Y(i,n);\; X(i,n)=o$$

for i=1,2,..,m. If we put

(69) $$T_i = [(io),(i1),..,(in)] \;;\; \overline{T}_i = [(\overline{i1}),(\overline{i2}),..,(\overline{in})],$$

(68) is equivalent to (cf. §2, Ia and b):

$$(\rho X+\sigma Y)T_i = \overline{T}_i L_n \qquad (i=1,2,..,m)$$

and

(7o) $$(\rho X+\sigma Y)T = \overline{T}(L_n \dotplus L_n \dotplus .. \dotplus L_n) = \overline{T}(mL_n)$$

where

$$T = [T_1,T_2,..,T_m] \;;\; T = [T_1,T_2,..,T_m]$$

~~T and $\overline{T}$~~

T and $\overline{T}$ are square matrices of non-zero determinants because their columns are just the vectors tabulated in(65a) and b). We have therefore proved:

THEOREM IX (case 5)

If $\alpha \neq o$

$$(W_m(\alpha)|L_n) \sim mL_n$$

By transposition we obtain the

COROLLARY (case 9)

If $\alpha \neq o$

$$(W_m(\alpha)|M_n) \sim mM_n$$

for $L'_n \sim M_n$ by(41) and $W'_m(\alpha) = J_m W_m(\alpha) J_m$

We can treat the case

$$(W_m(\alpha)|N_n) = \rho X + \sigma Y$$

briefly as it is very similar to the preceding one. Putting

$$(\rho C + \sigma D)S = \overline{S}\ N_n$$

we now have the chain

(71) $o = Ds_1;\ \overline{s}_1 = Cs_1 = Ds_2;\ \ldots;\overline{s}_{n-1} = Cs_{n-1} = Ds_n;\ \overline{s}_n = Cs_n \neq o$

where

$$S = [s_1, s_2, \ldots, s_n] \text{ and } \overline{S} = [\overline{s}_1, \overline{s}_2, \ldots, \overline{s}_n]$$

are non-singular and square. Again, let

(72) $\quad (\overline{i,j}) = (E^j r_i \times s_j) \quad (i=1,2,\ldots,m; j=1,2,\ldots,n)$

and

(73) $\quad (\overline{h,k}) = (E^k r_h \times \overline{s}_k) \quad (h=1,2,\ldots,m; k=1,2,\ldots,n)$

It then follows that

(74) $Y(i,1) = o;\ X(i,j) = Y(i,j+1); \ldots;\ X(i,n) \neq o$

which yields the vector chains

(75) $o=Y(i1);(\overline{i1})=X(i2);..;(\overline{i,n-1})=X(i,n-1)=Y(i,n);X(i,n)\neq o$
for $i=1,2,..,m$; i.e. we have m vector chains each of which represents a submatrix N_n(§2. type ~~IIIb~~ N). (75) is equivalent to

(76) $\quad (\rho X+\sigma Y)T_i = \overline{T}_i N_n \quad (i=1,2,..,m)$

where
$$T_i=[(i1),(i2),\ldots,(in)] \; ; \; \overline{T}_i=[(\overline{i1}),(\overline{i2}),\ldots,(\overline{in})]$$

Hence
$$(\rho X+\sigma Y)T = \overline{T}(N_n\dot{+}N_n\dot{+}..\dot{+}N_n)=\overline{T}(mN_n)$$

As before we can show that T and $\overline{T}$ are square matrices of non-zero determinant and we have proved

<u>THEOREM X</u> (case 12)

If $\alpha\neq o$
$$(W_m(\alpha)\backslash N_n)\sim mN_n$$

This is the last of the fifteen pencils(§3,26) whose canonical forms we wished to determine; hence we have obtained a method of finding the canonical form of any pencil of Zehfuss matrices, since the most general such pencil can be transformed into an aggregate of those fifteen pencils which we have considered.

§5.

Pencils of Compound Matrices.

If A is a matrix of type $m\times n$, its p^{th} compound* $A^{(p)}$ $(p\leq m; p\leq n)$ is a matrix whose elements are the

*ref,1,p.355

$\binom{m}{p}\times\binom{n}{p}$ minors of degree p that can be formed from the elements of A. The general element of $A^{(p)}$ will be denoted* by

$$\binom{i_1, i_2, \ldots, i_p}{j_1, j_2, \ldots, j_p} = \binom{i^{(p)}}{j^{(p)}} = \begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_p} \\ & & \\ a_{i_p j_1} & \cdots & a_{i_p j_p} \end{vmatrix}$$

The two groups of indices $i^{(p)}$ and $j^{(p)}$ are supposed to be arranged in dictionary order and the $\binom{m}{p}$ combinations of indices $i^{(p)}$ refer to the rows of $A^{(p)}$ while the $\binom{n}{p}$ combinations $j^{(p)}$ specify the columns. As is well known, compound matrices obey the multiplicative law

(1) $$(AB)^{(p)} = A^{(p)}B^{(p)}$$

We also note the rule

(2) $$(A')^{(p)} = (A^{(p)})'$$

Again, if A is of non-zero determinant, then so is $A^{(p)}$; this follows from (1) by putting $B=A^{-1}$.

We shall now consider the p^{th} compound of a direct sum. Let

$$m=m_1+m_2$$
$$n=n_1+n_2$$

and

$$A = B\dotplus C = \begin{bmatrix} B & \cdot \\ \cdot & C \end{bmatrix}$$

where B is of type $m_1\times n_1$, and C of type $m_2\times n_2$. We can then prove the

*ref.1, p.365

THEOREM XI (cf. Littlewood loc.cit.§3)

Two permutation matrices P and R can be found such that

(3) $P(B\dot{+}C)^{(p)}R = \sum_{s=0}^{p}(B^{(p-s)}\times C^{(s)})$

where P and R depend only on the number of rows and columns in B and C and not on their elements.

In(3)we have to put $B^{(0)}=C^{(0)}=1$ and to omit every term that is meaningless,i.e.in which the upper index exceeds the number of either rows or columns.Littlewood loc.cit.proves the theorem for the case of a square matrix;his arguments equally hold for a rectangular matrix and would also show that P and R do not depend on the elements of this matrix.But we will here deduce the theorem directly from the definition of a compound matrix.According to the partition of A we shall denote the elements by

(4) $\begin{pmatrix} i_1 & i_2 & \cdot\cdot & i_{p-s} & k_1 & k_2 & \cdot\cdot & k_s \\ j_1 & j_2 & \cdot\cdot & j_{p-t} & h_1 & h_2 & \cdot\cdot & h_t \end{pmatrix} = \begin{pmatrix} i^{(p-s)} & k^{(s)} \\ j^{(p-t)} & h^{(t)} \end{pmatrix}(s=0,1,2,..p)$

where

(5) $1\leq i_1<i_2 \cdot\cdot \leq m_1$; $1\leq j_1<j_2<\cdot\cdot\leq n_1$

(6) $1\leq k_1<k_2 \cdot\cdot \leq m_2$; $1\leq h_1<h_2<\cdot\cdot\leq n_2$

First of all it is easy to see that an element is zero unless s=t(Aitken loc.cit.p.366)and any non-zero minor of A is a product of a minor of B and a minor of C(one of which must in the limiting cases p=s and s=o be replaced by unity);in fact,we have

(7) $\begin{pmatrix} i^{(p-s)} & k^{(s)} \\ j^{(p-s)} & h^{(s)} \end{pmatrix} = \begin{pmatrix} i^{(p-s)} \\ j^{(p-s)} \end{pmatrix} \cdot \begin{pmatrix} k^{(s)} \\ h^{(s)} \end{pmatrix}$

Next, we pick out all combinations of indices $i^{(p-s)}$ $j^{(p-s)}, k^{(s)}, h^{(s)}$ which are possible for a fixed value of s and arrange them in each of these four groups in dictionary order. By (7) the corresponding elements form the matrix

$$(8) \qquad B^{(p-s)} \times C^{(s)}$$

This is to be done for $s=o,1,2,..,p$. Since the sub-matrices belonging to different values of s are evidently isolated, we obtain

$$(9) \qquad \sum_{s=o}^{p} B^{(p-s)} \times C^{(s)}$$

The final arrangement of <u>all</u> indices is, of course, not the dictionary order; but we may say that (9) is derived from $A^{(p)}$ by a permutation of rows and columns which depends only on the partition of A into B and C. This proves the theorem.

In this paragraph we shall deal with pencils of the form

$$(1o) \qquad (H;p) = \rho A^{(p)} + \sigma B^{(p)}$$

when the invariants of

$$(11) \qquad H = \rho A + \sigma B$$

are known. We may assume that (11) is already in canonical form, because if

$$S(\rho A + \sigma B)T = \rho \overline{A} + \sigma \overline{B}$$

i.e. $\quad SAT = \overline{A} \quad ; \quad SBT = \overline{B}$

we obtain by (1)

$$S^{(p)}(\rho A^{(p)} + \sigma B^{(p)})T^{(p)} = \rho \overline{A}^{(p)} + \sigma \overline{B}^{(p)}$$

Again, if the given pencil can be written as a direct sum

$$\rho A+\sigma B = \rho(A_1\dot{+}A_2)+\sigma(B_1\dot{+}B_2)$$

we have by theorem XI :

$$PA^{(p)}R = \sum_{s=0}^{p} A_1^{(p-s)}\times A_2^{(s)}$$

$$PB^{(p)}R = \sum_{s=0}^{p} B_1^{(p-s)}\times B_2^{(s)}$$

with the same matrices P and R for A and B. Hence on multiplying by ρ and σ and adding we get

$$P(\rho A^{(p)}+\sigma B^{(p)})R = \sum_{s=0}^{p}\rho(A_1^{(p-s)}\times A_2^{(s)})+\sigma(B_1^{(p-s)}\times B_2^{(s)})$$

or in the notation of §3,17

$$\rho A^{(p)}+\sigma B^{(p)} \sim \sum_{s=0}^{p}(\rho A_1^{(p-s)}+\sigma B_1^{(p-s)}|\rho A_2^{(p)}+\sigma B_2^{(s)})$$

We may assume that $\rho A_1+\sigma B_1$ is one of the elementary submatrices(§1)and since the case of direct products was fully treated in the last section,we have only to find the canonical form of

$$(H;p) = \rho A^{(p)}+\sigma B^{(p)}$$

when $H = \rho A+\sigma B$ is one of the elementary pencils

$$L_m, M_n, N_n, Q_n, W_n(\alpha)$$

(§1,(2a),(2b),(3),and(8)).The last two of these pencils also occur in the canonical forms of a nonsingular pencil and have been dealt with by Aitken* and Littlewood*. We state their results in our notation:

---

*ref.1 and 3.

THEOREM XII (Littlewood loc.cit, theorem IV)

If $\alpha \neq o$

$$(W_n(\alpha);p) \sim \sum_{s=o}^{\left[\frac{q-1}{2}\right]} c_s W_{q-2s}(\alpha^*)$$

where $q = p(n-p)+1$, and, c_s is the number of partitions of s minus the number of partitions of s-1 into $\leq p$ parts each $\leq n-p$. $\left[\frac{q-1}{2}\right]$ as usual, denotes the greatest integer less than or equal to $\frac{q-1}{2}$.

THEOREM XIII (Littlewood loc.cit.*, theorem V)

$$(Q_n;p) \sim \sum_{s=p}^{n} \binom{s-2}{p-2} Q_{n+1-s} \quad .$$

If we interchange ρ and σ in theorem XIII and transpose the matrices, we obtain the

COROLLARY

$$(N_n;p) \sim \sum_{s=p}^{n} \binom{s-2}{p-2} N_{n+1-s}$$

We now turn to the discussion of $\rho A + \sigma B$ when

$$\rho A + \sigma B \sim L_n$$

It will be useful first to consider a simple example. Let $p = 3$ and

(12) $\qquad (\rho A + \sigma B)R = \overline{R} \cdot L_4$

where

(13) $\quad R = [r_0, r_1, r_2, r_3, r_4]$ and $\overline{R} = [\overline{r}_1, \overline{r}_2, \overline{r}_3, \overline{r}_4]$

are non-singular matrices(e.g.unit matrices). Accor-

---

*Littlewood denotes the coefficient of $Q_{n-p+1-i}$ by c_i and defines c_i as the number of partitions into i parts each $\leq$m-1. But it is easy to see that this number is eqqual to $\binom{i+p-2}{i}$.

ding to §2, equat.(12) above is equivalent to a vector chain of type (L), viz.

(14) $o=Br_o; Ar_o=Br_1=\overline{r}_1; Ar_1=Br_2=\overline{r}_2; Ar_2=Br_3=\overline{r}_3; Ar_3=Br_4=\overline{r}_4; Ar_4=o$

Put

(15) $$(i,k,h) = [r_i, r_k, r_h]^{(3)} \quad (o \leqq i<k<h \leqq 4)$$

These are $\binom{5}{3}$ column vectors of degree $\binom{5}{3}$. They are linearly independent because by definition they are the columns of $R^{(3)}$. Similarly, we define the $\binom{4}{3}$ vectors

(16) $$\{i,k,h\} = [\overline{r}_i, \overline{r}_k, \overline{r}_h]^{(3)} \quad (1 \leq i<k<h \leq 4)$$

of degree $\binom{4}{3}$ which are likewise linearly independent, being the columns of $\overline{R}^{(3)}$. By (14) we have

$$A[r_{i-1}, r_{k-1}, r_{h-1}] = B[r_i, r_k, r_h] = [\overline{r}_i, \overline{r}_k, \overline{r}_h]$$

for $1 \leqq i<k<h \leqq 4$; or taking the third compound we obtain:

(17) $$A^{(3)}(i-1,k-1,h-1) = B^{(3)}(i,k,h) = \{i,k,h\}$$

Again,

(18) $$B^{(3)}(o,k,h)=o \; ; \; A^{(3)}(i,k,4)=o$$

Equations (17) and (18) enable us to set up the following vector chains for the pencil $\rho A^{(3)} + \sigma B^{(3)}$:

(19a) $B^{(3)}(o12)=o; A^{(3)}(o12)=B^{(3)}(123)=\{123\}; A^{(3)}(123)=B^{(3)}(234)=\{234\};$
$A^{(3)}(234)=o$

(19b) $B^{(3)}(o13)=o; A^{(3)}(o13)=B^{(3)}(124)=\{124\}; A^{(3)}(124)=0$

(19c) $B^{(3)}(o14)=o; A^{(3)}(o14)=o$

(19d) $B^{(3)}(o23)=o; A^{(3)}(o23)=B^{(3)}(134)=\{134\}; A^{(3)}(134)=o$

(19e) $B^{(3)}(o24)=o; A^{(3)}(o24)=o$

(19f) $B^{(3)}(o34)=o; A^{(3)}(o34)=o$

These six chains are all of type ~~Ib~~ $\mathcal{L}$(§2)and correspond to terms L_2,L_1,L_0,L_1,L_0,L_0 in the canonical form, where L_0 means that the canonical form contains a zero column due to a constant vector which annihilates all matrices of the pencil.Now,let

$$T_{12}=[(012),(123),(234)] \qquad \overline{T}_{12}=[\{123\},\{234\}]$$
$$T_{13}=[(013),(124)] \qquad \overline{T}_{13}=[\{124\}]$$
$$T_{14}=[(014)]$$
$$T_{23}=[(023)(134)] \qquad \overline{T}_{23}=[\{134\}]$$
$$T_{24}=[(024)]$$
$$T_{34}=[(034)]$$

Equations(19)are equivalent to

$$(\rho A^{(3)}+\sigma B^{(3)})T_{12}=\overline{T}_{12}L_2$$
$$(\rho A^{(3)}+\sigma B^{(3)})T_{13}=\overline{T}_{13}L_1$$
$$(\rho A^{(3)}+\sigma B^{(3)})T_{23}=\overline{T}_{23}L_1$$
$$(\rho A^{(3)}+\sigma B^{(3)})T_{14}=(\rho A^{(3)}+\sigma B^{(3)})T_{24}=(\rho A^{(3)}+\sigma B^{(3)})T_{34}=0$$

Hence

$$(\rho A^{(3)}+\sigma B^{(3)})T=\overline{T}(3L_0\dotplus 2L_1\dotplus L_2)$$

where

$$T=[T_{14},T_{24},T_{34},T_{13},T_{23},T_{12}] \text{ and } \overline{T}=[\overline{T}_{13},\overline{T}_{23},\overline{T}_{12}]$$

T is a square matrix of non-zero determinant;its columns are all the vectors (i,k,h)in a certain order.Similarly,we see that the columns of $\overline{T}$ are all the vectors $\{i,k,h\}$ which likewise make up a non-singular matrix.The canonical form of $\rho A^{(3)}+\sigma B^{(3)}$ is therefore

$$3L_0\dotplus 2L_1\dotplus L_2$$

Instead of writing down the matrices $T_{12},\overline{T}_{12},$..etc., explicitly it would have been sufficient to count the number of vectors (i,k,h) and $\{i,k,h\}$ which occur in all the chains together and to convince ourselves that all of them have been used. For since no vector appears twice in the above chains, we can then evidently construct a non-singular square matrix which transforms the pencil into the aggregate of submatrices represented by the vector chains. The exact shape of the transforming matrix is irrelevant.

Consider, now, the general case

$$\rho A+\sigma B \sim L_n \qquad \text{i.e.}$$

(2o) $$(\rho A+\sigma B)R = \overline{R}L_n$$

where

(21) $$R=[r_o, r_1, \ldots, r_n] \quad \text{and } \overline{R} =[\overline{r}_1, \overline{r}_2, \ldots, \overline{r}_n]$$

are non-singular matrices of degree n+1 and n resp. (2o) is equivalent to

(22) $$Br_o=o; Ar_o=Br_1=\overline{r}_1; Ar_1=Br_2=\overline{r}_2; \ldots; Ar_{n-1}=Br_n=\overline{r}_n; Ar_n=o$$

Next, we introduce $\binom{n+1}{p}$ column vectors of degree $\binom{n+1}{p}$

(23) $$(i_1, i_2, \ldots, i_p)=[r_{i_1}, r_{i_2}, \ldots, r_{i_p}]^{(p)} (o\leq i_1<i_2<\ldots<i_p\leq n)$$

which are linearly independent because they are the columns of $R^{(p)}$. Similarly, we define $\binom{n}{p}$ vectors of degree $\binom{n}{p}$ viz.

(24) $$\{j_1, j_2, \ldots, j_p\}=[\overline{r}_{j_1}, \overline{r}_{j_2}, \ldots, \overline{r}_{j_p}]^{(p)} \quad (1\leq j_1<j_2 \ldots <j_p\leq n)$$

which are also linearly independent being the columns of $\overline{R}^{(p)}$. By (22) we have

$$A[r_{i_1-1}, r_{i_2-1}, \ldots, r_{i_p-1}] =B[r_{i_1}, r_{i_2}, \ldots, r_{i_p}]=[\overline{r}_{i_1}, \overline{r}_{i_2}, \ldots, \overline{r}_{i_p}]$$

and taking the p^{th} compound of the matrices on either side we get

(25) $A^{(p)}(i_1-1,i_2-1,..,i_p-1)=B^{(p)}(i_1,i_2,..,i_p)=\{i_1,i_2,..,i_p\}$

for $1\leq i_1<i_2<...<i_p\leq n$. The same method yields

(26) $B^{(p)}(o,i_2,..,i_p)=o$; $A^{(p)}(i_1,i_2,..,i_{p-1},n)=o$

Equations(25)and(26)furnish the means of establishing vector chains for the pencil $\rho A^{(p)}+\sigma B^{(p)}$, the typical one being:

$$
\begin{aligned}
&B^{(p)}(o,i_2,..,i_p)=o\\
&A^{(p)}(o,i_2,..,i_p)=B^{(p)}(1,i_2+1,..,i_p+1)=\{1,i_2+1,..,i_p+1\}\\
&A^{(p)}(1,i_2+1,.,i_p+1)=B^{(p)}(2,i_2+2,.,i_p+2)=\{2,i_2+2,.,i_p+2\}\\
(27)\quad &\ldots\quad \ldots\quad \ldots\quad \ldots\quad \ldots\quad \ldots\\
&A^{(p)}(n-i_p-1,i_2+n-i_p-1,.,n-1)=B^{(p)}(n-i_p,i_2+n-i_p,.,n)\\
&\qquad\qquad\qquad\qquad =\{n-i_p,n-i_p+i_2,..,n\}\\
&A^{(p)}(n-i_p,n-i_p+1_2,..,n)=o
\end{aligned}
$$

where

(28) $1\leq i_2<i_3<..<i_p\leq n$

Each of these chains is of type $\mathcal{L}$ (§2)the initial and the final link being zero.The number of chains is equal to $\binom{n}{p-1}$ since the p-1 indices $i_2,i_3,..i_p$ are only subject to the condition(28).The chain(27)is of length $n-i_p$ i.e. $n-i_p+1$ vectors of the kind(23)and $n-i_p$ vectors of the kind(24)occur in it(see§2,Ib).It therefore gives rise to the submatrix

$$L_{n-i_p}$$

in the canonical form.The smallest value for i_p is obviously p-1;let

$$i_p=k \qquad (k=p-1,p,..,n)$$

According to(28)the remainig p-2 indices then have to fulfil the conditions

$$1 \leqq i_2 < i_3 < \dots < i_{p-1} \leqq k-1 .$$

Hence there are $\binom{k-1}{p-2}$ chains for which $i_p = k$, i.e. which are of length n-k; they correspond to the aggregate of submatrices

$$(29) \qquad \sum_{k=p-1}^{n} \binom{k-1}{p-2} L_{n-k}$$

In order to show that(29)is already the canonical form of $\rho A^{(p)} + \sigma B^{(p)}$ we have only to prove that the number of rows and columns is the same in $\rho A^{(p)} + \sigma B^{(p)}$ and in(29), i.e. equal to $\binom{n}{p}$ and $\binom{n+1}{p}$ resp. Indeed, since L_k has k rows and k+1 columns, the pencil(29)has

$$(30) \qquad \sum_{k=p-1}^{n} (n-k) \binom{k-1}{p-2} = \binom{n}{p}$$

rows and

$$(31) \qquad \sum_{k=p-1}^{n} (n-k+1) \binom{k-1}{p-2} = \binom{n+1}{p}$$

columns. (30)and(31)can be evaluated by using the formulae*:

$$(32) \qquad \sum_{i=a}^{b} \binom{i}{s} = \binom{b+1}{s+1} - \binom{a}{s+1}$$

$$(33) \qquad \sum_{i=a}^{b} i \binom{i-1}{s} = (s+1) \left\{ \binom{b+1}{s+2} - \binom{a}{s+2} \right\}$$

We have therefore proved the

---

*These formulae can be obtained by comparing the coefficients of x^s in the identity

$$\sum_{i=a}^{b} (1+x)^i = \frac{1}{x} \left[(1+x)^{b+1} - (1+x)^a \right]$$

and in its derivative with respect to x.

THEOREM XIV

$$(L_n;p) \sim \sum_{k=p-1}^{n} \binom{k-1}{p-2} L_{n-k}$$

Since $L'_n \sim M_n$, we obtain by transposition the

COROLLARY

$$(M\ ;p) \sim \sum_{k=p-1}^{n} \binom{k-1}{p-2} M_{n-k}$$

This concludes the investigation of pencils based on compound matrices.

§6.

Pencils of Induced Matrices.

The treatment of induced matrices does not lead to any new difficulties since a close connection between induced and compound matrices will make it possibie to refer to prₑvious results.The procedure is exactly as in the last section.*

Let A be a matrix of type $m \times n$ and consider the transformation

$$y = Ax$$

where y and x are column vectors of degree m and n resp.,viz.

(1) $y = \{y_1, y_2, \cdots, y_m\}$ $\qquad$ $x = \{x_1, x_2, \cdots, x_n\}$

The $\binom{m+p-1}{p}$ products and powers of degree p which can be formed from the quantities $y_1, y_2, \cdots y_m$ will then be aggregates of the $\binom{n+p-1}{p}$ products and powers of degree p constructed from the variables $x_1, x_2, \cdots, x_n$; we assume that these products and powers are arranged in dictio-

*cf. Littlewood's paper, ref. 2.

nary order and write this transformation as

(2) $$y^{[p]} = A^{[p]}x^{[p]}$$

Accordingly, $y^{[p]}$ and $x^{[p]}$ are column vectors and $A^{[p]}$ is a matrix of type $\binom{m+p-1}{p} \times \binom{n+p-1}{p}$; it is called the p^{th} induced matrix of A. We mention the following properties:

(3) $$(AB)^{[p]} = A^{[p]}B^{[p]} \qquad \text{(multiplicative law)}$$

(4) $$(A')^{[p]} = (A^{[p]})'$$

Further, if A is square and of non-zero determinant, then so is $A^{[p]}$; this can easily be deduced from (3) (see p. 39).

Next, we consider the induced matrix of a direct sum. Let

$$m = m_1 + m_2$$

$$n = n_1 + n_2$$

and

$$A = B \dot{+} C = \begin{bmatrix} B & \cdot \\ \cdot & C \end{bmatrix}$$

where B is of type $m_1 \times n_1$ and C of type $m_2 \times n_2$. We shall then prove the

THEOREM XV

Two permutation matrices P and R can be found such that

$$P(B \dot{+} C)^{[p]}Q = \sum_{s=0}^{p} B^{[p-s]} \times C^{[s]}$$

P and Q depend only on the number of rows and columns of B and C and not on their elements.

Proof: Let

(5) $$u = Bx$$

(6) $$v = Cy$$

where

$$u = \{u_1, u_2, \dots, u_{m_1}\} \qquad x = \{x_1, x_2, \dots, x_{n_1}\}$$

$$v = \{v_1, v_2, \dots, v_{m_2}\} \qquad y = \{y_1, y_2, \dots, y_{n_2}\}$$

are column vectors of degrees m_1, n_1, m_2, n_2 resp., and the elements of x and y are supposed to be variables. Put

$$\begin{bmatrix} x \\ y \end{bmatrix} = \{x, y\} \qquad \begin{bmatrix} u \\ v \end{bmatrix} = \{u, v\}$$

We have then

$$\{u, v\} = A\{x, y\}$$

and taking the p^{th} induced of either side:

(7) $$\{u, v\}^{[p]} = A^{[p]}\{x, y\}^{[p]}$$

$\{x\ y\}$ is a column vector whose elements are the $\binom{n}{p}$ products and powers which can be formed by the elements of x and y. Obviously, the vectors

(8) $$\{z\} = \{x^{[p]}, x^{[p-1]} \times y, x^{[p-2]} \times y^{[2]} \dots . y^{[p]}\}$$

and $\{x\ y\}$ have the same elements apart from the order. We can therefore find a permutation matrix R such that

(9) $$\{x, y\}^{[p]} = R\{z\}$$

Similarly,

(1o) $$P\{u, v\}^{[p]} = \{w\}$$

where

(11) $$\{w\} = \{u^{[p]}, u^{[p-1]} \times v, u^{[p-2]} \times v^{[2]}, \dots , v^{[p]}\}$$

Hence by (7) and (9) and (1o)

$$\{w\} = PA^{[p]}R\{z\}$$

On the other hand, we evidently have

$$\{w\} = \left(\sum_{s=0}^{p} B^{[p-s]} \times C^{[s]}\right)\{z\}$$

by (3), (5) and (6). Since there is no linear relation between the elements of $\{z\}$ it follows that

$$PA^{[p]}R = \sum_{s=0}^{p} B^{[p-s]} \times C^{[s]}$$

If

$$H = \rho A + \sigma B$$

is a given pencil, we define a new pencil

$$[H;p] = \rho A^{[p]} + \sigma B^{[p]} .$$

As before, we can show that

$$[\overline{H};p] \sim [H;p]$$

if

$$\overline{H} \sim H$$

We may therefore assume that H appears in canonical shape. Again, if

$$H = H_1 \dot{+} H_2 = \rho(A_1 \dot{+} A_2) + \sigma(B_1 \dot{+} B_2) ; \quad H_1 = \rho A_1 + \sigma B_1 ; \; H_2 = \rho A_2 + \sigma B_2 ;$$

we have by the last theorem

(12) $$P(A_1 \dot{+} A_2)^{[p]} R = \sum_{s=0}^{p} A_1^{[p-s]} \times A_2^{[s]}$$

(13) $$P(B_1 \dot{+} B_2)^{[p]} R = \sum_{s=0}^{p} B_1^{[p-s]} \times B_2^{[s]}$$

with the same matrices P and R in (12) and (13). Hence multiplying (12) by ρ and (13) by σ and adding we get

$$[H_1 \dot{+} H_2; p] \sim \sum_{s=0}^{p} ([H_1; p-s] \mid [H_2; s])$$

It is therefore sufficient to consider the pencil

$$[H;p]$$

when H is one of the elementary submatrices L, M, N, Q, W(α).

As regards the last two of these cases, Littlewood (loc. cit. §4) has proved the following theorems (cf. §5 theorems XII and XIII)

THEOREM XVI

$$[W_n(\alpha);p] \sim (W^{(\alpha)}_{n+p-1};p) \sim \sum_{s=0}^{\left[\frac{t-1}{2}\right]} c_s W_{r-s}$$

where $t = p(n-1)+1$, and c_s is the number of partitions of s minus the number of partitions of s-1 into $\leq$ p parts each $\leq$ n-1

and

THEOREM XVII

$$[Q_n;p] \sim (Q_{n+p-1};p) \sim \sum_{s=0}^{n-1}\binom{p-2+s}{p-2} Q_{n-s}$$

By transposing and interchanging ρ and σ we obtain the

COROLLARY

$$[N_n;p] \sim (N_{n+p-1};p) \sim \sum_{s=0}^{n-1}\binom{p-2+s}{p-2} N_{n-s}$$

We shall now show that similar relations hold for the singular submatrices (cf. §5, theorem XIV)

THEOREM XVIII

$$[L_n;p] = (L_{n+p-1};p) \sim \sum_{s=0}^{n}\binom{s+p-2}{p-2} L_{n-s}$$

and by transposition

COROLLARY

$$[M_n;p] = (M_{n+p-1};p) \sim \sum_{s=0}^{n}\binom{s+p-2}{p-2} M_{n-s}$$

Proof: By §1,2 we have

$$L_n = \rho F_n + \sigma G_n$$

where

(14) $\qquad F_n = [\cdot\ I_n]_{n,n+1} \qquad G_n = [I_n\ \cdot]_{n,n+1}$

Hence

$$[L_n;p] = \rho F_n^{[p]}+\sigma G_n^{[p]}$$
$$(L_{n+p-1};p)= \rho F_{n+p-1}^{(p)}+\sigma G_{n+p-1}^{(p)}$$

In order to prove theorem XVIII we have to show that

$$F_n^{[p]} = F_{n+p-1}^{(p)} \tag{15}$$

$$G_n^{[p]} = G_{n+p-1}^{(p)} \tag{16}$$

Consider the transformations

(17) a) $y = F_n x = [I_n \cdot] x$; b) $y = G_n x = [\cdot\, I_n] x$

where

$$x = \{x_1\ x_2\ \dots x_{n+1}\} \qquad y = \{y_1\ y_2\ \dots y_n\}$$

We write(17)down in full:

(19) a) F_n: $\begin{array}{l} y_1 = x_1 \\ y_2 = x_2 \\ \dots\dots \\ y_n = x_n \end{array}$ b) G_n: $\begin{array}{l} y_1 = x_2 \\ y_2 = x_3 \\ \dots\dots \\ y_n = x_{n+1} \end{array}$

$F_n^{[p]}$ is the matrix which expresses the products

(2o) $y_1^{j_1}\, y_2^{j_2} \dots y_n^{j_n}$ $(j_1+j_2+\dots+j_n = p; j_\nu \geqq 0)$

in terms of the products

(21) $x_1^{i_1} x_2^{i_2} \dots x_{n+1}^{i_{n+1}}$ $(i_1+i_2+\dots+i_{n+1}=p; i_\nu \geqq 0)$

We associate the products(2o)with the partitions

(22) $(1^{j_1}\, 2^{j_2} \dots n^{j_n})$

the p parts of which are arranged in non-decreasing order.

If we add zero to the first part,unity to the second part , two to the third part etc....p-1 to the last part, we obtain a partition

(22a) $(h_1, h_2, \ldots, h_p)$ $(1 \leq h_1 < h_2 \ldots h_p \leq n+p-1)$

These partitions are in one-one-correspondence to (22) and hence also to (20). We may therefore put

(23) $y_1^{j_1} y_2^{j_2} \ldots y_n^{j_n} = (h_1, h_2, \ldots, h_p)_y$

In the same way we introduce the notation:

(24) $x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} x_{n+1}^{i_{n+1}} = (k_1, k_2, \ldots k_p)_x$

where the partition $(k_1, k_2, \ldots, k_p)$ is derived from

$$(1^{i_1} 2^{i_2} \ldots n+1^{i_{n+1}})$$

exactly as (22a) is obtained from (22). The parts k satisfy the inequalities

$$1 \leq k_1 < k_2 < k_3 < \ldots k_p \leq n+p .$$

The effect of $F_n^{[p]}$ can now be described by the equations

$$y_1^{j_1} y_2^{j_2} \ldots y_n^{j_n} = x_1^{j_1} x_2^{j_2} \ldots x_n^{j_n}$$

or

(25a) $F_n^{[p]}$: $(h_1, h_2, \ldots, h_p)_y = (h_1, h_2, \ldots, h_p)_x$

In the same manner we get for $G_n^{[p]}$:

$$y_1^{j_1} y_2^{j_2} \ldots y_n^{j_n} = x_2^{j_1} x_3^{j_2} \ldots x_{n+1}^{j_n}$$

or

(25b) $G_n^{[p]}$: $(h_1, h_2, \ldots, h_p)_y = (h_1+1, h_2+1, \ldots, h_p+1)_x$

For the discussion of the compound matrices $F_{n+p-1}^{(p)}$ and $G_{n+p-1}^{(p)}$ we consider the transformations

(26) a) $y_{(r)} = F_{n+p-1} x_{(r)}$; b) $y_{(r)} = G_{n+p-1} x_{(r)}$ $(r=1,2,\ldots,p)$

where $x_{(r)}$ and $y_{(r)}$ are column vectors of degrees n+p and n+p-1.

Let

$$x_{(r)}=\{x_{1r}\ x_{2r}\ \cdots x_{n+p,r}\} \text{ and } y_{(r)}=\{y_{1r}\ y_{2r}\ \cdots y_{n+p-1,r}\}$$

As is well known, $F^{(p)}_{n+p-1}$ is the matrix which expresses the determinants

$$[h_1,h_2,\ldots,h_p]_y = \begin{vmatrix} y_{h_1 1} & y_{h_1 2} & \cdots & y_{h_1 p} \\ \\ y_{h_p 1} & y_{h_p 2} & \cdots & y_{h_p p} \end{vmatrix} (1\leq h_1<h_2<\ldots h_p\leq n+p-1)$$

in terms of the determinants

$$[k_1,k_2,\ldots,k_p]_x = \begin{vmatrix} x_{k_1 1} & x_{k_1 2} & \cdots & x_{k_1 p} \\ \\ x_{k_p 1} & x_{k_p 2} & \cdots & x_{k_p p} \end{vmatrix} (1\leq k_1<k_2<\ldots k_p\leq n+p)$$

when the y's are transformed according to (26), i.e. when

a) $y_{hr} = x_{hr}$ b) $y_{hr} = x_{h+1,r}$

The effect of $F^{(p)}_{n+p-1}$ and $G^{(p)}_{n+p-1}$ can therefore be described as:

(27a) $F^{(p)}_{n+p-1}$: $[h_1,h_2,\ldots,h_p]_y = [h_1,h_2,\ldots h_p]_x$

(27b) $G^{(p)}_{n+p-1}$: $[h_1,h_2,\ldots,h_p]_y = [h_1+1,h_2+1,\ldots,h_p+1]_x$

A comparison ~~of~~ between (25) and (27) shows that the transformations belonging to $F^{[p]}_n$ and $G^{[p]}_n$ differ only in the notation of the variables from those associated with $F^{(p)}_{n+p-1}$ and $G^{(p)}_{n+p-1}$. The respective matrices are therefore identical and theorem XVIII is proved.

REFERENCES:

(1) A.C.Aitken: The Normal Form of Compound and Induced Matrices(Proc.London Math.Soc.2,Vol.38) (1933)

(2) D.E.Littlewood:Note on Induced and Compound Matrices (Proc. London Math. Soc (2) Vol 40 (1935))

(3) W.E.Roth: On Direct Product Matrices(Bull.Am.Math. Soc.,1934 p.461ff.)

(4) H.W.Turnbull and A.C.Aitken:An Introduction to the Theory of Canonical Matrices(Glasgow 1932)

(5) H.W.Turnbull:On the Reduction of Singular Matrix Pencils(Proc.Edinb.Math.Soc.2,Vol.4) (1934)

(6) W.Ledermann: Reduction of Singular Pencils of Matrices(Proc.Edinb.Math.Soc.2,Vol.4) (1934) (= part (I) of this thesis).