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

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

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

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

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,

(7) $$Z = P + QZK,$$

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

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

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

(7′) $$Y = QYK$$

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

$$Y_0 = QY_0 K.$$

By iterating this equation we get

$$Y_0 = QY_0 K = Q^2 Y_0 K^2 = \dots = Q^{k-1} Y_0 K^{k-1} = Q^k Y_0 K^k = 0,$$

since $K^k = 0$.

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

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

(1) $$z_1, z_2, \dots, z_k.$$

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

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

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

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

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

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press

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

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

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

$$\begin{aligned} & x_1 B \subset (0), \quad (i.e.\ x_1 B = 0) \\ & x_2 B \subset (x_1 A), \\ & x_3 B \subset (x_1 A_1 x_2 A), \\ & x_4 B \subset (x_1 A_1, x_2 A_1, x_3 A), \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & x_k B \subset (x_1 A_1, x_2 A_1, \ldots, x_{k-1} A). \end{aligned} \tag{1}$$

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

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

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

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press

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 n 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)

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

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press

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)

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

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

The last matrix is non-singular, because

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

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

Hence the pencil

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

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

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

where

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

Hence

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

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 I 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

$$uA = \bar{u}B, \tag{1}$$

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.)

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

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

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

It is convenient to put

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

making the equation $v_{\nu-1} A = v_\nu B$ valid also for zero and negative integers. 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

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

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

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

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

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

We have therefore constructed the chain

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

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

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

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

(*a*) The lengths are increasing

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

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

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press

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

We now construct the chain

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

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

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press

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

It follows by (7) that

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

and

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

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

Hence

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

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press

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 n). 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.*

https://doi.org/10.1017/S0013091500008051 Published online by Cambridge University Press