University of Oxford

Algebraic subjects

Theory of Matrices; and applications

Being a thesis submitted to the Board of the Faculty of Physical Sciences in application to supplicate for the degree of Doctor of Philosophy

By

S.N. Afriat

The Queen's College

BODLEIAN LIBRARY OXFORD
DEPOSITED THESIS
18 xii 53

Supervisor: Dr. W.L. Ferrar

Trinity Term, 1953

Algebraic subjects. Theory of Matrices; and applications

ABSTRACT

This thesis consists of six independent parts, each of which develops a topic in the theory of matrices.

The first part, 'Compounds, adjugates and partitioned determinants', attempts, with the introduction of suitable notation, and the proving of some general results, to contribute to and complete certain aspects of the general theory of determinants. It is specially concerned with the derived systems of matrices, the compounds and adjugates of a matrix, whose elements are formed from the minors and cofactors of that matrix, with the dual hybrid compounds, which form a generalization of these, and their determinants, and also with the expansions of bordered and partitioned determinants in terms of the compounds and adjugates of their component parts. The multiplication theorem for compounds and adjugates of Binet and Cauchy is proved, and its application is fundamental throughout. One of the results is that if matrices $\underset{\sim}{a}$, $\underset{\sim}{b}$ are square of order m, n and $\underset{\sim}{x}$, $\underset{\sim}{y}$ are of order mxn, nxm then

$$\begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{b} \end{vmatrix} = |\underset{\sim}{a}||\underset{\sim}{b}| + \sum_{r=1}^{k-1} (-1)^r \operatorname{trace}(\underset{\sim}{a}^{[r]}\underset{\sim}{x}^{(r)}\underset{\sim}{b}^{[r]}\underset{\sim}{y}^{(r)}) + (-1)^k \begin{vmatrix} \underset{\sim}{x}^{(k)} & \underset{\sim}{b}^{[k]} & \underset{\sim}{y}^{(k)} \\ & |\underset{\sim}{x}||\underset{\sim}{y}| & \\ \underset{\sim}{y}^{(k)} & \underset{\sim}{a}^{[k]} & \underset{\sim}{x}^{(k)} \end{vmatrix}$$

according as $m<n$, $m=n$ or $m>n$, where $k=\min(m,n)$.

The second part, 'Commutative matrices, latent vectors and characteristic values', discusses commutative sets of matrices, their common latent vectors, and the sets of characteristic values which correspond to them. Also it makes a study of commutative matrix algebras with a unit over the complex field. A set of characteristic values all of which correspond to a single common latent vector of a set of commutative matrices defines a characteristic set of that matrix set. Two fundamental results proved are that every commutative matrix set has at least one characteristic set, and that the characteristic sets of any commutative matrix set can be extended to give the characteristic sets of any commutative extension of that set. Also it is proved that the matrices of any commutative set can all be reduced to triangular form by the same unitary transformation, and the sets of similarly situated diagonal elements in the triangular transforms are the characteristic sets with certain repetitions. Through these repetitions the multiplicity of a characteristic set is defined so as to generalize the concept of the multiplicity of a characteristic value of a single matrix as its multiplicity as a root of the characteristic equation. The rank of a characteristic set is defined as the rank of the space of corresponding simultaneous latent vectors of the matrices of the set, the corresponding axial space. It is shown that the rank of

a characteristic set is at most its multiplicity. The index of a characteristic set is defined so as to generalize the concept of the index of a characteristic value of a single matrix as its multiplicity as a root of the minimum function, and it is shown that no characteristic set has zero index. Proof is given to the theorem that the characteristic values of any rational function of commutative matrices are given by the values of that function when the argument ranges over the characteristic sets, and the multiplicity of any characteristic value is determined as the sum of the multiplicities of the characteristic sets which give that value. In any commutative matrix algebra K with a unit over the complex field there exist independent idempotents e_n which are irreducible in K and which form a reduction of the unit. They are called the principal idempotents in K, and there are subalgebras K_n in K, called the principal sections of K, in which the units are e_n respectively, and which are such that $K_n . K_\rho = 0$ $(\rho \neq n)$ and $K = \Theta K_n$. There are multiplicities m_n attached, such that, for any matrix $\underset{\sim}{a}$ in K, $\underset{\sim}{a}\underset{\sim}{e}_n$ has just one characteristic value α_n, and the multiplicity of any value α as a characteristic value of $\underset{\sim}{a}$ is given by $m_\alpha = \sum_{\alpha_n = \alpha} m_n$. It follows, by the Cayley-Hamilton theorem, that $\prod_n (\underset{\sim}{a} - \alpha_n \underset{\sim}{I})^{m_n} = \underset{\sim}{0}$ for every $\underset{\sim}{a}$ in K. If indices i_n are the least exponents such that $\prod_n (\underset{\sim}{a} - \alpha_n \underset{\sim}{I}) = \underset{\sim}{0}$ for all $\underset{\sim}{a}$ in K, then $0 < i_n \leq m_n$

and

$$(\underset{\sim}{a}-\alpha_n\underset{\sim}{I})^r(\underset{\sim}{b}-\beta_n\underset{\sim}{I})^s\ldots\ldots\underset{\sim}{e}_n=\underset{\sim}{0}, \quad r+s+\ldots \geqslant i_n,$$

for all matrices $\underset{\sim}{a}$, $\underset{\sim}{b}$,... in K. It is taken that m_n defines the <u>multiplicity</u> and i_n the <u>index</u> of any <u>characteristic plane</u> n of K. For each characteristic plane n of K there corresponds an axial space of K, which is denoted by $\mathfrak{X}_n$, and is such that $\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{X}\alpha_n$ for each matrix $\underset{\sim}{a}$ of K and each vector $\underset{\sim}{X}$ of $\mathfrak{X}_n$. Also K determines a splitting of the fundamental vector space into subspaces $\mathcal{U}_n$ which are called the <u>principal spaces</u> relative to K, and are the ranges of the principal idempotents $\underset{\sim}{e}_n$, of rank m_n respectively. Each principal space contains its corresponding axial space, that is to say $\mathfrak{X}_n \subset \mathcal{U}_n$ for each n. The principal idempotents in the algebra generated by a single matrix $\underset{\sim}{a}$ define the partial resolvents $\underset{\sim}{e}_\alpha$ of that matrix, indexed by its characteristic values α. They are unique matrices such that $\underset{\sim}{a}\underset{\sim}{e}_\alpha=\underset{\sim}{e}_\alpha\underset{\sim}{a}$, $\underset{\sim}{e}_\alpha^2=\underset{\sim}{e}_\alpha\neq\underset{\sim}{0}$, $\sum_\alpha \underset{\sim}{e}_\alpha$ and $(\underset{\sim}{a}-\alpha\underset{\sim}{I})\underset{\sim}{e}_\alpha$ is nullpotent. Any characteristic values of commutative matrices which belon to a common latent vector are said to be <u>latently associated</u>. If a commutative pair of matrices are such that there is a unique characteristic value of one which is latently associated with any characteristic value of the other, then we say that the one is <u>resolved</u> by the other. It is proved that is a commutative pair of matrices are such that one resolves the other, then the independent idempotent resolution of the unit matrix relative to

the one refines that of the other; and conversely. It appears as a consequence that a commutative pair of matrices are mutually resolved if and only if they have the same resolvents. If commutative matrices $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... have resolvents $\underset{\sim}{f}_\alpha$, $\underset{\sim}{f}_\beta$, ..., then $\underset{\sim}{f}\,\underset{\sim}{f}\ldots\neq\underset{\sim}{0}$ if and only if α, β, ... are latently associated characteristic value of $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... respectively. The resolvent quantities, the axial and principal spaces of a commutative matrix algebra are expressed in terms of the same for any generating subset of matrices.

The third part, 'Axial and normal matrix sets', deals with matrix sets, the algebras generated by which span the fundamental vector space by their latent vectors, or which include among their members the transposed conjugat of each member.

The fourth part, 'Composite matrices', considers matrices whose elements are commutative matrices, the ground field being the complex numbers. A <u>composite matrix</u> of <u>order</u> n and <u>degree</u> m is an nth order matrix whose elements, called its <u>components</u>, are themselves matrices of order m. Let the i,jth component $\underset{\sim}{A}^*_{ij}$ of a composite matrix $\underset{\sim}{A}^*$ of order n and degree m have r,sth element A^{rs}_{ij}. Then correlative to $\underset{\sim}{A}^*$ there is defined the <u>dual</u> composite matrix $\underset{\sim}{A}^*_*$, of order m and degree n, whose r,sth component $\underset{\sim}{A}^{rs}_*$ has i,jth element A^{rs}_{ij}. Any

composite matrix has a ground matrix, which is that simple matrix from which it may be derived by a partitioning. A composite matrix is called commutative if its components all commute, and doubly-commutative if its dual is also commutative. The determinant of a commutative composite matrix of order n and degree m is a determinant of order n whose value is a matrix of order m; and the determinant of this determinant defines the double-determinant of that composite matrix. When the components of a commutative composite matrix are replaced by their respective characteristic values in a latently associated set, that is to say which form a characteristic set for the commutative set of components, a characteristic dual component of the composite matrix is obtained, with a multiplicity attached to it equal to the multiplicity of the characteristic set formed by its elements. The characteristic dual components are linearly dependent on the dual components. When the composite matrix is doubly commutative, a set of characteristic components are also defined. Three of the principal results are as follows. The characteristic values of the ground matrix of a commutative composite matrix are the characteristic values of its characteristi dual components, with multiplicities attached correspondingly. The double determinant of a commutative composite matrix is equal to the determinant of its

ground matrix. If $m_{\underset{\sim}{\alpha}^*}$, $m_{\underset{\sim}{\alpha}_*}$ denote the multiplicities attached to the characteristic and the characteristic dual components $\underset{\sim}{\alpha}^*$, $\underset{\sim}{\alpha}_*$ of a doubly commutative dual pair of composite matrices $\underset{\sim}{A}^*$, $\underset{\sim}{A}_*$ then

$$\prod_{\underset{\sim}{\alpha}^*}(\underset{\sim}{A}^*-\underset{\sim}{\alpha}^*\underset{\sim}{I}^*)^{m_{\underset{\sim}{\alpha}^*}}\underset{\sim}{\alpha}^*=\underset{\sim}{O}^*, \quad \prod_{\underset{\sim}{\alpha}_*}(\underset{\sim}{A}_*-\underset{\sim}{\alpha}_*\underset{\sim}{I}_*)^{m_{\underset{\sim}{\alpha}_*}}\underset{\sim}{\alpha}_*=\underset{\sim}{O}_*$$

and if

$$\underset{\sim}{C}^*(\lambda)=|\underset{\sim}{A}^*-\lambda\underset{\sim}{I}^*|, \quad \underset{\sim}{C}_*(\lambda)=|\underset{\sim}{A}_*-\lambda\underset{\sim}{I}_*|$$

then

$$\underset{\sim}{C}^*(\underset{\sim}{\alpha}^*)=\underset{\sim}{O}^*, \quad \underset{\sim}{C}_*(\underset{\sim}{\alpha}_*)=\underset{\sim}{O}_*.$$

The fifth part, 'Functions of a complex matrix variable', investigates functions with a complex matrix for argument, which extend an arbitrary analytic function of a complex variable. It shows results which are generalizations of formulae familiar in the theory of functions of a complex variable, and besides these there are also results which have no analogue in this theory.

The sixth part, 'The moduli of a complex matrix', investigates and sets out the properties of two non-negative real valued functions defined over the complex matrices, called the upper and lower moduli. These functions supply the fitting analogue for complex matrice of the concept of the modulus of a complex number. The upper modulus $|\underset{\sim}{a}|^*$ and the lower modulus $|\underset{\sim}{a}|_*$ of any matrix $\underset{\sim}{a}$ with complex elements are defined as the non-negative real numbers whose squares are the maximum and the minimum characteristic values respectively of the

non-negative definite Hermitian matrix $\bar{\underset{\sim}{a}}'\underset{\sim}{a}$, and the absolute trace $|\underset{\sim}{a}|_{(1)}$ and the absolute determinant $|\underset{\sim}{a}|_{(k)}$ are defined as the non-negative real numbers whose square are the trace and the determinant respectively of $\bar{\underset{\sim}{a}}'\underset{\sim}{a}$.

contents

I Compounds, adjugates and partitioned determinants ;

II Commutative matrices, latent vectors and characteristic values ; 28

III Axial and normal matrix sets ; 85

IV Composite matrices ; 96

V Functions of a complex matrix variable ; 117

VI The moduli of a complex matrix ; 170

Published parts are included in an appendix.

COMPOUNDS, ADJUGATES AND PARTITIONED DETERMINANTS.

I. Introduction.

In the theory of determinants and matrices some difficulty is sometimes encountered through the absence of a notation which is readily sufficient for giving certain expressions or for performing certain operations. A notation introduced here diminishes these obstacles, and leads to some new results. It also permits a useful restatement of some familiar parts of the subject. One of the principal objects is to obtain the double Laplace expansion of a determinant symmetrically partitioned in the form $\begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{b} \end{vmatrix}$, where the matrices $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are square, the result being expressed in terms of the compounds of $\underset{\sim}{x}$ and $\underset{\sim}{y}$ and the adjugates of $\underset{\sim}{a}$ and $\underset{\sim}{b}$. The simplest example of this procedure is the Cauchy expansion of a determinant [1], with respect to the elements in a single row and column: thus if $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are column vectors, k is scalar, and $\underset{\sim}{a}^*$ is the adjugate of $\underset{\sim}{a}$, then $\begin{vmatrix} k & \underset{\sim}{x}' \\ \underset{\sim}{y} & \underset{\sim}{a} \end{vmatrix} = k|\underset{\sim}{a}| - \underset{\sim}{x}'\underset{\sim}{a}^*\underset{\sim}{y}$. The multiplication theorem of Binet and Cauchy and other fundamental identities for compounds and adjugates are proved, and some identities are established for bordered and partitioned determinants. Also there is an account of the hybrid compound matrices and determinants, which have been studied mainly by Bazin and Reiss. This

account ends with a theorem of Reiss which contains the various well known results which apply here and which thus provides a unification and summary of this part of the subject. As illustrations of the present methods, brief proofs are given of Grassmann uniqueness and duality theorems for the coordinates of linear spaces. Also there is a note on the discussion of matrix rank, which follows easily on the background given here.

2. Indices and signs.

The $\binom{n}{k}$ subseries of k elements of a series of n elements have an order, called the order of last differences, in which the priority of any pair of the subseries is determined by the priority, in the original series, of their last differing elements. Thus let $\underset{\sim}{i}=(i_1,\ldots,i_k)$, $\underset{\sim}{j}=(j_1,\ldots,j_k)$ be subseries of k elements of the series $\underset{\sim}{n}=(I,\ldots,n)$, so that $I\leqslant i_1<\ldots<i_k\leqslant n$, $I\leqslant j_1<\ldots<j_k\leqslant n$; and let the last elements in which they differ be their rth, so that $i_r\neq j_r$ but $i_s=j_s$ when $s>r$. Then, for the order of last differences, write $\underset{\sim}{i}<\underset{\sim}{j}$ or $\underset{\sim}{j}<\underset{\sim}{i}$ according as $i_r<j_r$ or $j_r<i_r$. From the series $\underset{\sim}{n}=(I,\ldots,n)$, the index series of order n, and for $k\leqslant n$, there is derived the series $\underset{\sim}{n}^{(k)}=((I,\ldots,k),\ldots,(n-k+I,\ldots,n))$, called the kth compound index series of order n, which is the series of the subseries of k elements of $\underset{\sim}{n}$ taken in their order of last differences.

Proposition (i). For $k\leqslant m\leqslant n$, $\underset{\sim}{m}^{(k)}$ is the subseries of the leading $\binom{m}{k}$ terms of $\underset{\sim}{n}^{(k)}$.

Firstly, $\underset{\sim}{m}^{(k)}$ is a subseries of $\underset{\sim}{n}^{(k)}$, of $\binom{m}{k}$ terms. Also any term of $\underset{\sim}{n}^{(k)}$ which does not belong to $\underset{\sim}{m}^{(k)}$ has its final element exceeding m, and so coming after all the elements of every term of $\underset{\sim}{m}^{(k)}$. Hence it follows that $\underset{\sim}{m}^{(k)}$ is a leading subseries of $\underset{\sim}{n}^{(k)}$.

The ith term of $\underset{\sim}{n}^{(k)}$ may be denoted by $n_i^{(k)}=(i_1,\ldots,i_k)$, where $I\leq i_1<\ldots<i_k\leq n$, and $I\leq i\leq\binom{n}{k}$. The set of elements $i_1,\ldots,i_k$ in the series $\underset{\sim}{n}$ may be called the ith set of k indices taken from $\underset{\sim}{n}$; or, by proposition (i), they may without ambiguity be called the ith set of k indices.

Proposition (ii). If $I\leq i_1<\ldots<i_k\leq n$, then $n_i^{(k)}=(i_1,\ldots,i_k)$, where

$$i=I+\sum_{s=1}^{k}\binom{i_s-I}{s}$$

Let I_s be the subseries of $\underset{\sim}{n}^{(k)}$ whose terms are series of the form $(x_1,\ldots,x_s,i_{s+1},\ldots,i_k)$. Evidently the terms of I_s form a consecutive set following, say, the N_s th term of $\underset{\sim}{n}^{(k)}$. Now I_0 consists of the one term $(i_1,\ldots,i_k)$, and this is the terms after the N_0th; that is to say it is the ith term of $\underset{\sim}{n}^{(k)}$, where $i=N_0+I$. Also we have that the series I_k is identical with the series $\underset{\sim}{n}^{(k)}$; so that $N_k=0$. Consider the subseries of I_s of terms which have $x_s<i_s$ and $x_s\leq i_s$ respectively. They form leading consecutive sets in I_s, the first having $\binom{i_s-I}{s}$ terms. Accordingly the subseries of terms of I_s which have $x_s=i_s$, which is I_{s-1}, is a series of consecutive terms following the $\binom{i_s-I}{s}$th term of I_s, that is to say it is the N_{s-1} th term of $\underset{\sim}{n}^{(k)}$, where

$$N_{s-1}=N_s+\binom{i_s-I}{s}.$$

Hence, by summation with respect to s, we have $N_0=N_k+N$, where $N=\sum_{s=1}^{k}\binom{i_s-I}{s}$; and then we have $i=I+N$, since $N_k=0$. [2]

For each i, the series $n_i^{(k)}$ has, with respect to $\underset{\sim}{n}$, a complementary subseries of n-k elements, which we denote by $n_i^{[k]}$. This complementary series is the ith term of a series $\underset{\sim}{n}^{[k]}=((k+I,\ldots,n),\ldots,(I,\ldots,n-k))$, called the kth adjugate index series, of order n, which will be seen to be the series of the $\binom{n}{k}$ subseries of n-k elements of $\underset{\sim}{n}$ in their reverse order of last differences, which is the series $\underset{\sim}{n}^{(n-k)}$ taken in its reverse order. Thus let $\underset{\sim}{\sigma}'$, $\underset{\sim}{\sigma}''$ be subseries of k elements of $\underset{\sim}{n}$, and let $\underset{\sim}{\rho}'$, $\underset{\sim}{\rho}''$ be their complementary subseries. It will be shown that if $\underset{\sim}{\sigma}'$ precedes $\underset{\sim}{\sigma}''$, in the order of last differences, then $\underset{\sim}{\rho}'$ succeeds $\underset{\sim}{\rho}''$. Let the last elements in which $\underset{\sim}{\sigma}'$ and $\underset{\sim}{\sigma}''$ differ be their rth, so that

$$\underset{\sim}{\sigma}'=(\sigma_1',\ldots,\sigma_r',\sigma_{r+1}',\ldots,\sigma_k),$$
$$\sigma''=(\sigma_1'',\ldots,\sigma_r'',\sigma_{r+1},\ldots,\sigma_k),$$

where $\sigma_r'<\sigma_r''<\sigma_{r+1}$. Now σ_r'' occurs in $\underset{\sim}{\rho}'$, as the sth element, say; and this will be the last element in which $\underset{\sim}{\rho}'$ differs from $\underset{\sim}{\rho}''$. Further, $\rho_s''<\sigma_r''$. Hence $\rho_s''<\rho_s'$, and accordingly $\underset{\sim}{\rho}''$ precedes $\underset{\sim}{\rho}'$. So we have:

Proposition (iii). The series $\underset{\sim}{n}^{[k]}$ is the series $\underset{\sim}{n}^{(n-k)}$ taken in its reverse order.

Again, for each i, the series $n_2^{(k)}$ and $n_2^{[k]}$ adjoin to give a derangement $(n_2^{(k)}n_2^{[k]})$ of $\underset{\sim}{n}$, by a permutation which we may denote by

$$\pi_2^{(k)}=(n_1^{(k)}n_2^{[k]})/\underset{\sim}{n}\quad;$$

denote the alternating character of this permutation.

and let $\epsilon_i^{(k)} = \alpha(n_i^{(k)})$. It will be convenient to write $\epsilon_{ij}^{(k)} = \epsilon_i^{(k)} \epsilon_j^{(k)}$, and $\epsilon_i^{[k]} = \epsilon_i^{(n-k)}$; and $\epsilon_i^{n-(k)}$ will denote $\epsilon_i^{(k)}$ with explicit reference to the order n.

Proposition (iv). $$\epsilon_i^{(k)} = (-I)^{\sum_{s=1}^{k} i_s - \frac{1}{2}k(k+I)} = (-I)^{\frac{1}{2}n(n+I) - \frac{1}{2}k(k+I) - \sum_{s=1}^{n-k} i_s'}$$

where $n_i^{(k)} = (i_1, \ldots, i_k)$, and $n_i^{[k]} = (i_1', \ldots, i_{n-k}')$; Also if $n_j^{(k)} = (j_1, \ldots, j_k)$, then $\epsilon_{ij}^{(k)} = (-I)^{\sum_{s=1}^{k} (i_s + j_s)}$

3. Submatrices.

Let $\underline{a}$ be any matrix, of order mxn. Then the submatrix of $\underline{a}$, of order pxq, whose elements fall in the ith set of p rows and the jth set of q columns of $\underline{a}$, is denoted by $\underline{a}_{ij}^{(pq)}$, and is called the i,jth pxq submatrix of $\underline{a}$. In particular we shall denote the submatrices of order pxn, mxq formed by the ith set of p complete rows, and by the jth set of q complete columns of $\underline{a}$, by $\underline{a}_i^{(p}$ and $\underline{a}_j^{q)}$ respectively, and their complementary submatrices, of m-p rows and n-q columns respectively, by $\underline{a}_i^{[p}$ and $\underline{a}_j^{q]}$. Moreover the ith row and the jth column of $\underline{a}$ are written $\underline{a}_{(i}$ and $\underline{a}_{j)}$ respectively. Now $\underline{a}_{ij}^{(pq)}=(\underline{a}_i^{(p})_j^{q)}=(\underline{a}_j^{q)})_i^{(p}$; and other submatrices can be written $\underline{a}_{ij}^{[pq)}=(\underline{a}_i^{[p})_j^{q)}$, and so on. Typical examples of expressions for products are given by $(\underline{ab})_{ij}=\underline{a}_{(i}\underline{b}_{j)}$ and $(\underline{abc})_{ij}^{(pq]}=\underline{a}_i^{(p}\underline{bc}_j^{q]}$.

4. Compounds and adjugates.

From any matrix $\underset{\sim}{a}$ we derive a series of further matrices, called its compounds.[5] Thus the kth compound $\underset{\sim}{a}^{(k)}$ of any matrix $\underset{\sim}{a}$ is the matrix, of order $\binom{m}{k} \times \binom{n}{k}$, whose i,jth element is the kth order subdeterminant $a_{ij}^{(k)} = |\underset{\sim}{a}_{ij}^{(kk)}|$ of $\underset{\sim}{a}$, called the i,jth kth order minor of $\underset{\sim}{a}$. We have that the kth compound of the submatrix of $\underset{\sim}{a}$ formed by its ith set of k rows is the ith row of its kth compound, and similarly for columns. Expressed symbolically, this is:

$$(I) \qquad (\underset{\sim}{a}_{i}^{(k)})^{(k)} = \underset{\sim}{a}_{(i}^{(k)}, \quad (\underset{\sim}{a}_{j}^{k)})^{(k)} = \underset{\sim}{a}_{j)}^{(k)}.$$

When $\underset{\sim}{a}$ is square, there can be derived another series of matrices, called the series of adjugates of $\underset{\sim}{a}$. The kth adjugate $\underset{\sim}{a}^{[k]}$ of $\underset{\sim}{a}$ has i,jth element $a_{ij}^{[k]} = \epsilon_{ij}^{(k)} |\underset{\sim}{a}_{ji}^{[kk]}|$, and is formed by replacing each element in the transpose of $\underset{\sim}{a}^{(k)}$ by its cofactor in the determinant of $\underset{\sim}{a}$. Hence

$$(2) \qquad \underset{\sim}{a}^{(k)} \underset{\sim}{a}^{[k]} = \underset{\sim}{a}^{[k]} \underset{\sim}{a}^{(k)} = |\underset{\sim}{a}| \underset{\sim}{I} \quad ,$$

since here we have the Laplace expansions on k rows or columns in the determinant of $\underset{\sim}{a}$, these being by true factors on the diagonal, to give $|\underset{\sim}{a}|$, and by alien factors off the diagonal, to give zero.*[6] Thus a compound and its adjugate of the same order are quasi-reciprocals, that is to say reciprocals but for a scalar factor. Each is non-singular when $|\underset{\sim}{a}|$ is non-singular, and then

$$(3) \qquad (\underset{\sim}{a}^{(k)})^{-1} = |\underset{\sim}{a}|^{-1} \underset{\sim}{a}^{[k]}, \quad (\underset{\sim}{a}^{[k]})^{-1} = |\underset{\sim}{a}|^{-1} \underset{\sim}{a}^{(k)}.$$

Now take determinants in (2); then by considering the degree of polynomials involved, and the primeness of a

determinants[7], we have

$$(4) \qquad |\underset{\sim}{a}^{(r)}| = |\underset{\sim}{a}|^{\binom{n-1}{r-1}}, \qquad |\underset{\sim}{a}^{[r]}| = |\underset{\sim}{a}|^{\binom{n-1}{r}}.$$

If $k>\min(m,n)$, $\underset{\sim}{a}^{(k)}$ is taken to be the null matrix. When $\underset{\sim}{a}$ is square, of order n, it will be taken that

$$(5) \qquad \underset{\sim}{a}^{(0)} = \underset{\sim}{I}, \qquad \underset{\sim}{a}^{[0]} = |\underset{\sim}{a}| ,$$
$$\underset{\sim}{a}^{(n)} = |\underset{\sim}{a}|, \quad \underset{\sim}{a}^{[n]} = \underset{\sim}{I} \qquad .$$

Any matrix is its own first compound. The first adjugate of any matrix $\underset{\sim}{a}$ will be called <u>the adjugate</u> of $\underset{\sim}{a}$, and it will be denoted by $\underset{\sim}{a}^{*}$.

5. Laplace expansion.

Suppose $\underline{a}$ is square, and $\underline{x}$ has the same number of rows. The determinant of $\underline{a}$ with the permutation $n_i^{(k)}$ applied to its columns is

$$(1) \qquad |\underline{a}_i^{(k)}\ \underline{a}_i^{[k]}| = \epsilon_i^{(k)} |\underline{a}| \ ,$$

which is the determinant of $\underline{a}$ modified by a sign $\epsilon_i^{(k)} = \pm 1$. Hence if the ith set of k columns in $|\underline{a}|$ be replaced by the jth set of k columns of $\underline{x}$, the determinant obtained has value $\epsilon_i^{(k)} |\underline{x}_j^{(k)}\ \underline{a}_i^{[k]}|$. But the Laplace expansion of this determinant by the minors on the k substituted columns is $\underline{a}_{/i}^{[k]} \underline{x}_j^{(k)}$, by the definition of $\underline{a}^{[k]}$. Hence we have

$$(2) \qquad \underline{a}_{/i}^{[k]} \underline{x}_j^{(k)} = \epsilon_i^{(k)} |\underline{x}_j^{(k)}\ \underline{a}_i^{[k]}| \ .$$

Here the left is an expression for the Laplace expansion of the determinant on the right with a partition at its kth column.

6. A lemma.

LEMMA. If $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are of order mxn and nxm respectively, so they border the unit matrix of order n to give a determinant of order m+n, then $\begin{vmatrix} \underset{\sim}{0} & -\underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{I} \end{vmatrix} = \underset{\sim}{x}^{(m)} \underset{\sim}{y}^{(m)}$.

Consider the Laplace expansion of the determinant with respect to its leading m rows. If m>n, all the determinant on these rows vanish, so that the determinant vanishes. If m=n, we have

$$\begin{vmatrix} \underset{\sim}{0} & -\underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{I} \end{vmatrix} = (-I)^{m(2m-1)} \begin{vmatrix} -\underset{\sim}{x} & \underset{\sim}{0} \\ \underset{\sim}{I} & \underset{\sim}{y} \end{vmatrix} = (-I)^{m(2m-1)} |-\underset{\sim}{x}|\,|\underset{\sim}{y}| = (-I)^{m(2m-1)} (-I)^{m} |\underset{\sim}{x}|\,|\underset{\sim}{y}| = |\underset{\sim}{x}|\,|\underset{\sim}{y}|$$

Suppose now that m<n. Then the non-vanishing minors are $|-\underset{\sim}{x}_i^{m)}| = (-I)^{m} \underset{\sim}{x}_i^{(m)}$ $(i = I, \ldots, \binom{n}{m})$. Hence, by a permutation applied to the columns of the determinant, and a similar permutation applied to its rows, so that its value remains unchanged, we have

$$\begin{vmatrix} \underset{\sim}{0} & -\underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{0} \end{vmatrix} = \begin{vmatrix} \underset{\sim}{0} & -\underset{\sim}{x}_i^{m)} & -\underset{\sim}{x}_i^{m]} \\ \underset{\sim}{y}_i^{(m} & \underset{\sim}{I} & \underset{\sim}{0} \\ \underset{\sim}{y}_i^{(m} & \underset{\sim}{0} & \underset{\sim}{I} \end{vmatrix} .$$

By permuting the columns, this becomes

$$(-I)^{m(2m-1)} \begin{vmatrix} -\underset{\sim}{x}_i^{m)} & \underset{\sim}{0} & -\underset{\sim}{x}_i^{m]} \\ \underset{\sim}{I} & \underset{\sim}{y}_i^{(m} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{y}_i^{(m} & \underset{\sim}{I} \end{vmatrix} ,$$

showing that the cofactor of $|-\underset{\sim}{x}_i^{m)}|$ is $(-I)^{m(2m-1)} |\underset{\sim}{y}_i^{(m}| = (-I)^{m(2m-1)} \underset{\sim}{y}_i^{(m)}$

Hence the considered determinant is

$$\sum_i (-I)^{m} \underset{\sim}{x}_i^{(m)} (-I)^{m(2m-1)} \underset{\sim}{y}_i^{(m)} = \sum_i \underset{\sim}{x}_i^{(m)} \underset{\sim}{y}_i^{(m)} = \underset{\sim}{x}^{(m)} \underset{\sim}{y}^{(m)} .$$

7. The multiplication theorem.

Let $\underline{x}$, $\underline{y}$ be of order kxn, nxk respectively, where $k \leq n$. By taking determinants on either side of the identity

$$\begin{pmatrix} \underline{0} & -\underline{x} \\ \underline{y} & \underline{I} \end{pmatrix} \begin{pmatrix} \underline{I} & \underline{0} \\ -\underline{y} & \underline{0} \end{pmatrix} = \begin{pmatrix} \underline{xy} & -\underline{x} \\ \underline{0} & \underline{I} \end{pmatrix},$$

and using the lemma of §6, we get $\underline{x}^{(k)}\underline{y}^{(k)} = (\underline{xy})^{(k)}$. This expresses the so called theorem on the multiplication of determinantal arrays.*(8) Now for any pair of matrices $\underline{a}$, $\underline{b}$ with product $\underline{ab}$, we have

$$(\underline{ab})^{(k)}_{ij} = (\underline{ab})^{(k\lambda)(k)}_{ij} = (\underline{a}^{(k}_{i}\,\underline{b}^{k)}_{j})^{(k)}$$

$$= (\underline{a}^{(k}_{i})^{(k)}(\underline{b}^{k)}_{j})^{(\lambda)} = \underline{a}^{(k)}_{i}\underline{b}^{(k)}_{j} = (\underline{a}^{(k)}\underline{b}^{(\lambda)})_{ij} \quad ;$$

so that $(\underline{ab})^{(k)} = \underline{a}^{(k)}\underline{b}^{(k)}$. Now suppose $\underline{a}$ and $\underline{b}$ are square, and suppose first that they are non-singular; so that, by taking reciprocals, we get $(\underline{ab})^{[k]} = \underline{b}^{[k]}\underline{a}^{[k]}$. Then this result is also true for when $\underline{a}$ and $\underline{b}$ are alowed to be singular, by the principle of the irrelevance of algebraical inequalities.*(8a) Now for a finite product $\underline{a}\ldots\underline{d}$, by induction, we have

$$(\underline{a}\ldots\underline{d})^{(k)} = \underline{a}^{(k)}\ldots\underline{d}^{(k)}, \quad (\underline{a}\ldots\underline{d})^{[k]} = \underline{d}^{[k]}\ldots\underline{a}^{[k]}.$$

This result is the multiplication theorem*(9) for compounds and adjugates.

If $\underline{x}$ is of order mxn, and r=min(m,n), then $\underline{a}^{(s)}$ is not defined for s>r. However we note that if $\underline{a}=\underline{bc}$, where $\underline{b}$ has r columns and $\underline{c}$ has r rows, and if s>r, then we may augment $\underline{b}$ by s-r null columns and $\underline{c}$ by s-r null rows, to make $\underline{a}^{(s)} = (\underline{bc})^{(s)} = \underline{b}^{(s)}\underline{c}^{(s)}$ where now $\underline{b}^{(s)}$ and $\underline{c}^{(s)}$ are properly defined, and both are null. Hence $\underline{a}^{(s)} = \underline{0}$ if s>r.

8. Repeated compounds and adjugates.

Take compounds and adjugates in the identity

$$\underline{a}\underline{a}^{-1} = \underline{a}^{-1}\underline{a} = I \ .$$

Then, by the multiplication theorem,

$$(\underline{a}^{-1})^{(k)} = (\underline{a}^{(k)})^{-1} \ , \qquad (\underline{a}^{-1})^{[k]} = (\underline{a}^{[k]})^{-1} \ .$$

Hence, by §4,(3),

$$|\underline{a}|^{-q}\underline{a}^{[p](q)} = (\underline{a}^{-1})^{(p)(q)} = |\underline{a}^{(p)}|^{-1}\underline{a}^{(p)[q]},$$
$$|\underline{a}|^{-q}\underline{a}^{(p)(q)} = (\underline{a}^{-1})^{[p](q)} = |\underline{a}^{[p]}|^{-1}\underline{a}^{[p][q]}.$$

Substitution from §4,(4), then gives the identities

$$\underline{a}^{(p)[q]} = |\underline{a}|^{\binom{n-1}{p-1}-q}\underline{a}^{[p](q)}, \qquad \underline{a}^{[p][q]} = |\underline{a}|^{\binom{n-1}{p}-q}\underline{a}^{(p)(q)}.$$

These show that, but for a scalar factor, which is a power of the determinant, a compound of a compound is the corresponding adjugate of an adjugate, and an adjugate of a compound is the corresponding compound of an adjugate.*(10) Putting p=I, q=k in the first of these identities, we have

$$\underline{a}^{*(k)} = |\underline{a}|^{k-1}\underline{a}^{[k]},$$

which is an expression of Jacobi's theorem on the minors of the adjugate, that any minor, of order k, in the adjugate of a matrix, is equal to the complementary minor in the transpose of that matrix, multiplied by the (k-I)th power of its determinant.

Also, for the adjugate of the adjugate of a matrix, we have

$$\underline{a}^{**} = |\underline{a}|^{n-2}\underline{a},$$

showing a duality between a matrix and its adjugate corresponding, in projective geometry, to the duality of the locus and the envelope of a quadric.*(10a)

9. The determinant of a sum.

Lemma. If $\underline{a}$ is any matrix, and $\underline{d}$ a diagonal matrix, then

$$|\underline{a}+\underline{d}| = \sum_{k} \sum_{i_1 \cdots i_k} d_{i_1} \cdots d_{i_k} a^{[k]}_{i_1 \cdots i_k} .$$

For the cofactor of $d_{i_1} \cdots d_{i_k}$ in $|\underline{a}+\underline{d}|$ is the cofactor of $a^{(k)}_{i_1 \cdots i_k}$ in $|\underline{a}|$.

Corollary. $|a+I| = \sum_{\nu} \text{trace } \underline{a}^{[\nu]} = \sum_{\nu} \text{trace } \underline{a}^{(\nu)}$.

THEOREM. $|\underline{a}+\underline{b}| = \sum_{\nu} \text{trace}(\underline{a}^{(\nu)}\underline{b}^{(\nu)}) = \sum_{\nu} \text{trace}(\underline{a}^{[\nu]}\underline{b}^{(\nu)})$.

Suppose $\underline{a}$ is non-singular. Then, using the corollary above, we have

$$|\underline{a}+\underline{b}| = |\underline{a}|\,|\underline{I}+\underline{b}\underline{a}^{-1}| = |\underline{a}| \sum_{\nu} \text{trace}(\underline{b}\underline{a}^{-1})^{[\nu]}$$
$$= |\underline{a}| \sum_{\nu} \text{trace}(\underline{a}^{-1})^{[\nu]}\underline{b}^{[\nu]} = \sum_{\nu} \text{trace}(\underline{a}^{(\nu)}\underline{b}^{(\nu)}) .$$

The theorem is thus proved for when a is regular. Hence, by the principle of the irrelevance of algebraic inequalities, it is also valid when $\underline{a}$ is singular.*(11)

COROLLARY I. $|\underline{a}-\alpha\underline{I}| = \sum_{t=0}^{n} (-\alpha)^{n-t} \text{trace } \underline{a}^{(t)}$.

COROLLARY II. $|\underline{a}\underline{b}-\lambda\underline{I}| = |\underline{b}\underline{a}-\lambda\underline{I}|$.

For

$$\text{trace}(\underline{a}\underline{b})^{(t)} = \text{trace}(\underline{a}^{(t)}\underline{b}^{(t)}) = \text{trace}(\underline{b}^{(t)}\underline{a}^{(t)}) = \text{trace}(ba)^{(t)}.$$

10. Symmetrically partitioned determinants.

THEOREM I. *If* $\underline{a}$ *is regular, then* $\begin{vmatrix} \underline{a} & \underline{x} \\ \underline{y} & \underline{b} \end{vmatrix} = |\underline{a}|\,|\underline{b}-\underline{y}\underline{a}^{-1}\underline{x}|$.

By the multiplication theorem for determinants, we have

$$\begin{vmatrix} \underline{a} & \underline{x} \\ \underline{y} & \underline{b} \end{vmatrix} = \begin{vmatrix} \underline{a} & \underline{0} \\ \underline{0} & \underline{I} \end{vmatrix} \begin{vmatrix} \underline{I} & \underline{a}^{-1}\underline{x} \\ \underline{y} & \underline{b} \end{vmatrix} = |\underline{a}| \begin{vmatrix} \underline{I} & \underline{0} \\ -\underline{y} & \underline{I} \end{vmatrix} \begin{vmatrix} \underline{I} & \underline{a}^{-1}\underline{x} \\ \underline{y} & \underline{b} \end{vmatrix}$$

$$= |\underline{a}| \begin{vmatrix} \underline{I} & \underline{a}^{-1}\underline{x} \\ \underline{0} & \underline{b}-\underline{y}\underline{a}^{-1}\underline{x} \end{vmatrix} = |\underline{a}|\,|\underline{b}-\underline{y}\underline{a}^{-1}\underline{x}| .$$

THEOREM II. *If* $\underline{a}$, $\underline{b}$ *are regular, of order* m, n *and* $\underline{x}$ *and* $\underline{y}$ *are of order* mxn, nxm *respectively, then*

$$\begin{vmatrix} \underline{a} & \underline{x} \\ \underline{y} & \underline{b} \end{vmatrix} = |\underline{a}\,\|\,\underline{b}| \sum_{\kappa=0}^{k} (-I)^{\kappa} \operatorname{trace} \underline{c}^{(\kappa)} ,$$

where $\underline{c}=\underline{a}^{-1}\underline{x}\underline{b}^{-1}\underline{y}$, *and* $k=\min(m,n)$.

This expresses the determinant $\begin{vmatrix} \underline{a} & \underline{x} \\ \underline{y} & \underline{b} \end{vmatrix}$ of order m+n in terms of the sum of the principal minors of order at most min(m,n) of the matrix $\underline{c}$, signed alternately according to their order. By the theorem I above,

$$\begin{vmatrix} \underline{a} & \underline{x} \\ \underline{y} & \underline{b} \end{vmatrix} = |\underline{a}|\,|\underline{b}-\underline{y}\underline{a}^{-1}\underline{x}| = |\underline{a}|\,|\underline{b}|\,|\underline{I}-\underline{b}^{-1}\underline{y}\underline{a}^{-1}\underline{x}| ;$$

and then, by application of the theorem of §9, this gives

$$\begin{vmatrix} \underline{a} & \underline{x} \\ \underline{y} & \underline{b} \end{vmatrix} = |\underline{a}\,\|\,\underline{b}| \sum_{\kappa} (-I)^{\kappa} \operatorname{trace}(\underline{b}^{-1}\underline{y}\underline{a}^{-1}\underline{x})^{(\kappa)} = |\underline{a}\,\|\,\underline{b}| \sum_{\kappa=0}^{k} (-I)^{\kappa} \operatorname{trace} \underline{c}^{(\kappa)},$$

since $\underline{c}^{(\kappa)}=\underline{0}$ if $\kappa > \min(m,n)$, by §7. It is evident that the formula is valid when the factors in c are given a cyclic permutation.

II. Cauchy's expansion.

THEOREM. If k is any scalar and $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are vectors, then[*12]

a
$$\begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y}' & k \end{vmatrix} = k|\underset{\sim}{a}| - \underset{\sim}{y}'\underset{\sim}{a}^*\underset{\sim}{x} ,$$

where a^* is the adjugate of $\underset{\sim}{a}$; and if, also, $\underset{\sim}{a}$ is regular, then[*13]

$$\begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y}' & \underset{\sim}{k} \end{vmatrix} = |\underset{\sim}{a}|(k - \underset{\sim}{y}'\underset{\sim}{a}^{-1}\underset{\sim}{x}) \quad .$$

This is the theorem I of the previous paragraph, where the matrix $\underset{\sim}{b}$ there is of order unity.

COROLLARY. If $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are any vectors, then $\begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y}' & \underset{\sim}{0} \end{vmatrix} = -\underset{\sim}{y}'\underset{\sim}{a}^*\underset{\sim}{x}$, where $\underset{\sim}{a}^*$ is the adjugate of $\underset{\sim}{a}$.

This formula is familiar in algebraic geometry for connecting the locus and envelope equations of a quadric.[*14]

12. Double Laplace expansion.

Using the results we have established, we can now prove the following theorem:*(15)

THEOREM. If $\underset{\sim}{a}$, $\underset{\sim}{b}$ are square, of order m, n and $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are of order mxn and nxm, then

$$\begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{b} \end{vmatrix} = |\underset{\sim}{a}|\,|\underset{\sim}{b}| + \sum_{K=1}^{k-1} (-1)^K \operatorname{trace}(\underset{\sim}{a}^{[K]} \underset{\sim}{x}^{(K)} \underset{\sim}{b}^{[K]} \underset{\sim}{y}^{(K)}) + (-1)^k \begin{cases} \underset{\sim}{x}^{(k)} \\ |\underset{\sim}{x}| \\ \underset{\sim}{y}^{(k)} \underset{\sim}{a}^{[k]} \end{cases} \begin{matrix} \underset{\sim}{b}^{[k]} \underset{\sim}{y}^{(k)} \\ |\underset{\sim}{y}| \\ \underset{\sim}{x}^{(k)} \end{matrix}$$

according as m<n, m=n or m>n, where k=min(m,n).

First suppose that $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are both regular. Then, by theorem II, § 10, and the multiplication theorem for compounds and adjugates, we have

$$\begin{aligned} \begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{x} \\ \underset{\sim}{y} & \underset{\sim}{b} \end{vmatrix} &= |\underset{\sim}{a}|\,|\underset{\sim}{b}| \sum_{K=0}^{k} (-1)^K \operatorname{trace}(\underset{\sim}{a}^{-1} \underset{\sim}{x} \underset{\sim}{b}^{-1} \underset{\sim}{y})^{(K)} \\ &= |\underset{\sim}{a}|\,|\underset{\sim}{b}| \sum_{K=0}^{k} (-1)^K \operatorname{trace}((\underset{\sim}{a}^{-1})^{(K)} \underset{\sim}{x}^{(K)} (\underset{\sim}{b}^{-1})^{(K)} \underset{\sim}{y}^{(K)}) \\ &= \sum_{K=0}^{k} \operatorname{trace}(|\underset{\sim}{a}|\,(\underset{\sim}{a}^{(K)})^{-1} \underset{\sim}{x}^{(K)} |\underset{\sim}{b}|\,(\underset{\sim}{b}^{(K)})^{-1} \underset{\sim}{y}^{(K)}) \\ &= \sum_{K=0}^{k} \operatorname{trace}(\underset{\sim}{a}^{[K]} \underset{\sim}{x}^{(K)} \underset{\sim}{b}^{[K]} \underset{\sim}{y}^{(K)}) \quad . \end{aligned}$$

This last expression for the considered determinant can be written in the required form shown in the theorem by using § 4; and then the validity of this result for when $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are allowed to be singular follows from Weyl's principle of the irrelevance of algebraic inequalities.

13. Bordered determinants.

A bordered determinant is formed by taking the determinant of a matrix which has been augmented by so-many extra rows and columns, called the borders, zero elements being placed in the intersection of the bordering rows and columns. Some fundamental formulae for bordered determinants are as follows:*(16)

THEOREM I. If $\underset{\sim}{a}$ is regular, of order n, and $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are of order nxm, then

$$\left|\begin{matrix}\underset{\sim}{a} & \underset{\sim}{x}\\ \underset{\sim}{y}' & \underset{\sim}{0}\end{matrix}\right| = (-1)^m |\underset{\sim}{a}|\,|\underset{\sim}{y}'\underset{\sim}{a}^{-1}\underset{\sim}{x}| \ .$$

This follows from theorem I, §10, by taking $\underset{\sim}{b}$ as the null matrix. It generalizes the well known formula in the corollary of the theorem in §11.

THEOREM II. If $\underset{\sim}{a}$ is square, of order n, and $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are of order nxm, then

$$\left|\begin{matrix}\underset{\sim}{a} & \underset{\sim}{x}\\ \underset{\sim}{y}' & \underset{\sim}{0}\end{matrix}\right| = (-1)^m \begin{cases} \underset{\sim}{x}^{(m)\prime}\,\underset{\sim}{a}^{[m]}\,\underset{\sim}{y}^{(m)} \\ |\underset{\sim}{x}|\,|\underset{\sim}{y}| \\ 0 \end{cases}$$

according as m<n, m=n or m>n.

If m>n, the formula is identically satisfied, since each side vanishes. Assume now that m≥n, and apply §4.(2), and §8 to the above theorem, supposing first that $\underset{\sim}{a}$ is regular. Then we have

$$\begin{matrix}\underset{\sim}{a} & \underset{\sim}{x}\end{matrix}$$

$$\left|\begin{matrix}\underset{\sim}{a} & \underset{\sim}{x}\\ \underset{\sim}{y}' & 0\end{matrix}\right| = (-1)^m |\underset{\sim}{a}|\, \underset{\sim}{y}'\underset{\sim}{a}^{-1}\underset{\sim}{x} = (-1)^m |\underset{\sim}{a}| (\underset{\sim}{y}'\underset{\sim}{a}^{-1}\underset{\sim}{x})^{(m)}$$
$$= (-1)^m |\underset{\sim}{a}|\, \underset{\sim}{y}^{(m)\prime} (\underset{\sim}{a}^{-1})^{(m)} \underset{\sim}{x}^{(m)} = (-1)^m |\underset{\sim}{a}|\, \underset{\sim}{y}^{(m)\prime} (\underset{\sim}{a}^{(m)})^{-1} \underset{\sim}{x}^{(m)}$$
$$= (-1)^m \underset{\sim}{y}^{(m)\prime} \underset{\sim}{a}^{(m)} \underset{\sim}{x}^{(m)}.$$

Now finally we have that the condition of the regularity of $\underset{\sim}{a}$ is irrelevant in the conclusion.*(17)

Another theorem about bordered determinants is given in §16 below.

14. Hybrid compound matrices.

Given a pair of matrices $\underset{\sim}{a}$ and $\underset{\sim}{b}$, which are square and of the same order, we form the matrix whose elements are the determinants formed by replacing k columns of $\underset{\sim}{b}$, in the determinant $|\underset{\sim}{b}|$, by k columns of $\underset{\sim}{a}$, in all possible ways, their ordering in rows and columns following the order of the replacing and the replaced sets of columns respectively.*(18) This matrix is denoted by $(b,a)^{[k)}$, and is called the kth hybrid compound of $\underset{\sim}{b}$ with $\underset{\sim}{a}$. By §5, we have that its i,jth element is

$$(\underset{\sim}{b},\underset{\sim}{a})^{[k)}_{ij}=\mathbf{e}^{(k)}_i\,|\,\underset{\sim}{a}^{k)}_j\ \underset{\sim}{b}^{k]}_i\,|=\underset{(i}{\underset{\sim}{b}}^{[k]}\underset{j)}{\underset{\sim}{a}}^{(k)}=(\underset{\sim}{b}^{[k]}\underset{\sim}{a}^{(k)})_{ij}\ ;$$

and accordingly we have

$$(I)\qquad (\underset{\sim}{b},\underset{\sim}{a})^{[k)}=\underset{\sim}{a}^{[k]}\underset{\sim}{b}^{(k)}\ .$$

The corresponding definition by rows provides the dual hybrid, given by $(\underset{\sim}{a},\underset{\sim}{b})^{(k]}=\underset{\sim}{a}^{(k)}\underset{\sim}{b}^{[k]}$; and the adjugate hybrid is the matrix $(\underset{\sim}{a},\underset{\sim}{b})^{[k)}$ which, with $(\underset{\sim}{b},\underset{\sim}{a})^{[k)}$, satisfies the quasi-reciprocal relation

$$(2)\qquad (\underset{\sim}{a},\underset{\sim}{b})^{[k)}(\underset{\sim}{b},\underset{\sim}{a})^{[k)}=(\underset{\sim}{b},\underset{\sim}{a})^{[k)}(\underset{\sim}{a},\underset{\sim}{b})^{[k)}=|\underset{\sim}{a}||\underset{\sim}{b}|\underset{\sim}{I},$$

which can be seen to generalize the adjugate relation shown in (2), §4. Bazin's matrix, formed from the determinants obtained by substituting one column in $|\underset{\sim}{b}|$ by one column of $\underset{\sim}{a}$, may be written $[\underset{\sim}{b},\underset{\sim}{a})$, with dual $(\underset{\sim}{a},\underset{\sim}{b}]$. It is identical with the first Reissian hybrid $(\underset{\sim}{b},\underset{\sim}{a})^{[1)}=\underset{\sim}{b}^{*}\underset{\sim}{a}$. Evidently its kth compound is $[b,a)^{(k)}=|\underset{\sim}{b}|^{k-1}(\underset{\sim}{b},\underset{\sim}{a})^{[k)}$, which is the kth Reissian hybrid, with a scalar factor $|\underset{\sim}{b}|^{k-1}$; and

its kth adjugate is $[\underline{b},\underline{a})^{(k)} = |\underline{b}|^{n-k}(\underline{a},\underline{b})^{[k)}$, the adjugate hybrid with a factor $|\underline{b}|^{n-k-1}$. . The hybrid determinants are defined as the determinants of hybrid matrices, and they are given by

$$(3) \qquad |(\underline{b},\underline{a})^{[k)}| = |\underline{b}|^{\binom{n-1}{k-1}}|\underline{a}|^{\binom{n-1}{k}}.$$

By §8, the compounds and adjugates of hybrids are related by the identity

$$(4) \qquad (\underline{b},\underline{a})^{[p)(q)} = |\underline{a}|^{q-\binom{n-1}{p}}|\underline{b}|^{q-\binom{n-1}{p-1}}(\underline{a},\underline{b})^{[p)[q]}.$$

This is the theorem of Reiss by which any minor in a hybrid compound is the complementary minor in the hybrid adjugate, except for a scalar factor. By giving $\underline{a},\underline{b}$, p and q special values, various theorems, associated mainly with Cauchy, Jacobi, Sylvester, Francke, Picquet, Bazin and Reiss,are obtained.*(18a) For example, the identities in §8 may be returned to by taking $\underline{a}=\underline{I}$, and $\underline{b}=\underline{I}$. The second of these identities is Francke's theorem on the minors of a compound determinant; and the first, with p=I, is Jacobi's theorem on the minors of the adjugate. Either imply both the identities (4) of §4, of which the first is Sylvester's theorem on compound determinants, and the second, with k=I, is Cauchy's theorem on the adjugate determinant.

I5. Reversal relations.

If $\underline{S}=(S_1,\ldots,S_N)$ is any series of N terms, let $[\underline{S}]=(S_N,\ldots,S_1)$, called the reverse of S, denote the series $\underline{S}$ taken in its reverse order. For an index i having the series N as its range, so that S_i defines a term of $\underline{S}$, the corresponding reverse index is defined by $[i]=N-i+1$; so that the ith term of $[S]$ is $[S]_i=S_{[i]}$.

Proposition (i). $\underline{n}^{[k]}=[\underline{n}^{(n-k)}]$; that is, $n_i^{[k]}=n_{[i]}^{(n-k)}$, where $i=\binom{n}{k}-i+1$.

This just restates the proposition (iii) of §2.

Proposition (ii). $\epsilon_{ij}^{[k]}=\epsilon_{[ij]}^{(k)}$.

Here $[ij]$ denotes the double reverse index $[i][j]$; and other definitions involved are made in §2. We have

$$\begin{aligned}\epsilon_i^{(k)}=\alpha(\pi_i^{(k)})&=\alpha((\pi_i^{(k)}\pi_i^{[k]})/\underline{n})\\&=\alpha((\pi_{[i]}^{[n-k]}\pi_{[i]}^{(n-k)})/n)\\&=(-1)^{k(n-k)}\alpha((\pi_{[i]}^{(n-k)}\pi_{[i]}^{[n-k]})/\underline{n})\\&=(-1)^{k(n-k)}\alpha(\pi_{[i]}^{(n)})=(-1)^{k(n-k)}\epsilon_{[i]}^{[k]}\ ;\end{aligned}$$

which is

$$(1)\qquad \epsilon_i^{(k)}=(-1)^{k(n-k)}\epsilon_{[i]}^{[k]}\ ,$$

and hence the proposition.

If $\underline{\alpha}$ is a square matrix whose order is a combinatorial $N=\binom{n}{k}=\binom{n}{n-k}$, two further matrices, which are denoted by $\alpha^{[k]}$, $[k]\alpha$ and called the k-signed reverse

and the complementary k-signed reverse respectively, can be obtained by reversing the order of the rows and columns of $\underset{\sim}{\alpha}$ and attaching certain signs to the elements. Thus $(\kappa[\underset{\sim}{\alpha}]\kappa)_{ij} = \epsilon_{ij}^{(\kappa)}\alpha_{[ij]}$. Similarly, from a column vector $\underset{\sim}{\xi}$ of N elements, there may be derived various vectors, say $(\kappa[\underset{\sim}{\xi})$, or $(\underset{\sim}{\xi}'\kappa]$; for example, $(\kappa[\underset{\sim}{\xi})_i = \epsilon_i^{(\kappa)}\xi_{[i]}$ and $(\underset{\sim}{\xi}'\kappa] = (\kappa[\underset{\sim}{\xi})'$.

Proposition (iii). $\underset{\sim}{\xi}'[\kappa[\underset{\sim}{\alpha}]\kappa]\underset{\sim}{\xi} = (\underset{\sim}{\xi}'\kappa]\underset{\sim}{\alpha}(\kappa[\underset{\sim}{\xi})$.

This follows from the definitions, and proposition (ii).

Proposition (iv). If $|\underset{\sim}{x}\ \underset{\sim}{y}|$ denotes a determinant of order m+n partitioned at its mth column, then $|\underset{\sim}{x}\ \underset{\sim}{y}| = \underset{\sim}{x}^{(m)}{}'(n[\underset{\sim}{y}^{(n)}$

In the determinant $|\underset{\sim}{x}\ \underset{\sim}{y}|$, $x_i^{(m)}$ and $\epsilon_{ij}^{(m)}y_{[j]}^{(n)}$ are cofactors, so that the proposition is an expression of Laplace's theorem, for the expansion of a determinant by the minors on any set of columns.

Proposition (v). $\underset{\sim}{a}^{[k]\prime} = (\kappa[\underset{\sim}{a}^{(n-\kappa)}]k)$.

This shows the explicit relation between the compounds and adjugates of a matrix, and it can be seen from the definitions and the proposition (iii) of §2. Another statement of this relation we shall require is as follows.

Proposition (vi). $\underset{\sim}{a}^{[r]}{}_{i}' = \epsilon_i^{(k)}\, \mathrm{cof}_k\left(\underset{\sim}{a}^{(n-r)}_{[i]}\right)$.

For we have

$$\begin{aligned}\underset{\sim}{a}^{[r]}{}_{i}' &= \epsilon_i^{(k)}\, \mathrm{cof}_k\left[\left(\underset{\sim}{a}_1^{r]}\right)^{(n-r)}\right)\\ &= \epsilon_i^{(k)}\, \mathrm{cof}_k\left[\left(\underset{\sim}{a}^{n-k)}_{[i]}\right)^{(n-r)}\right)\\ &= \epsilon_i^{(k)}\, \mathrm{cof}_k\left[\underset{\sim}{a}^{(n-k)}_{[i]}\right).\end{aligned}$$

Proposition (vi). If $\underset{\sim}{a}$ is of order mxn, then

$$\left(\begin{matrix}\underset{\sim}{I}\\ \underset{\sim}{a}\end{matrix}\right)^{(n)} = \mathrm{cof}_m\left[\left(\begin{matrix}-\underset{\sim}{a}'\\ \underset{\sim}{I}\end{matrix}\right)^{(m)}\right)$$

Consider the determinant $\left|\begin{matrix}\underset{\sim}{I} & -\underset{\sim}{a}'\\ \underset{\sim}{a} & \underset{\sim}{I}\end{matrix}\right|$. By proposition (iv) it is equal to

$$\left(\begin{matrix}\underset{\sim}{I}\\ \underset{\sim}{a}\end{matrix}\right)^{(n)} \mathrm{cof}_m\left[\left(\begin{matrix}-\underset{\sim}{a}'\\ \underset{\sim}{I}\end{matrix}\right)^{(m)}\right),$$

this being its expression as a Laplace expansion; and then, by the double-Laplace expansion theorem of §7, it is also

$$\sum_k \mathrm{trace}\left(\underset{\sim}{a}'\underset{\sim}{a}\right)^{(k)} = \sum_{i,j,k} \left(a_{ij}^{(k)}\right)^2 .$$

A typical element of $\left(\begin{matrix}\underset{\sim}{I}\\ \underset{\sim}{a}\end{matrix}\right)^{(n)}$ is $\epsilon a_{ij}^{(k)}$, where $\epsilon = \pm I$. By comparison of the above expansions, the corresponding element in $\mathrm{cof}_m\left[\left(\begin{matrix}-\underset{\sim}{a}'\\ \underset{\sim}{I}\end{matrix}\right)^{(m)}\right)$ is $\epsilon' a_{ij}^{(k)}$, where $\epsilon\epsilon' = I$; and this establishes the proposition.

I6. Reciprocity of bordered determinants.

Now we have them means of giving a formal proof of the following theorem, which shows a reciprocal relation in a set of pairs of bordered determinants.*(19)

THEOREM. If matrices $\underline{a}$, $\underline{x}$ and $\underline{y}$, of order n, have reciprocals $\underline{b}$, $\underline{u}$ and $\underline{v}$ respectively, then

$$\begin{vmatrix} \underline{a} & \underline{x}^{(r)} \\ \underline{y}_1^{(k)} & \underline{0} \end{vmatrix} = w \begin{vmatrix} \underline{b} & \underline{v}_1^{[k]} \\ \underline{u}_1^{[k]} & \underline{0} \end{vmatrix},$$

for $k=0,1,\ldots,n$, where $w=(-1)^{\frac{1}{2}n(n-1)}|\underline{x}||\underline{a}||\underline{y}|$.

By §13, we have

$$\begin{vmatrix} \underline{a} & \underline{x}_1^{(k)} \\ \underline{y}_1^{(r)} & \underline{0} \end{vmatrix} = (-1)^k \underline{y}_{(1}^{(k)} \underline{a}^{[k]} \underline{x}_{1)}^{(k)}.$$

But

$$\underline{a}^{[k]} = |\underline{a}|\,(\underline{a}^{(k)})^{-1} = |\underline{a}|\,(\underline{a}^{-1})^{(k)} = |\underline{a}|\,\underline{b}^{(k)},$$

and similarly for $\underline{x}$ and $\underline{y}$; so that substitution leads to the expression

$$(-1)^k |\underline{x}||\underline{a}||\underline{y}|\, \underline{v}_{(1}^{[k]} \underline{b}^{(r)} \underline{u}_{1)}^{[k]}.$$

By transposing, and putting $w=(-1)^k |\underline{x}||\underline{a}||\underline{y}|$, this gives

$$(-1)^{n-k} w (\underline{u}^{[k]\prime})_{(1} \underline{b}^{(r)\prime} (\underline{v}^{[r]\prime})_{1)}.$$

The proposition (iv) of §15 may now be used, to obtain

$$(-1)^{n-k} w (\underline{u}_{(1}^{(n-k)}]_k) \underline{b}^{(k)}{}_{1)}{}^{(r}[\underline{v}_{1)}^{(n-k)}) \;;$$

and, reversing the order of the summation in this expression, and transferring the sign-factors, this transforms, by the proposition (i) of §15, to

$$(-1)^{n-k} w\; \underline{u}_{(1}^{(n-k)} [r|\underline{b}^{(r)}{}_{1)}{}^{|k]}\; \underline{v}_{1)}^{(n-k)}.$$

Proposition (ii) of §13 then gives

$$(-1)^{n-k} w (\underline{u}_1^{[k]})^{(n-k)} \underline{b}^{(r)}{}_{1)}{}^{k} (\underline{v}_1)^{(n-k)};$$

which, by § I3, is

$$w \begin{vmatrix} \underset{\sim}{b} & \underset{\sim}{v}_{,}^{[k]} \\ \underset{\sim}{u}_{,}^{[k} & \underset{\sim}{0} \end{vmatrix}.$$

Thus the theorem is proved.

It may be seen from this theorem that if $\underset{\sim}{X}'\underset{\sim}{a}\underset{\sim}{X}$ is an n-ary quadratic form, with reciprocal form $\underset{\sim}{U}'\underset{\sim}{b}\underset{\sim}{U}$, and if $\underset{\sim}{p}$ is the coordinate vector of a linear k-space with dual coordinate vector $\underset{\sim}{q}$, then

$$\underset{\sim}{p}'\underset{\sim}{a}^{[k]}\underset{\sim}{p} = (-I)^{n} |\underset{\sim}{a}| c^{2} \underset{\sim}{q}'\underset{\sim}{b}^{[n-k]}\underset{\sim}{q} ,$$

where c is the proportionality factor which appears in Grassmann's duality theorem applied to $\underset{\sim}{p}$ and $\underset{\sim}{q}$.

COROLLARY.[*(20)]

$$\left|\underset{\sim}{x}_{,}^{(k}\underset{\sim}{a}\underset{\sim}{y}_{,}^{k)}\right| = \begin{cases} (-I)^{k} |\underset{\sim}{b}|^{-1} \begin{vmatrix} \underset{\sim}{b} & \underset{\sim}{y}_{,}^{k)} \\ \underset{\sim}{x}_{,}^{(k} & \underset{\sim}{0} \end{vmatrix} \\ (-I)^{n-k} |\underset{\sim}{u}|^{-1} |\underset{\sim}{v}|^{-1} \begin{vmatrix} \underset{\sim}{a} & \underset{\sim}{u}_{,}^{k]} \\ \underset{\sim}{v}_{,}^{[k} & \underset{\sim}{0} \end{vmatrix} \end{cases}$$

This follows by a combination of the present theorem with theorem I, §9.

I8. Coordinates of linear spaces.

If $\mathfrak{X}$ is any linear n-vector space of rank r, we may form an nxr matrix $\underset{\sim}{P}$ whose columns are a maximal linearly independent set of vector of $\mathfrak{X}$, and then determine the $\binom{n}{r}$-vector $\underset{\sim}{p}=\underset{\sim}{P}^{(r)}$, the ratios of whose elements are a set of coordinates of $\mathfrak{X}$.*(21) Grassmann's theorem, that linear spaces are uniquely determined by their coodinate vectors, is a consequence of the proposition: If $\underset{\sim}{P}$ and $\underset{\sim}{Q}$ are of order nxr, and if $\underset{\sim}{P}^{(r)}=\underset{\sim}{Q}^{(r)}$, then $\underset{\sim}{Q}=\underset{\sim}{P}\underset{\sim}{u}$ where $|\underset{\sim}{u}|=1$. A proof of this is as follows. Without loss in generality, suppose $\underset{\sim}{P}=\begin{pmatrix}\xi\\ x\end{pmatrix}$ and $\underset{\sim}{Q}=\begin{pmatrix}\eta\\ y\end{pmatrix}$ where $|\xi|=|\eta|\neq 0$. Then $\underset{\sim}{M}\underset{\sim}{P}=\begin{pmatrix}\xi\\ 0\end{pmatrix}$, $\underset{\sim}{M}\underset{\sim}{Q}=\begin{pmatrix}\eta\\ \sigma\end{pmatrix}$, where $\underset{\sim}{M}=\begin{pmatrix}I & 0\\ -x\xi^{-1} & I\end{pmatrix}$ and $\sigma=y-x\xi^{-1}\eta$. We have $(\underset{\sim}{M}\underset{\sim}{P})^{(r)}=\underset{\sim}{M}^{(r)}\underset{\sim}{P}^{(r)}=\underset{\sim}{M}^{(r)}\underset{\sim}{Q}^{(r)}=(\underset{\sim}{M}\underset{\sim}{Q})^{(r)}$, and hence $\begin{pmatrix}\xi\\ 0\end{pmatrix}^{(r)}=\begin{pmatrix}\eta\\ \sigma\end{pmatrix}^{(r)}$. This implies that the determinants on $r-1$ rows of η with one row of σ all vanish; or $\sigma\eta^{*}=0$, where η^{*} is the adjugate of η, regular since η is regular. It follows that $\sigma=0$, so that $\underset{\sim}{Q}=\underset{\sim}{P}\underset{\sim}{u}$, where $\underset{\sim}{u}=\xi^{-1}\eta$.

Now let $\underset{\sim}{x}$ and $\underset{\sim}{u}$ be reciprocal matrices, so that $\underset{(i}{x}\,\underset{j)}{u}=\underset{(i}{u}\,\underset{j)}{x}=\delta_{ij}$. Then $\underset{\sim}{x}_{1}^{(k)}$ and $\underset{\sim}{u}_{1}^{[k}$ are each of full rank, and such that $\underset{\sim}{u}_{1}^{[k}\,\underset{\sim}{x}_{1}^{k)}=0$. The columns of $\underset{\sim}{x}_{1}^{(k)}$ are a maximal independent set which is annihilated with the rows of $\underset{\sim}{u}_{1}^{[k}$; and the rows of $\underset{\sim}{u}_{1}^{[k}$ are a maximal independent set which is annihilated with the columns of $\underset{\sim}{x}_{1}^{(k)}$. Thus the columns of $\underset{\sim}{x}_{1}^{(k)}$ and the rows of $\underset{\sim}{u}_{1}^{[k}$ span dual linear spaces. Hence $\underset{\sim}{p}=(\underset{\sim}{x}_{1}^{(k)})^{(k)}=\underset{\sim}{x}_{1)}^{(k)}$ and $\underset{\sim}{q}=(\underset{\sim}{u}_{1}^{[k}\,)^{(n-k)}=\underset{[1]}{u}^{(n-k)}$ are a typical pair of dual coordinate vectors; also we have ~~that they~~ have the rela-

tion

$$\underset{\sim}{x}^{(k)}_{(i)} = |\underset{\sim}{u}|^{-1}\,\underset{\sim}{u}^{[k]}_{(i)} = |\underset{\sim}{u}|^{-1}\left(\underset{\sim}{u}^{(m-k)}_{[i]}\right)^{(k)}\,,$$

by (vi), I5; or $p_i = |\underset{\sim}{u}|^{-1} c_i^{(k)} q_{[i]}$, which shows the proposition: the ith elements of p, and the reverse ith elements of q with the signe $c_i^{(k)}$, are proportional. This is a statement of Grassmann's duality theorem for dual coordinates or linear spaces. As an alternative procedure we have a typical pair of dual coodinate given by $\underset{\sim}{p} = \begin{pmatrix} \underset{\sim}{I} \\ \underset{\sim}{a} \end{pmatrix}^{(k)}$, $\underset{\sim}{q} = \begin{pmatrix} -\underset{\sim}{a}' \\ \underset{\sim}{I} \end{pmatrix}^{(m)}$; and the duality theorem, which is that $\underset{\sim}{p} = (m\lceil\underset{\sim}{q})$, is then shown in proposition (vi), § II.

I8. Decomposition of matrices.

The fundamental theory of matrix rank follows easily from the formula to be given now.:

THEOREM I. *If* $a^{(r)}_{\lambda k}\neq 0$, *but* $\underset{\sim}{a}^{(r+1)}=\underset{\sim}{0}$, *then* $\underset{\sim}{a}=\underset{\sim}{a}^{r)}_{k}\,\underset{\sim}{a}^{(r)-1}_{\lambda k}\,\underset{\sim}{a}^{(r}_{\lambda}$.

If $\underset{\sim}{a}^{(r+1)}=\underset{\sim}{0}$, then the minors of $\underset{\sim}{a}$ of order r+I, in particular those which contain the non-vanishing minor $a^{(r)}_{\lambda k}$, all vanish. Thus, by the Cauchy determinant expansion, §II, we have, for each i,j

$$0=\begin{vmatrix} \underset{\sim}{a}^{(r)}_{\lambda k} & \underset{\sim}{a}^{(r1)}_{\lambda_j} \\ \underset{\sim}{a}^{(1r)}_{ik} & a_{ij} \end{vmatrix} = a^{(r)}_{\lambda k}\left(a_{ij}-\underset{\sim}{a}^{(r)}_{ik}\,\underset{\sim}{a}^{(r)-1}_{\lambda k}\,\underset{\sim}{a}^{(r1)}_{\lambda_j}\right) ;$$

then the formula follows immediately.

This formula shows explicitly the linear dependence of the rows and columns of $\underset{\sim}{a}$ on its hth set of r rows and kth set of r columns respectively.

We note that if $\underset{\sim}{a}$ has a maximal linearly independent set of r columns, say the kth set $a^{r)}_{\lambda k}$, then $0\neq(\underset{\sim}{a}^{r)}_{\lambda})^{(r)}=\underset{\sim}{a}^{(r)}_{\lambda)}$, or, by theorem I, we should be able to express all the columns of $\underset{\sim}{a}^{r)}_{\lambda k}$ as linear combinations of a subset. Thus $\underset{\sim}{a}^{(r)}\neq\underset{\sim}{0}$, by the linear independence of r columns. But every set of r+I columns are linearly dependent; that is, for each l, we have $\underset{\sim}{a}^{r+1)}_{l}\underset{\sim}{X}=\underset{\sim}{0}$ for some vector $\underset{\sim}{X}\neq 0$. Elimination of $\underset{\sim}{X}$ then gives $\underset{\sim}{0}=(\underset{\sim}{a}^{r+1)}_{l})^{(r+1)}=\underset{\sim}{a}^{(r+1)}_{l)}$, for each l. Thus $\underset{\sim}{a}^{(r+1)}=\underset{\sim}{0}$, by the linear dependence of every r+I columns. Further, suppose $\underset{\sim}{a}$, of order mxn,

is a product $\underset{\sim}{\gamma}\underset{\sim}{\rho}$ of matrices $\underset{\sim}{\gamma}, \underset{\sim}{\rho}$ of order mxr, rxn respectively. Then, by the multiplication theorem for compounds, §7, we have $\underset{\sim}{a}^{(r+1)}=0$. If, moreover, $\underset{\sim}{\gamma}$ has its r columns independent, and $\underset{\sim}{\rho}$ its r rows, so that $\underset{\sim}{\gamma}^{(r)}\neq 0$ and $\underset{\sim}{\rho}^{(r)}\neq 0$, then we have $\underset{\sim}{a}^{(r)}=\underset{\sim}{\gamma}^{(r)}\underset{\sim}{\rho}^{(r)}\neq 0$, again by the multiplication theorem.

Now we see the equivalence of the following five conditions, applied to any matrix $\underset{\sim}{a}$:

$$\underset{\sim}{a}^{(r)}\neq 0,\quad \underset{\sim}{a}^{(r+1)}=\underset{\sim}{0}.$$

An r-th order minor of $\underset{\sim}{a}$ is non-vanishing, while every minor of order r+1 containing it vanishes.

$\underset{\sim}{a}$ has a maximal independent set of r rows.

$\underset{\sim}{a}$ has a maximal independent set of r columns.

$\underset{\sim}{a}$ is a product of a pair of matrices with r independent rows and columns respectively.

Any matrix which satisfies these conditions will be said to be of rank r. A matrix which has either all its rows independent or all its columns independent is said to be non-degenerate, or of full rank; otherwise it is said to be degenerate. We have that the mxn matrices of degenerate rank r are the products of pairs of matrices of order mxr, rxn, of full rank. Also we have the following:

THEOREM II.*(22) If the rank r of any matrix $\underset{\sim}{a}$ be partitioned into a sum $\sum_{\rho} r_{\rho}$, where $r_{\rho}>0$, then the matrix $\underset{\sim}{a}$ can be

Partitioned correspondingly into a sum $\sum_\nu \underline{a}_\nu$ of linearly independent matrices $\underline{a}_\nu$, of rank r_ν respectively, such that the rows and columns of each $\underline{a}_\nu$ belong to the row space and the column space of $\underline{a}$ respectively.

The expression $\underline{a}=\underline{\gamma}\underline{\rho}$ for a can be partitioned in a way corresponding to the partitioning of r as $\sum_\nu r_\nu$. Thus

$$\underline{a}=\underline{\gamma}\underline{\rho}=(\underline{\gamma}_1,\ldots,\underline{\gamma}_N)\begin{pmatrix}\underline{\rho}_1\\ \vdots\\ \underline{\rho}_N\end{pmatrix}=\sum_\nu \underline{\gamma}_\nu\underline{\rho}_\nu=\sum_\nu \underline{a}_\nu\ ,$$

where $\underline{a}_\nu=\underline{\gamma}_\nu\underline{\rho}_\nu$ is of rank r_ν. If the $\underline{a}_\nu$ were linearly dependent, there would be a non-null diagonal matrix $\underline{\Lambda}$ such that $\underline{\gamma}\underline{\Lambda}\underline{\rho}=\underline{0}$. Since $\underline{\rho}$ is of full rank, this would imply $\underline{\gamma}\underline{\Lambda}=\underline{0}$; and since $\underline{\gamma}$ is of full rank, this in turn would imply $\underline{\Lambda}=\underline{0}$, contradicting the hypothesis that $\underline{\Lambda}\neq\underline{0}$. Hence the $\underline{a}_\nu$ are linearly independent.

THEOREM III. *If $\underset{\sim}{a}$ is of full rank r, and $s \leqslant r$, then $\underset{\sim}{a}^{(s)}$ is of full rank $\binom{r}{s}$.*

For then there is a regular submatrix $\underset{\sim}{\alpha}$ of $\underset{\sim}{a}$ of the maximum order r, and then $\alpha^{(s)}$ is a regular submatrix of $a^{(s)}$ of the maximum order $\binom{r}{s}$. Thus $\underset{\sim}{a}^{(s)}$ is of full rank $\binom{r}{s}$.

THEOREM IV. *If $\underset{\sim}{a}$ is of rank r, then $\underset{\sim}{a}^{(s)}$ is of rank $\binom{r}{s}$ if $s \leqslant r$, and of zero rank if $s > r$.*[*(23)]

For then $\underset{\sim}{a}=\underset{\sim}{\gamma}\underset{\sim}{\rho}$, where $\underset{\sim}{\gamma}$, $\underset{\sim}{\rho}$ are of full rank r; so that $\underset{\sim}{a}^{(s)}=\underset{\sim}{\gamma}^{(s)}\underset{\sim}{\rho}^{(s)}$ is of rank $\binom{r}{s}$, if $s \leqslant r$, and null if $s > r$, for $\underset{\sim}{\gamma}^{(s)}$, $\underset{\sim}{\rho}^{(s)}$ are of full rank $\binom{r}{s}$ when $s \leqslant r$, and are null when $s > r$.

THEOREM V. *If $\underset{\sim}{a}$ is Hermitian, of rank r, then $\underset{\sim}{a}=\bar{\underset{\sim}{\beta}}'\underset{\sim}{\alpha}\underset{\sim}{\beta}$, where $\underset{\sim}{\alpha}$ is regular Hermitian of order r, and $\underset{\sim}{\beta}$ is of full rank.*

This follows from theorem I, and enables an n-ary Hermitian form $\underset{\sim}{X}'\underset{\sim}{a}\underset{\sim}{X}$ of degenerate rank r to be expressed as a non-degenerate r-ary Hermitian: thus $\underset{\sim}{X}'\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{\xi}'\underset{\sim}{\alpha}\underset{\sim}{\xi}$, where $\underset{\sim}{\xi}=\underset{\sim}{\beta}\underset{\sim}{X}$

THEOREM VI. *If $\underset{\sim}{a}$ is Hermitian of rank I, then $a=c\underset{\sim}{\theta}\bar{\underset{\sim}{\theta}}'$, where $\underset{\sim}{\theta}$ is a column vector, and $c=\pm I$.*[*(24)]

We have $c=\alpha/|\alpha|^{-1}$, and $\bar{\theta}'=|\alpha|^{\frac{1}{2}}\underset{\sim}{\beta}$.

NOTES

p.

1. I. Cauchy (I).

4 2. It may be noted that for the conventional lexicographical order, or the order of first differences, the corresponding formula is $i=\binom{n}{k}-\sum_{s=1}^{\lambda}\binom{n-i_s}{k-s+1}$. Either of these formulae incidentally establish the combinatorial identity $\binom{n}{k}=1+\sum_{s=1}^{\lambda}\binom{n-s}{k-s+1}$.

6 3. Lederrmann (2), 74.

6 4. This permutation belongs to a class of determinantal permutations; see Turnbull (3),37.

8 5. The term'compound' in the present conection originates historically in the study of compound determinants, or determinants whose elements are themselves determinants. The word now is less satisfactorily taken to refer to a matrix whose elements are determinants. Possibly a better terminology is, with any matrix, to have various derived matrices, the 'systemes derivées' of Cauchy. The usual terminology is used here, however. See Aitken(4 ch.V.

8 6. Aitken(4), 91.

9 7. See Franke (22). This is Sylvesters theorem on compound determinants.

12 8. Aitken (4), 86.

12 8a. Weyl (6), 4.

12 9. See Muir (7), Vol.I, ch.iv, where this theorem is discussed at length, and the original papers are quoted: Binet (5) and Cauchy (I).

13 I0. Aitken (4), I01. Then second of these identities is a statement of Franke's theorem on the minors of the compound of a matrix.

13 I0a. Todd (I0), I05.

14 II. Weyl (6), 4. Aitken (4), I02.

16 I2. Cauchy (I).

16 I3. See Mirlat (9) for an application of this formula.

16 I4. Todd (I0), I05.

17 I5. Some considerations approaching this result are apparent in Kowalewski (I).

18 I6. Aitken (4), 8I.

19 I7. Muir attributes a statement which is essentially this result to Arnaldi (i3). See Muir (I3), 295, and Aitken (4), ch.iv.

20 I8. Turnbull (3), I07-9 and Aitken (4), I0I-3.

25 I9. c.f. Turnbull , where perhaps the present result is intended, though for a matrix appearing there we would have its transpose.

26 20. See Mirlat (I4) for a bearing of this result. This formula leads to the further condition

$$(-1)^{n-m}\begin{vmatrix} \underset{\sim}{A} & \underset{\sim}{U}_{n-m} \\ \underset{\sim}{U}'_{n-m} & \underset{\sim}{0} \end{vmatrix} > 0 \quad (m=1,\ldots,k)$$

for $\underset{\sim}{X}'\underset{\sim}{A}\underset{\sim}{X}>0$ on $\mathfrak{X}_k$, where $\underset{\sim}{U}_{n-k}$ is a base for the space $\mathcal{U}_{n-k}$. This gives $\underset{\sim}{X}'\underset{\sim}{A}\underset{\sim}{X} > 0$ on $\mathfrak{X}_k$ if and only if

$$\underset{\sim}{q}'_{n-m}\ \underset{\sim}{A}^{[n-m]}\ \underset{\sim}{q}_{n-m} > 0 \quad (m=1,\ldots,k)$$

where $\underset{\sim}{q}_{n-m}$ is a dual coordinate vector for $\mathfrak{X}_m$.

27. 21. Hodge and Pedoe (I6), Vol.I, ch viii.

30. 22. cf. Ghosh (I7), where it is proved that if $\underset{\sim}{M}$ is mxn, of rankr r, then $\underset{\sim}{M} = \sum_{\lambda=1}^{r} \underset{\sim}{C}^{\lambda} \underset{\sim}{R}_{\lambda}$, where $\underset{\sim}{C}^{\lambda}$, $\underset{\sim}{R}_{\lambda}$ are column and row vectors belonging to the column and row spaces of $\underset{\sim}{M}$.

32 23. Niccoletti (I8).

32 24. See Afriat (2I) for an application of this result.

REFERENCES

I. A. Cauchy, Journ. de l'Ec. Polyt., x, cah.I7(I8I2), 29-II2.

2. W. Ledermann, Theory of Finite Groups (Edinburgh I949), 74.

3. H. W. Turnbull, The Theory of Determinants, Matrices, and Invariants (London I945).

4. A. C. Aitken, Determinants and Matrices (Edinburgh I942).

5. Binet, Journ. de l'Ec. Polyt., ix, cah.I6 (I8I2)? 280-302.

6. H. Weyl, The Classical Groups (Princeton I939), 4.

7. T. Muir, Theory of determinants in the historical order of development (London I906).

8. J. M. H. Wedderburn, Lectures on Matrices (Amer. Math. Soc. Colloquem Publ.,XVIII, I934)

9. J. Plemelj, Mh. Math. Phys. Vol.I2 (I90I), 82-96).

9. S. N. Arriat, An iterative process for the numerical determination of characteristic values of certain matrices, Quart.J.Math.Oxford(2),2(I95I),I2I-2.

I0. J. A. Todd, Projective and Analytical Geometry (London I947), I05.

II. Kowalewski, Einführung in die Determinantentheorie (Leipsig I909), 89-99.

I2. T. Muir, Contributions to the History of Determinants (London I930), 295.

13. Arnaldi, Giornale di Mat., xxxiv (1896) 209-14.

14. S.N. Afriat, The quadratic form positive definite on a linear manifold, Proc. Camb. Phil. Soc., Vol. 47, Pt.I.

15.

16. W.V.D. Hodge and D. Pedoe, Methods in Algebraic Geometry (Cambridge 1947), Vol.I, ch.viii.

17. N.N. Ghosh, On a new reduction theorem of matrices, Bull. Calcutta Math. Soc., 37(1945), 33-36.

18. O. Niccoletti, Atti Accad. Sci. Torino, Vol.37 (1901-1902), 655-659.

19. E. Pascal, Repertorium der Hohere Mathematik, Vol.I 138-146.

20. S.N. Afriat, The rank and multiplicity theorem for the reduction of quadratic forms, Math. Gazette,

21. H. W. Turnbull and A.C. Aitken, The Theory of Canonical matrices (London 1932).

22. E. Franke, J.reine angew. Math. Vol.61(1863), 350-356.

23. J.Williamson, Proc.Edinburgh Math.Soc. II,vol.2(1929) 240-251.

COMMUTATIVE MATRICES, LATENT VECTORS AND CHARACTERISTIC VALUES

contents

I Introduction

2 Vector spaces and matrices

3 Latent vectors and characteristic sets

4 Rational functions

5 Triangular unitary transforms

6 Rank, multiplicity and index

7 The independence of axial spaces

8 Idempotents

9 The resolvents

IO Commutative matrix algebras

II The reduction of a commutative algebra

I2 Nullpotents

I3 The normal form for a matrix

I4 Projectors and principal spaces

I. Introduction.

An historical interest is attached to the algebra of matrices through its close resemblance to ordinary algebra with the contrast of the non-commutivity of its multiplication. It gave a significant, and simple generalization of what was understood as algebra, and contributed to the development of abstract algebra, where it preserves a central position. A certain importance, arising in this way, and also in other more material connections, centres on the condition of commutivity between matrices. Moreover, there are simplicities in commutative matrix algebra over the complex field, and analogies to the simpler algebra of its ground field, which it is our object to set out here. We consider matrices over the complex field, which in this case is called the field of scalars. They are coordinate with the endomorphisms of a vector space, which is referred to as the fundamental vector space. Matrices are unlike the scalars in that they do not generally commute in their multiplications, that is to say as operators throughout the fundamental vector space. However if any matrices resemble the scalars in that they do all commute, so that they satisfy all the axioms of the scalar algebra and thus their substitution in any algebrain scalar identity would yield a commutative

matrix identity, it appears further that together they will effectively be a set of scalars when their operation is restricted to certain subspaces of the fundamental vector space. An equivalent statement is that they have certain axial or latent vectors in common. Any of their characteristic values which belong to a common axial vector are said to be latently associated, and the association to be through that vector. The maximal subspaces on which a commutative matrix set is effectively some scalar set define the axial spaces of that matrix set. The axial spaces are shown to be all independent. The respective scalar sets, formed by latently associated characteristic values of the matrices, define the corresponding characteristic sets. We show an extent of the analogy which exists between processes with scalars and with commutative matrices in which these sets of characteristic values of the matrices have a fundamental role. Also there is discussion of commutative matrix sets in such a way as to generalize the familiar theory of latent vectors and characteristic values of a single matrix. Two fundamental results are proved: that every commutative matrix set has at least one characteristic set, and that the characteristic sets of any commutative matrix set can be extended to give the characteristic set of any commutative extension of that set. Also it is proved that the matrices of any commutative set can all

be reduced to triangular form by the same unitary transformation, and the sets of similarly situated diagonal elements of the triangular transforms are the characteristic sets with certain repetitions. Through these repetitions the multiplicity of a characteristic set is defined so as to generalize the concept of the multiplicity of a characteristic value of a single matrix as its multiplicity as a root of the characteristic equation. The rank of a characteristic set is defined as the rank of the corresponding axial space. It is shown that the rank of a characteristic set is at most its multiplicity. The index of a characteristic set is defined in a way which generalizes the concept of the index of a characteristic value of a single matrix as its multiplicity as a root of the minimum equation, and it is shown that no characteristic set has zero index. Also it is proved that the characteristic values of any rational function of commutative matrices are given by the values of that function when the argument ranges over the characteristic sets, and the multiplicity of any characteristic value is determined as the sum of the multiplicities of the characteristic sets which yield that value.

An idempotent is a matrix $\underset{\sim}{e}$ such that $\underset{\sim}{e}^2=\underset{\sim}{e}$; and idempotents $\underset{\sim}{e}$, $\underset{\sim}{f}$ are independent if $\underset{\sim}{e}\underset{\sim}{f}=\underset{\sim}{f}\underset{\sim}{e}=\underset{\sim}{0}$. Any idempotents $\underset{\sim}{e}_n$ form a reduction of an idempotnet $\underset{\sim}{e}$

if they are independent idempotents such that $\sum_n e_n = e$. In any commutative matrix algebra K over the complex field which has the unit matrix for unit, there are independent idempotents e_n which are irreducible in K and which form a reduction of the unit matrix. They are called the principal idempotents in K, or the resolvents of K. There are subalgebras K_n in K, the principal subalgebras in K, in which the units are e_n respectively, such that $K_n . K_\rho = 0$ $(\rho \neq n)$, and $K = \sum_n K_n$. Each of these algebras is congruent to the scalar algebra, modulo a nullpotent algebra. Thus for any matrix a in K the matrix ae_n has just one distinct characteristic value α_n. Moreover if a ranges in any set s in K, then α_n describes the characteristic sets α_n of s; and there are multiplicities m_n such that the multiplicity of any σ as a characteristic set of s is given by $m_\sigma = \sum_{\sigma_n = \sigma} m_n$. It follows, by the Cayley-Hamilton theorem, that $\prod_n (a - \alpha_n I)^{m_n} = 0$ for every matrix a in K. Let i_n be least exponents such that $\prod_n (a - \alpha_n I)^{i_n} = 0$ for all a in K. Then $0 < i_n \leqq m_n$, and

$$(a - \alpha_n I)^r (b - \beta_n I)^s \ldots . e_n = 0, \quad r+s+\ldots \geq i_n ,$$

for all matrices a, b, ... in K. It is taken that m_n defines the multiplicity, and i_n the index of the characteristic plane n of K; and the rank r_n is defined as the rank of the corresponding axial space $\mathfrak{X}_n$ of K, which is such that $aX = X\alpha_n$ for each matrix a of K and each vector X of $\mathfrak{X}_n$. Also K determines a splitting of the

fundamental vector space into subspaces U_n, which are called the principal spaces relative to K, and are the ranges of the principal idempotents e_n, of rank m_n, respectively. Each principal space contains its corresponding axial space, or $\mathfrak{A}_n \subset U_n$ for each n; so that the rank of each plane is at most its multiplicity. The principal idempotnets in the algebra generated by a single matrix define the resolvents of that matrix, indexed by its characteristic values. If a commutative pair of matrices are such that there is a unique characteristic value of one which is latently associated with any characteristic value of the other, then we say the one matrix is resolved by the other. It is proved that if a commutative pair of matrices is such that one matrix resolves the other, then the independent idempotent resolution of the unit matrix relative to the one refines that of the other, and conversely. It appears as a consequence that a pair of commutative atrices are mutually resolved if and only if they have the same resolvents. If commutative matrices a, b, ... have resolvents f_α, g_β, ..., then $f_\alpha g_\beta \ldots \neq 0$ if and only if α, β, ... are latently associated characteristic values of a, b, ... respectively. The resolvents, the axial and the principal spaces of a commutative matrix algebra are expressed in terms of the same for any generating subset of matrices.

2. Vector spaces and matrices.

Let $\mathfrak{A}$ denote the space of complex n-vectors, called the fundamental complex vector space, of order n. A space $\mathfrak{X}$ in $\mathfrak{A}$ is defined by a set of vectors $\mathfrak{X} \subset \mathfrak{A}$ such that $\lambda\underset{\sim}{U}+\mu\underset{\sim}{V}$ belongs to $\mathfrak{X}$ whenever vectors $\underset{\sim}{U}$, $\underset{\sim}{V}$ belong to $\mathfrak{X}$, and λ,μ are any scalars. The rank r of $\mathfrak{X}$ is the maximum number of vectors of $\mathfrak{X}$ which can be linearly independent. Any maximal linearly independent set of vectors of $\mathfrak{X}$ forms the columns of an nxr matrix $\underset{\sim}{P}$ which defines a base for $\mathfrak{X}$. It is such that the n-vector $\underset{\sim}{X}=\underset{\sim}{P}\underset{\sim}{t}$ describes the space $\mathfrak{X}$ as $\underset{\sim}{t}$ ranges over the space $\mathfrak{J}_r$ of r-vectors, and $\underset{\sim}{P}\underset{\sim}{t}=\underset{\sim}{0}$ only if $\underset{\sim}{t}=\underset{\sim}{0}$. Thus there is the isomorphism $\underset{\sim}{P}: \mathfrak{A}_r \to \mathfrak{X}$ between $\mathfrak{X}$ and $\mathfrak{A}_r$ where $\underset{\sim}{t}$ in $\mathfrak{A}_r$ has image $\underset{\sim}{X}=\underset{\sim}{P}\underset{\sim}{t}$ in $\mathfrak{X}$. We may regard $\mathfrak{A}_r$ as a parametric space for $\mathfrak{X}$, the parameter $\underset{\sim}{t}$ being relative to the base $\underset{\sim}{P}$. For equivalent bases we have the matrices $\underset{\sim}{P}\underset{\sim}{u}$, where $|\underset{\sim}{u}|\neq 0$ and the corresponding parameters are $\underset{\sim}{u}^{-1}\underset{\sim}{t}$ respectively.

The nth order matrices are coorninate with the endomorphisms of $\mathfrak{A}$. The space $\mathfrak{X}$ is invariant in a matrix $\underset{\sim}{a}$ if it contains its image under $\underset{\sim}{a}$, that is if $\underset{\sim}{a}\mathfrak{X} \subset \mathfrak{X}$; the the endomorphism $\underset{\sim}{a}: \mathfrak{A} \to \mathfrak{A}$, in which any vector $\underset{\sim}{X}$ has image $\underset{\sim}{a}\underset{\sim}{X}$, induces an endomorphism $\underset{\sim}{a}: \mathfrak{X} \to \mathfrak{X}$ on $\mathfrak{X}$. When $\underset{\sim}{a}$ gives an automorphism of $\mathfrak{X}$ it is called regular on $\mathfrak{X}$; otherwise it is singular on $\mathfrak{X}$. The image space $\mathfrak{X}=\underset{\sim}{a}\mathfrak{A}$ is the range, and the maximal space whose image is the zero space $\mathfrak{O}$ is the null-space $\mathfrak{N}$ of $\underset{\sim}{a}$ on $\mathfrak{X}$. Thus $\underset{\sim}{a}$ is

regular on $\mathfrak{X}$ if and only if $\mathfrak{N} = \mathfrak{O}$, that is if and only if $a\mathfrak{X}=\mathfrak{X}$, or equivalently when $\mathfrak{X}$ is identical with its image under $\underset{\sim}{a}$.

The minimum function $\phi_{a,\mathfrak{X}}(\zeta)$ of a matrix $\underset{\sim}{a}$ on a space $\mathfrak{X}$ which it leaves invariant is the monic polynomial of least degree which is null in $\underset{\sim}{a}$ on $\mathfrak{X}$, that is $\phi_{a,\mathfrak{X}}(\underset{\sim}{a})\mathfrak{X}=\mathfrak{O}$. Such a polynomial exists, its degree being bounded by the degree of the minimum polynomial of the matrix, with respect to the fundamental space. It is unique. For were ϕ another, then $\phi_{a,\mathfrak{X}} - \phi$ would be of degree less than the degree of $\phi_{a,\mathfrak{X}}$, there being cancellation of terms of the highest degree, and also it would be null in $\underset{\sim}{a}$ on $\mathfrak{X}$; giving a contradiction. The <u>index</u> $i_{a,\mathfrak{X}}$ of $\underset{\sim}{a}$ on $\mathfrak{X}$, is defined by the degree of $\phi_{a,\mathfrak{X}}$.

THEOREM I. <u>The minimum function of a matrix on a space which it leaves invariant divides every polynomial which is null in that matrix with respect to the space, in particular it divides the minimum function of the matrix, and the index of the matrix on the space is at most the rank of the space, so the index with respect to any vector of the space is at most the rank of the space.</u>

If $\underset{\sim}{a}\mathfrak{X} \subset \mathfrak{X}$, then $\mathfrak{X}$, which is spanned by the vectors which compose its base matrix $\underset{\sim}{P}$, contains its image space $\underset{\sim}{a}\mathfrak{X}$, which is spanned by the vectors which compose the matrix $\underset{\sim}{a}\underset{\sim}{P}$. It follows that $\underset{\sim}{a}\underset{\sim}{P}=\underset{\sim}{P}\underset{\sim}{a}^*$ for some matrix $\underset{\sim}{a}^*$

of order r, and $\underset{\sim}{a}^*$ is called the <u>value</u> of $\underset{\sim}{a}$ on $\mathfrak{X}$ relative to the base $\underset{\sim}{P}$. The endomorphism $\underset{\sim}{a}^*: \mathfrak{I}_r \to \mathfrak{I}_r$ of the parametric space of $\mathfrak{X}$ is a representation of the endomorphism $\underset{\sim}{a}:\mathfrak{X} \to \mathfrak{X}$ induced by $\underset{\sim}{a}$ on $\mathfrak{X}$; here the vector $\underset{\sim}{X}=\underset{\sim}{P}\underset{\sim}{t}$ in $\mathfrak{X}$, with parameter $\underset{\sim}{t}$, has as image under $\underset{\sim}{a}$ the vector $\underset{\sim}{X}^*=\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{a}\underset{\sim}{P}\underset{\sim}{t}=\underset{\sim}{P}\underset{\sim}{a}^*\underset{\sim}{t}=\underset{\sim}{P}\underset{\sim}{t}^*$ in $\mathfrak{X}$, with parameter $\underset{\sim}{t}^*=\underset{\sim}{a}^*\underset{\sim}{t}$ which is the image under $\underset{\sim}{a}^*$ of the parameter $\underset{\sim}{t}$. The values of $\underset{\sim}{a}$ on $\mathfrak{X}$ relative to the various bases $\underset{\sim}{P}\underset{\sim}{u}$, $|\underset{\sim}{u}|\neq 0$, of $\mathfrak{X}$, form a class of equivalent matrices $\underset{\sim}{u}\underset{\sim}{a}^*\underset{\sim}{u}^{-1}$. All these have the same minimum function ϕ^*, given, say, by the minimum function of the representative $\underset{\sim}{a}^*$, the index of which is at most its order r. Now if ϕ denotes any polynomial, then we have $\phi(\underset{\sim}{a})\underset{\sim}{P}=\underset{\sim}{P}\phi(\underset{\sim}{a}^*)$, so that $\phi(\underset{\sim}{a})$ has value $\phi(\underset{\sim}{a}^*)$ on $\mathfrak{X}$ relative to $\underset{\sim}{P}$. Further, $\phi(\underset{\sim}{a})\mathfrak{X}=0$ if and only if $\phi(\underset{\sim}{a})\underset{\sim}{P}=\underset{\sim}{0}$, which is if and only if $\phi(\underset{\sim}{a}^*)=\underset{\sim}{0}$; and, again, this is if and only if ϕ^* divides ϕ. Accordingly $\phi(\underset{\sim}{a})\mathfrak{X}=0$ if and only if ϕ^* divides ϕ. It follows that $\phi_{a,\mathfrak{X}}=\phi^*$. Further, since $\phi_a(\underset{\sim}{a})=\underset{\sim}{0}$, where ϕ_a denotes the minimum function of $\underset{\sim}{a}$, we have $\underset{\sim}{0}=\phi_a(\underset{\sim}{a})\underset{\sim}{P}=\underset{\sim}{P}\phi_a(\underset{\sim}{a}^*)$; so that $\phi_a(\underset{\sim}{a}^*)=\underset{\sim}{0}$, and ϕ^* divides ϕ_a.

Vectors whose rays are invariant in a matrix are called <u>axial</u> or <u>latent</u> vectors of that matrix. Thus if $\underset{\sim}{X}$ is a latent vector of $\underset{\sim}{a}$, we have $\underset{\sim}{a}\underset{\sim}{X}=\lambda\underset{\sim}{X}$ for some scalar α; and α is called the <u>characteristic value</u> of $\underset{\sim}{a}$ corresponding to $\underset{\sim}{X}$.

THEOREM II. *Any matrix has a latent vector in any space which it leaves invariant.*

This is to say that any invariant space contains an invariant ray. Let $\underset{\sim}{a}\mathfrak{X} \subset \mathfrak{X}$, and let $\underset{\sim}{a}^*$ be the value of $\underset{\sim}{a}$ on $\mathfrak{X}$ relative to some base $\underset{\sim}{P}$, so that $\underset{\sim}{a}\underset{\sim}{P}=\underset{\sim}{P}\underset{\sim}{a}^*$. If $\underset{\sim}{a}$ is not nullpotent on $\mathfrak{X}$, so that $\underset{\sim}{a}^*$ is not a nullpotent matrix, let α be any non-zero characteristic value of $\underset{\sim}{a}^*$. Let $\underset{\sim}{t}$ be a corresponding latent vector of $\underset{\sim}{a}^*$, so that $\underset{\sim}{a}^*\underset{\sim}{t}=\underset{\sim}{t}\alpha$; and let $\underset{\sim}{X}=\underset{\sim}{P}\underset{\sim}{t}$ be the vector in $\mathfrak{X}$ with parameter $\underset{\sim}{t}$. Then we have $\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{a}\underset{\sim}{P}\underset{\sim}{t}=\underset{\sim}{P}\underset{\sim}{a}^*\underset{\sim}{t}=\underset{\sim}{P}\underset{\sim}{t}\alpha=\underset{\sim}{X}\alpha$, showing the non-null vector $\underset{\sim}{X}$ of $\mathfrak{X}$ to be a latent vector of $\underset{\sim}{a}$, with characteristic value α. If, however, $\underset{\sim}{a}$ is nullpotent on $\mathfrak{X}$, say with index i, so that $\underset{\sim}{a}^i\mathfrak{X}=0$, but $\mathfrak{U} = \underset{\sim}{a}^{i-1}\mathfrak{X} \neq 0$, then we have $\mathfrak{U} \subset \mathfrak{X}$, and $\underset{\sim}{a}\mathfrak{U} = 0$, and any vector of $\mathfrak{U}$ is a latent vector of $\underset{\sim}{a}$, as required.

3. Latent vectors and characteristic sets.

Let $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... denote any finite number of commutative matrices. We say α, β, ... are latently associated characteristic values of $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... respectively, if there exists a vector $\underset{\sim}{X}$ which is a latent vector of $\underset{\sim}{a}$ corresponding to α, and $\underset{\sim}{b}$ corresponding to β, and....., that is to say $\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{X}\alpha$, $\underset{\sim}{b}\underset{\sim}{X}=\underset{\sim}{X}\beta$, ... The sets of latently associated characteristic values of the matrices in a set define the characteristic sets of that matrix set. We shall now prove two fundamental results concerning the existence and extension of the characteristic sets of finite commutative matrix sets, which will subsequently extended to infinite sets of matrices.

THEOREM I. (i) Every commutative matrix set has at least one characteristic set.

(ii) The characteristic sets of any commutative matrix set can be extended to give the characteristic sets of any commutative extension of that set.

First we shall prove that if $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... is any set of commutative matrices, and α, β, ... one of its characteristic sets, if there are any, and if $\underset{\sim}{f}$ is any further matrix which commutes with each of $\underset{\sim}{a}$, $\underset{\sim}{b}$, ..., then there exists a characteristic value ϕ of $\underset{\sim}{f}$ such that α, β, . .ϕ is a characteristic set of $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... It will then follow that any n+1 commutative matrices will have at least one characteristic set if this is the case for any n.

Part (i) of the theorem for infinite sets of matrices follows directly from the case for finite sets. For every infinite set has a finite generating subset, and every latent vector of the generating subset is a latent vector of the set generated. The extension of (ii) is similarly obtained. Thus let $\underset{\sim}{t}$ denote an extension of a matrix set $\underset{\sim}{s}$, and $\underset{\sim}{m}$ a generating subset of $\underset{\sim}{t}$ which extends a generating subset $\underset{\sim}{l}$ of $\underset{\sim}{s}$. Then by (ii) for finite sets the characteristic sets μ of $\underset{\sim}{m}$ are obtained by taking extensions of each of the characteristic sets λ of $\underset{\sim}{l}$; and since the latent vectors of any generating subset of a matrix set contain all the latent vectors of the complete set, it follows that the characteristic sets of $\underset{\sim}{t}$ are obtained by taking extensions of the characteristic sets of $\underset{\sim}{s}$.

Then (1) will be proved, since any single matrix has at least one characteristic value. Moreover (ii) will follow directly, again by an inductive application of this proposition. Thus, if α, β, ... is a characteristic set of $\underset{\sim}{a}$, $\underset{\sim}{b}$, ..., let $\mathfrak{X}$ be the space of latent vectors which correspond, so that $\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{X}\alpha$, $\underset{\sim}{b}\underset{\sim}{X}=\underset{\sim}{X}\beta$, ... for each vector $\underset{\sim}{X}$ of $\mathfrak{X}$. Then, since $\underset{\sim}{f}$ commutes with each of $\underset{\sim}{a}$, $\underset{\sim}{b}$, ..., we have $\underset{\sim}{a}\underset{\sim}{f}\underset{\sim}{X}=\underset{\sim}{f}\underset{\sim}{a}\underset{\sim}{X}=\underset{\sim}{f}\underset{\sim}{X}\alpha$, $\underset{\sim}{b}\underset{\sim}{f}\underset{\sim}{X}=\underset{\sim}{f}\underset{\sim}{b}\underset{\sim}{X}=\underset{\sim}{f}\underset{\sim}{X}\beta$, ... for each vector $\underset{\sim}{X}$ of $\mathfrak{X}$; that is $\underset{\sim}{f}\underset{\sim}{X}$ belongs to $\mathfrak{X}$ for each vector $\underset{\sim}{X}$ of $\mathfrak{X}$, or $\mathfrak{X}$ is invariant in $\underset{\sim}{f}$. It follows by theorem 2.11 that there exists a vector $\underset{\sim}{Y}$ in $\mathfrak{X}$ which is a latent vector of $\underset{\sim}{f}$, say with characteristic value ϕ. Then $\underset{\sim}{Y}$ is a latent vector of $\underset{\sim}{a}$, $\underset{\sim}{b}$, .., $\underset{\sim}{f}$, and α, β, ...ϕ is the corresponding characteristic set. This proves the theorem for finite sets.

The following theorem is easily proved:

THEOREM II. *If $\underset{\sim}{s}U\underset{\sim}{t}$ denotes a commutative extension of a commutative matrix set $\underset{\sim}{s}$, and $\underset{\sigma}{t}$ is the set of values relative to some base of the matrices of $\underset{\sim}{t}$ on the axial space of $\underset{\sim}{s}$ corresponding to a characteristic set σ, then the characteristic sets of $\underset{\sim}{s}U\underset{\sim}{t}$ are the sets $\sigma U\tau_\sigma$ where σ is any characteristic set of $\underset{\sim}{s}$, and then τ_σ is any characteristic set of $\underset{\sigma}{t}$.*

4. Rational functions.

If any matrices $\underline{a},\underline{b}$ have a common latent vector $\underline{X}$, and α,β are their corresponding characteristic values, so that $\underline{a}\underline{X}=\underline{X}\alpha$, $\underline{b}\underline{X}=\underline{X}\beta$, then for any scalars λ,μ we have

$$\underline{a}^{-1}\underline{X}=\underline{X}\alpha^{-1}, \quad \underline{a}\underline{b}\underline{X}=\underline{X}\alpha\beta, \quad (\lambda\underline{a}+\mu\underline{b})\underline{X}=\underline{X}(\lambda\alpha+\mu\beta).$$

Hence $\underline{X}$ is a latent vector of $\underline{a}^{-1}$, $\underline{a}\underline{b}$, and $(\lambda\underline{a}+\mu\underline{b})$, and the characteristic values are α^{-1}, $\alpha\beta$, and $(\lambda\alpha+\mu\beta)$. It follows by induction that if $f(\underline{s})$ is a matrix constructed by finite rational operations in a finite set $\underline{s}$ of matrices which have a vector $\underline{X}$ in common as an axial vector, and if σ is the set of the corresponding characteristic values, then $f(\underline{s})\underline{X}=\underline{X}f(\sigma)$; so then $\underline{X}$ is a latent vector of $f(\underline{s})$, with $f(\sigma)$ as characteristic value. Accordingly if $f(\omega)$ is any rational function in a set ω of scalar indeterminates, and if $\underline{s}$ is a similar set of commutative matrices such that no characteristic set σ of $\underline{s}$ gives a singularity of the function $f(\omega)$, or $f(\sigma)$ is always defined, its denominator being non-zero, then the members of $\underline{s}$ can be substituted for ω, and the matrix $f(\underline{s})$ obtained is defined is, since $\underline{s}$ is commutative and since the denominator in its expression is non-singular.

Now if σ is any characteristic set of $\underline{s}$, let $\underline{X}$ be a corresponding latent vector. Then $f(\underline{s})\underline{X}=\underline{X}f(\sigma)$, showing $f(\sigma)$ to be a characteristic value of $f(\underline{s})$. Moreover the values of $f(\sigma)$ where σ ranges over the characteristic sets of $\underline{s}$ are all the characteristic values of $f(\underline{s})$.

For the matrix set $(\underline{f},\underline{s})$ formed by augmenting $\underline{s}$ by a further member $\underline{f}=f(\underline{s})$, is commutative; and hence, by the extension theorem for characteristic sets 4,II, there exists , for each characteristic value ϕ of $\underline{f}$ a characteristic set σ of $\underline{s}$ such that (ϕ,σ) is a characteristic set of $(\underline{f},\underline{s})$. Let $\underline{X}$ be any corresponding latent vector. Then $\underline{X}\phi=\underline{f}\underline{X}=f(\underline{s})\underline{X}=\underline{X}f(\sigma)$; and hence $\phi=f(\sigma)$. We have proved the following theorem:*[5]

THEOREM. The characteristic values of any rational function of a commutative matrix set are the values of that function when the argument ranges over the characteristic sets of that matrix set.

We note now that the axial vectors of $\underline{f}$ include all the axial vectors of $\underline{s}$, but that not all the axial vectors of $\underline{f}$ are necessarily axial vectors of $\underline{s}$; however in any axial space of $\underline{f}$ there are axial vectors of $\underline{s}$. Thus we have that all the axial spaces of $\underline{s}$ are subspaces of the axial spaces of $\underline{f}$, and each axial space of $\underline{f}$ contains at least one axial space of $\underline{s}$.

If $\underline{f}$ is any matris rationally dependent on a commutative matrix set $\underline{s}$, then values ϕ_σ are the characteristic values of $\underline{f}$ indexed relative to $\underline{s}$ if they may be given as values $\phi_\sigma=f(\sigma)$, where $f(\omega)$ is a rational function such that $f=f(\underline{s})$, and σ ranges over the characteristic sets

of $\underline{s}$. Further, we may say $\underline{f}$ is <u>resolved</u> relative to $\underline{s}$ if $\phi_\lambda \neq \phi_\mu$ $(\lambda \neq \mu)$.

5. Triangular reduction of commutative sets.

If $\underline{X}$ is a unit vector, let $\underline{U}_1=\underline{X}$, and let $\underline{U}_2,\ldots,\underline{U}_n$ be a set of mutually orthogonal unit vectors which are orthogonal to $\underline{X}$. Then $\underline{U}=(\underline{U}_1\ldots\underline{U}_n)$ is a unitary matrix, which may be called a unitary matrix with initial vector $\underline{X}$; it is such that $\underline{X}=\underline{U}\underline{e}$, where $\underline{e}=\begin{pmatrix}\underline{I}\\ \underline{0}\end{pmatrix}$ denotes the first fundamental unit vector. Now let α be any characteristic value of a matrix $\underline{a}$, and $\underline{X}$ a corresponding unit latent vector, so that $\underline{a}\underline{X}=\underline{X}\alpha$; and let $\underline{U}$ be a unitary matrix with initial vector $\underline{X}$. Then $\underline{X}=\underline{U}\underline{e}$, so that $(\underline{U}^{-1}\underline{a}\underline{U})\underline{e}=\underline{e}\alpha$; and this implies that the initial column of $\underline{U}^{-1}\underline{a}\underline{U}$ is $\underline{e}\alpha$, or that $\underline{U}^{-1}\underline{a}\underline{U}=\begin{pmatrix}\alpha & \underline{\xi}\\ \underline{0} & \underline{a}'\end{pmatrix}$. Now the following theorem can be proved:* [6]

THEOREM I. The matrices of a commutative matrix set can all be reduced to triangular form by the same unitary transformation.

This theorem is vacuously true for sets of matrices of order one. It will be proved inductively by showing it for sets of matrices of some arbitrary order, for its hypothesis for sets in which the order is less by one. Thus let $\underline{a},\underline{b},\ldots$ be a set of matrices which all commute, so that, by theorem 4,I, they have some axial vector $\underline{X}$ in common, say for characteristic values $\alpha,\beta,\ldots$ respectively Let $\underline{U}_0$ be a unitary matrix with initial vector $\underline{X}$. Then

$$\underline{U}_0^{-1}\underline{a}\underline{U}_0=\begin{pmatrix}\alpha & \underline{\xi}\\ \underline{0} & \underline{a}_1\end{pmatrix},\quad \underline{U}_0^{-1}\underline{b}\underline{U}_0=\begin{pmatrix}\beta & \underline{\eta}\\ \underline{0} & \underline{b}_1\end{pmatrix}\ldots\ldots$$

and since $\underline{a}$, $\underline{b}$, ... commute, so do $\underline{a}_1$, $\underline{b}_1$, ..., and their orders are less by one. Hence, by the inductive hypothesis, there arexists a unitary matrix $\underline{U}_1$ for which $\underline{U}_1^{-1}\underline{a}_1\underline{U}_1$, $\underline{U}_1^{-1}\underline{b}_1\underline{U}_1$, ... are all triangular. Let $\underline{U}=\underline{U}_0\begin{pmatrix}\underline{I} & \underline{0}\\ \underline{0} & \underline{U}_1\end{pmatrix}$. Then $\underline{U}$ is a unitary matrix such that $\underline{U}^{-1}\underline{a}\underline{U}$, $\underline{U}^{-1}\underline{b}\underline{U}$, ... are all triangular.

The <u>diagonal sets</u> of a matrix set are the sets of scalars formed by the similarly situated diagonal elements of the matrices; and they are certain sets of characteristic values if the matrices are triangular. The <u>multiplicity</u> m_τ of any scalar set τ as a diagonal set of a matrix set t is defined as the multiplicity of the repetition with which τ occurs among the diagonal sets of t.

THEOREM II. <u>For a commutative matrix set s, the diagonal sets of any set $t=\underline{U}^{-1}s\underline{U}$, of the transforms of the members of s by a matrix $\underline{U}$, are the same for all the unitary matrices $\underline{U}$ which make all the members of t triangular, and are the characteristic sets of s with certain fixed multiplicities.</u>

Let f range in the linear closure of s while its characteristic values ϕ_σ indexed relative to s remain all separate. Then $\underline{U}^{-1}f\underline{U}$ is triangular, and the ϕ_σ are its characteristic values. Since f and $\underline{U}^{-1}f\underline{U}$ have the same characteristic equation , each ϕ_σ is a ϕ_τ and each ϕ_τ a ϕ_σ

Hence each τ is a σ, and each σ a τ. Thus, since the ϕ_σ are separate, we have that the ϕ_τ are separate, and are the characteristic values of multiplicity m_τ respectively of both $\underset{\sim}{U}^{-1}\underset{\sim}{f}\underset{\sim}{U}$ and $\underset{\sim}{f}$, irrespective of the particular $\underset{\sim}{U}$ which makes $\underset{\sim}{U}^{-1}\underset{\sim}{s}\underset{\sim}{U}$ triangular.

6. Rank, multiplicity and index.

THEOREM I. *There are multiplicities m_σ attached to the characteristic sets σ of any commutative matrix set $\underline{s}$, such that if ϕ has multiplicity m_ϕ as a characteristic value of $\underline{f}$, then $m_\phi = \sum_{\phi_\sigma = \phi} m_\sigma$ for every matrix $\underline{f}=f(\underline{s})$ which is a rational combination of $\underline{s}$, and its characteristic values $\phi_\sigma = f(\sigma)$ indexed relative to $\underline{s}$; so that*

$$\prod_\sigma (\underline{f} - \phi_\sigma \underline{1})^{m_\sigma} = \underline{0}.$$

By theorem 7, there exists a unitary matrix $\underline{U}$ by which the matrices of $\underline{s}$ transform into a set $\underline{t}=\underline{U}^{-1}\underline{s}\underline{U}$ of triangular matrices. Any rational combination $\underline{f}=f(\underline{s})$ of $\underline{s}$ has, under $\underline{U}$, the transform $\underline{U}^{-1}\underline{f}\underline{U}=f(\underline{U}^{-1}\underline{s}\underline{U})=f(\underline{t})$, which is triangular; and its characteristic values are the diagonal elements $f(\tau)$ of $\underline{U}^{-1}\underline{f}\underline{U}$, where the τ are the sets of corresponding diagonal elements in the members of $\underline{t}$. Now, by theorem 6, the values $f(\sigma)$ are also the characteristic values of $\underline{f}$, for all rational combinations. So we have the sets τ consist of the characteristic sets σ with certain repetitions, say σ repeated with a multiplicity m_σ. A value ϕ appears with an m_ϕ-fold repetition on the diagonal of $\underline{U}^{-1}\underline{f}\underline{U}$, where $m_\phi = \sum_{\phi_\sigma = \phi} m_\sigma$; and this gives the multiplicity m_ϕ of ϕ as a characteristic value of $\underline{f}$. Now, by the Cayley-Hamilton theorem, $\prod_\phi (\underline{f} - \phi \underline{1})^{m_\phi} = \underline{0}$; so

$$\underline{0} = \prod_\phi (\underline{f} - \phi \underline{1})^{\sum_{\phi_\sigma = \phi} m_\sigma} = \prod_\phi \prod_{\phi_\sigma = \phi} (\underline{f} - \phi_\sigma \underline{1})^{m_\sigma} = \prod_\sigma (\underline{f} - \phi_\sigma \underline{1})^{m_\sigma}.$$

Take m_σ as the multiplicity of σ as a characteristic set of $\underline{s}$; define the rank r_σ of σ as a characteristic set of $\underline{s}$ as the rank of the corresponding axial space $\mathfrak{X}_\sigma$ of $\underline{s}$; and with the smallest exponents i_σ such that $\prod_\sigma (\underline{f}-\phi_\sigma \underline{1})^{i_\sigma} = 0$ for all rational combinations $\underline{f}$ of $\underline{s}$ with characteristic values ϕ_σ indexed relative to $\underline{s}$, let i_σ define the index of σ as a characteristic set of $\underline{s}$. Then we have $0<i_\sigma \leqslant m_\sigma$, and $0<r_\sigma \leqslant m_\sigma$, shown as follows:

THEOREM II. The index of any characteristic set of a commutative matrix set is greater than zero and at most its multiplicity.

Firstly, we have that there is no characteristic value σ, of any matrix $\underline{a}$, which has zero index. For if $\underline{aX}=\underline{X}\alpha$, then $p(\underline{a})\underline{X}=\underline{X}p(\alpha)$, so that $p(\alpha)=0$ if $p(\underline{a})=\underline{0}$; so that any characteristic value of $\underline{a}$ matrix is a zero of its minimum function, as required.*(7) Now let $\underline{f}$, rationally dependent by $\underline{s}$, with characteristic values ϕ_σ, be such that $\phi_\sigma \neq \phi_\omega$ $(\omega \neq \sigma)$; and let i_ϕ be the index of $\phi=\phi_\sigma$. Then $i_\phi \leqslant i_\sigma$; and, since we know $i_\phi > 0$, we know have $i_\sigma > 0$.

THEOREM III. The rank of any characteristic set of a commutative matrix set is greater than zero and at most its multiplicity.

Firstly, the rank r_α of any characteristic value α, of a matrix $\underline{a}$, is at most its multiplicity. For we have $t=0$ is a root of multiplicity m_α of the equation

$$|\underline{a}_\alpha - t\underline{I}| = \sum_{t=0}^{n} (-t)^{n-t} \text{ trace } \underline{a}_\alpha^{(t)} = 0,$$

where $\underline{a}_\alpha = \underline{a} - \alpha\underline{I}$. Hence trace $\underline{a}_\alpha^{(n-m_\alpha)} \neq 0$; and thus $\underline{a}_\alpha^{(n-m_\alpha)} \neq \underline{0}$. Thus the matrix $\underline{a}_\alpha$ has rank at least $n-m_\alpha$, and the rank r_α of the space $\mathfrak{X}_\alpha$, of vectors $\underline{X}$ such that $\underline{a}_\alpha \underline{X} = \underline{0}$, that is $\underline{a}\underline{X} = \underline{X}\alpha$, is at most m_α. Now let $\underline{f}$, rationally generated by a commutative matrix set $\underline{s}$, with characteristic values ϕ_σ indexed relative to $\underline{s}$, be such that $\phi_\sigma \neq \phi_\omega$ $(\omega \neq \sigma)$. Then by theorem I, $m_\rho = m_\sigma$ is the multiplicity of $\phi = \phi_\sigma$ as a characteristic value of $\underline{f}$. Now every latent vector of $\underline{s}$ is a latent vector of $\underline{f}$, so that $r_\sigma \leq r_\rho$. But we have $r_\phi \leq m_\phi$. Hence $r_\sigma \leq m_\sigma$.

7. Independence of axial spaces.

THEOREM. The axial spaces of any commutative matrix set are independent.

We first show that the axial spaces $\mathfrak{X}_\alpha$, of any matrix $\underline{a}$, with characteristic values α, are all independent. If they are not independent, there are vectors $\underline{X}_\alpha$, from the spaces $\mathfrak{X}_\alpha$ respectively, such that $X = \sum_\alpha \underline{X}_\alpha = 0$; so that we have $0 = \underline{a}^m \underline{X} = \sum_\alpha \underline{X}_\alpha \alpha^m$. This gives $\sum_\alpha \xi_\alpha \alpha^m = 0$ when the ξ_α are elements, not all zero, of the vectors $\underline{X}_\alpha$ respectively. Elimination of the ξ_α from a set $m = 0, 1, \ldots$ of these equations gives the vanishing of the simple alternant in the α's; which is impossible, since the α's are distinct. Hence $\underline{X} \neq 0$, and the hypothesis of dependence is false.

Now let $\underline{f}$ be a rational combination of $\underline{s}$ with distinct ϕ_σ. If $\mathfrak{X}_\phi$, $\mathfrak{U}_\sigma$ are the axial spaces of $\underline{f}$ for ϕ, $\underline{s}$ for σ respectively, then $\mathfrak{U}_\sigma \subset \mathfrak{X}_\phi$, where σ is the unique characteristic set of $\underline{s}$, such that ϕ_σ is the characteristic value ϕ of f. So, every axial space of $\underline{s}$ lies in some axial space of $\underline{f}$, and there is just one axial space of $\underline{s}$ contained in each axial space of $\underline{f}$. Now the independence of the $\mathfrak{U}_\sigma$ follows from the independence of the $\mathfrak{X}_\phi$.

COROLLARY. If $\underline{f}$ is rationally dependent on a commutative matrix set $\underline{s}$, and has characteristic values ϕ_σ indexed relative to $\underline{s}$, then $r_\phi \geq \sum_{\phi_\sigma = \phi} r_\sigma$, for the rank r_ϕ of any value ϕ as a characteristic value of $\underline{f}$.

8 . Idempotents.

An idempotent is a matrix $\underline{e}$ such that $\underline{e}^2=\underline{e}$; so that $\underline{e}$ acts as a unit in the algebra of matrices of the form $\underline{a}_e=\underline{a}\underline{e}$, since $\underline{e}\underline{a}_e=\underline{a}_e\underline{e}=\underline{a}$; also $\underline{e}$ defines a projection $\underline{e}:1\to 1$, of 1, onto the subspace $R_e=\underline{e}1$, of vectors $\underline{X}$ such that $\underline{e}\underline{X}=\underline{X}$, parallel to the subspace N_e, of vectors $\underline{X}$ such that $\underline{e}\underline{X}=0$. An idempotent has only zero or unit characteristic values, and its axial vectors are those vectors which it either annihilates or leaves invariant.

THEOREM I. If $\underline{e}^2=\underline{e}$, then the range of $\underline{e}$ is the null-space of $\underline{I}-\underline{e}$.

If $\underline{X}$ is in the null-space $\mathfrak{N}_e$ of $\underline{I}-\underline{e}$, then $\underline{X}=\underline{e}\underline{X}$, so that $\underline{X}$ belongs to the range $\mathfrak{R}_e$ of $\underline{e}$. Thus $\mathfrak{N}_e \subset \mathfrak{R}_e$. Moreover, if $\underline{X}$ is any vector of $\mathfrak{R}_e$, then $\underline{X}=\underline{e}\underline{Y}$, where $\underline{Y}$ is any vector; so that $\underline{e}\underline{X}=\underline{e}^2\underline{Y}=\underline{e}\underline{Y}=\underline{X}$, and $(\underline{I}-\underline{e})\underline{X}=0$, or $\underline{X}$ belongs to $\mathfrak{N}_e$. Thus $\mathfrak{R}_e \subset \mathfrak{N}_e$. Hence $\mathfrak{R}_e=\mathfrak{N}_e$.

COROLLARY i. The space of invariant vectors of an idempotent matrix is identival to the range.

This restates the theorem above.

COROLLARY ii. If $\underline{e}^2=\underline{e}$, then $\mathfrak{R}_e \cap \mathfrak{R}_{1-e}=0$ and $R_e \oplus R_{1-e}=1$.

For $\mathfrak{R}_e$ and $\mathfrak{R}_{1-e}$ are the range and null-space respectively of a matrix, explicitly $\underline{e}$; so they form a dual pair of spaces, which are independent, and which together span the unit space.

COROLLARY iii. The vectors which any idempotent either annihilates or leaves invariant together span the unit space.

Another expression of this proposition is that any idempotent is axially complete.

Idempotents f,g are said to be independent if $gf=fg=0$.

THEOREM II. If $f,g,\ldots$ are any independent idempotents, then

$$R_f \cap R_g = \theta, \ldots\ldots, \text{ and } R_f \oplus R_g \oplus \ldots = R_{f+g+\ldots}$$

We show the theorem for two idempotents, and complete the proof by induction for an arbitrary number. Thus for any two independent idempotents f,g we have $Z=(f+g)Y$ is any vector of R_{f+g}, where Y is any vector; also $Z=fY+gY$, which belongs to $R_f \oplus R_g$. Thus $R_{f+g} \subset R_f \oplus R_g$. Now $Z=fX+gY$ is any vector of $R_f \oplus R_g$, where X and Y are any vectors, and

$$(f+g)Z=(f+g)(fX+gY)=fX+gY=Z,$$

since $f^2=f$, $g^2=g$, and $gf=fg=0$; that is, Z is an invariant vector of the idempotent $f+g$, and hence, by theorem I Z belongs to the range R_{f+g}. Thus $R_f \oplus R_g \subset R_{f+g}$, and hence $R_f \oplus R_g = R_{f+g}$. Now, for the inductive step, we show the theorem for arbitrary independent idempotents $f,g,\ldots$ on the hypothesis of the theorem for these idempotents with f omitted; and the proof of the theorem will then

be complete. Thus, we have

$$R_f \oplus R_g \oplus \ldots = R_f \oplus (R_g \oplus \ldots) = R_f \oplus R_{g+\ldots} = R_{f+(g+\ldots)} = R_{f+g+\ldots}$$

by the associative law, the inductive hypothesis, the theorem for two idempotents, and the associative law respectively.

THEOREM III. *If $\sum_\lambda e_\lambda = I$ is a decomposition of the unit matrix into a sum of independent idempotents e_λ, of rank r_λ respectively, then there exists a regular matrix P such that*

$$P^{-1} e_\lambda P = \begin{pmatrix} 0 & & 0 \\ & I_{r_\lambda} & \\ 0 & & 0 \end{pmatrix}.$$

Let P_λ be a base for the space $R_{e_\lambda} = e_\lambda 1$. Then since, by theorem II, we have $\bigoplus_\lambda R_{e_\lambda} = R_{\sum_\lambda e_\lambda} = R_1 = 1$, it appears that the matrix $P = (P_1 \ldots P_\lambda \ldots P_l)$ is regular. Moreover e_λ has the value I on R_{e_λ}, and the value 0 on R_{e_μ} $(\mu \neq \lambda)$. Hence

$$e_\lambda P = (0 \ldots I_{r_\lambda} \ldots 0) = P \begin{pmatrix} 0 & & 0 \\ & I_{r_\lambda} & \\ 0 & & 0 \end{pmatrix};$$

which proves the theorem.

An idempotent e is *reduced* by an idempotent e' if there exists an idempotent e'' such that $e = e' + e''$.

THEOREM IV. *If e, e' are idempotents which commute, then e' reduces e if and only if $e'e = e'$.*

9. Resolvent polynomials of a matrix.

The roots of the minimum function $\phi_a(\zeta)$ of a matrix $\underline{a}$ are just the characteristic values of $\underline{a}$; that is,

$$\phi_a(\zeta)=\prod_{\alpha}(\zeta-\alpha)^{i_\alpha}, \quad 0<i_\alpha\leq m_\alpha,$$

where m_α is the multiplicity, and i_α defines the index of any characteristic value α. For the minimum function divides the characteristic function $C_a(\tau)=|\underline{a}-\tau\underline{I}|$, whose roots are the values α such that there exists a vector $\underline{X}$ such that $\underline{a}\underline{X}=\underline{X}\alpha$; and if $\underline{a}\underline{X}=\underline{X}\alpha$ for some vector $\underline{X}$, so α is a characteristic value of $\underline{a}$, then $\underline{0}=\phi_a(\underline{a})\underline{X}=\underline{X}\phi_a(\alpha)$, so that $\phi_a(\alpha)=0$. Further, since $\phi_a(\zeta)$ is defined as a polynomial of least degree which is null in $\underline{a}$, and divides every polynomial which is null in $\underline{a}$, we have

$$\psi(\underline{a})\prod_{\omega\neq\alpha}(\underline{a}-\omega\underline{I})^{i_\omega}=0$$

if and only if $(\zeta-\alpha)^{i_\alpha}$ is a factor of $\psi(\zeta)$.

Write $\underline{a}_\alpha=\underline{a}-\alpha\underline{I}$, and let $\underline{f}_\alpha=\prod_{\omega\neq\alpha}\underline{a}_\omega^{i_\omega}$. Then each $\underline{f}_\alpha$ is a polynomial in $\underline{a}$, which is not null, with the characteristic values $f_\alpha(\lambda)$ of multiplicity m_λ respectively, where $f_\alpha(\lambda)=0$ $(\lambda\neq\alpha)$, but $f_\alpha(\alpha)\neq 0$. Thus the matrix $\underline{f}=\sum_\alpha\underline{f}_\alpha$ is regular, having characteristic values $f(\lambda)=\sum_\alpha f_\alpha(\lambda)=f_\alpha(\alpha)$, none of which is zero. Accordingly we may define matrices $\underline{e}_\alpha=\underline{f}_\alpha/\underline{f}$; and all these are expressible as polynomials in $\underline{a}$, since $\underline{f}^{-1}$ can be expressed as a polynomial in $\underline{f}$, and $\underline{f}$ and the $\underline{f}_\alpha$ are polynomials in $\underline{a}$. Directly we have the following.

THEOREM I. $\underline{e}_\alpha^2 = \underline{e}_\alpha \neq 0$, $\underline{e}_\lambda \underline{e}_\mu = 0$ $(\lambda \neq \mu)$, $\sum_\alpha \underline{e}_\alpha = \underline{I}$, $\underline{a}^{i_\alpha} \underline{e}_\alpha = 0$, and $\phi(\underline{a})\underline{e}_\alpha \neq \underline{0}$ for every polynomial $\phi(\zeta)$ of degree less than i_α.

Thus the $\underline{e}_\alpha$ define an independent idempotent decomposition of the unit matrix, which we will call the resolution of the unit matrix relative to the matrix a. The uniqueness of the resolvents e_α of $\underline{a}$, with respect to certain of their properties, is shown in the next theorem.

THEOREM II. If any matrices $\underline{f}_\alpha$, indexed by the characteristic values α of a matrix $\underline{a}$, are such that

$\underline{a}\underline{f}_\alpha = \underline{f}_\alpha \underline{a}$, $\underline{f}_\alpha^2 = \underline{f}_\alpha \neq 0$, $\sum_\alpha \underline{f}_\alpha = \underline{I}$, and $\underline{a}_\alpha \underline{f}_\alpha$ is nullpotent, then $\underline{f}_\alpha = \underline{e}_\alpha$, where the $\underline{e}_\alpha$ are the resolvent polynomials of $\underline{a}$.

Since $\underline{a}\underline{f}_\alpha = \underline{f}_\alpha \underline{a}$, and the $\underline{e}_\alpha$ are polynomials in $\underline{a}$, we have that $\underline{e}_\mu \underline{f}_\lambda = \underline{f}_\lambda \underline{e}_\mu$; so if $\underline{g}_{\lambda\mu} = \underline{f}_\lambda \underline{e}_\mu$, then $\underline{g}_{\lambda\mu}^2 = \underline{g}_{\lambda\mu}$ and $\underline{g}_{\lambda\mu}$ is nullpotent only if it is null. Now the matrices $\underline{p}_\mu = \underline{a}_\mu \underline{e}_\mu$, $\underline{q}_\lambda = \underline{a}_\lambda \underline{f}_\lambda$ are nullpotent. Hence the matrix

$$(\lambda - \mu)\underline{g}_{\lambda\mu} = \underline{p}_\mu \underline{f}_\lambda - \underline{q}_\lambda \underline{e}_\mu$$

is nullpotent. But if $\lambda \neq \mu$ this implies that $\underline{g}_{\lambda\mu}$ is nullpotent, and hence null. That is, $\underline{f}_\lambda \underline{e}_\mu = \underline{0}$ $(\lambda \neq \mu)$; and Thus

$$\underline{e}_\alpha = \underline{e}_\alpha \sum_\lambda \underline{f}_\lambda = \underline{e}_\alpha \underline{f}_\alpha = \underline{f}_\alpha \sum_\mu \underline{e}_\mu = \underline{f}_\alpha .$$

We have, further, the following independence theorem.

THEOREM III. The matrices $\underline{a}_\alpha^{\dagger}\underline{e}_\alpha$, $(r<i_\alpha, \alpha)$, form a linearly independent set.

Suppose they are linearly dependent, and there exist scalars $\lambda_{\alpha,r}$, not all zero, such that

$$\underline{L}=\sum_{\omega}\sum_{r<i_\omega}\lambda_{\omega,r}\underline{a}_\omega^{\dagger}\underline{e}_\omega=0.$$

Then for each characteristic value α we have, by theorem I,

$$0=\underline{e}_\alpha\underline{L}=\sum_{r<i_\alpha}\lambda_{\alpha,r}\underline{a}_\alpha^{\dagger}\underline{e}_\alpha=\phi_\alpha(\underline{a})\underline{e}_\alpha,$$

where $\phi_\alpha(\underline{a})=\sum_{r<i_\alpha}\lambda_{\alpha,r}\underline{a}_\alpha^{\dagger}$. But $\phi_\alpha(\tau)$ is a polynomial of degree less than the index i_α of α, so that $\phi_\alpha(\underline{a})\underline{e}_\alpha\neq 0$ by theorem I; and this makes a contradiction. Hence the hypothesis of linear dependence is false.

We say $\alpha,\beta,\ldots$ are latently associated characteristic values of commutative matrices $\underline{a},\underline{b},\ldots$ if there exists a vector $\underline{X}$ which is a latent vector of $\underline{a}$ corresponding to α, $\underline{b}$ corresponding to $\beta,\ldots\ldots$, that is $\underline{a}\underline{X}=\underline{X}\alpha$, $\underline{b}\underline{X}=\underline{X}\beta,\ldots$, in which case the latent association is through the vector $\underline{X}$; or equivalently if $\alpha,\beta,\ldots$ form a characteristic set of $\underline{a},\underline{b},\ldots$ We may write $\alpha\backsim\beta$ for when α,β are latently associated characteristic values of $\underline{a}$ and $\underline{b}$.

THEOREM IV. If $\underline{a},\underline{b},\ldots$ are commutative matrices, with resolvents $f_\alpha, g_\beta,\ldots$ respectively, then $f_\alpha g_\beta\ldots\neq 0$ if and only if $\alpha,\beta,\ldots$ are latently associated characteristic values of $\underline{a},\underline{b},\ldots$

Since $\underline{a},\underline{b},\ldots$ commute, so do their polynomials; so that $\underline{f}_\alpha,\underline{g}_\beta,\ldots$ commute. Therefore $(\underline{f}_\alpha\underline{g}_\beta\ldots)^2=\underline{f}_\alpha\underline{g}_\beta\ldots$, and hence $\underline{f}_\alpha\underline{g}_\beta\ldots$ is not nullpotent unless it is null. But the characteristic values of $\underline{f}_\alpha\underline{g}_\beta\ldots$ are, by theorem 6, $f_\alpha(\lambda)g_\beta(\mu)\ldots=\delta_{\alpha\lambda}\delta_{\beta\mu}\ldots$, where $\lambda,\mu,\ldots$ are latently associated characteristic values of $\underline{a},\underline{b},\ldots$ Therefore we have that $\underline{f}_\alpha\underline{g}_\beta\ldots$ has only zero characteristic values, and so is nullpotent and hence null, unless $\alpha,\beta,\ldots$ are a set of latently associated characteristic values of $\underline{a},\underline{b},\ldots$; in which case $\underline{f}_\alpha\underline{g}_\beta\ldots$ has one non-zero characteristic value, and hence is non-null.

If a commutative matrix pair $\underline{a},\underline{b}$ are such that there is a unique characteristic value β of $\underline{b}$ latently associated with each characteristic value α of $\underline{a}$, then we say that the matrix $\underline{a}$ <u>resolves</u> the matrix $\underline{b}$. It follows from the last theorem that if the resolvents of one matrix of a commutative pair reduce the resolvents of the other, or, as we may say, the resolution of the one <u>refines</u> the resolution of the other, then the one matrix resolves the other.

COROLLARY. <u>If $\underline{a},\underline{b}$ is a pair of commutative matrices, with resolvents $\underline{f}_\alpha$, $\underline{g}_\beta$ respectively, then $\underline{a}$ resolves $\underline{b}$ if the idempotents $\underline{f}_\alpha$ reduce the idempotents $\underline{g}_\beta$.</u>

If the $\underline{f}_\alpha$ reduce the $\underline{g}_\beta$, then, by theorem IO, IV, $\underline{f}_\alpha \underline{g}_\beta = \underline{f}_\alpha$ or $\underline{f}_\alpha \underline{g}_\beta = 0$. If $\underline{f}_\alpha \underline{g}_\beta = \underline{f}_\alpha$, then $\underline{f}_\alpha \underline{g}_\mu = \underline{f}_\alpha \underline{g}_\beta \underline{g}_\mu = 0$ $(\mu \neq \beta)$ so, by the present theorem, there is a unique β of $\underline{b}$ latently associated with each α of $\underline{a}$, which is that a resolves $\underline{b}$.

The next theorem shows the converse, that if any commutative pair of matrices are such that one resolves the other, then the independent idempotent resolution of the unit matrix relative to the one refines that of the other.

THEOREM V. <u>If matrices $\underline{a}, \underline{b}$ commute, and for each characteristic value α of $\underline{a}$ there is a unique latently associated characteristic value β of $\underline{b}$, then $\underline{f}_\alpha \underline{g}_\beta = \underline{f}_\alpha$, so that $\underline{g}_\beta = \sum_{\alpha \frown \beta} \underline{f}_\alpha$, where $\underline{f}_\alpha, \underline{g}_\beta$ are the resolvents of $\underline{a}$ and $\underline{b}$ respectively.</u>

If $\underline{f}_\alpha \underline{g}_\beta \neq 0$, we have, by theorem IV, that α and β are latently associated characteristic values, that is $\alpha \frown \beta$; and there is a unique such β corresponding to each α. Thus the non-null of the $\underline{e}_{\alpha\beta} = \underline{f}_\alpha \underline{g}_\beta$ may be regarded as indexed by the α. They commute with $\underline{a}$. Moreover, we have

$$\underline{e}_{\alpha\beta}^2 = \underline{e}_{\alpha\beta}, \quad \sum_{\alpha\beta} \underline{e}_{\alpha\beta} = \Big(\sum_\alpha \underline{f}_\alpha\Big)\Big(\sum_\beta \underline{g}_\beta\Big) = \underline{I}, \quad (\underline{a}_\alpha \underline{e}_{\alpha\beta})^{i_\alpha} = \underline{a}_\alpha^{i_\alpha} \underline{f}_\alpha \underline{g}_\beta = 0;$$

which shows, by theorem II, that $\underline{e}_{\alpha\beta} = \underline{f}_\alpha$. Thus either $\underline{f}_\alpha \underline{g}_\beta = \underline{f}_\alpha$ or $\underline{f}_\alpha \underline{g}_\beta = 0$. Hence

$$\underline{g}_\beta = \underline{g}_\beta \sum_\alpha \underline{f}_\alpha = \sum_\alpha \underline{f}_\alpha \underline{g}_\beta = \sum_{f_\alpha g_\beta \neq 0} \underline{f}_\alpha = \sum_{\alpha \frown \beta} \underline{f}_\alpha .$$

For commutative matrices $\underline{a},\underline{b},\ldots$, with resolvents $\underline{f}_\alpha, \underline{g}_\beta, \ldots$ respectively, and forming a matrix set $\underline{s}$, let $\alpha,\beta,\ldots$ be any of their latently associated characteristic values, forming a characteristic set σ of $\underline{s}$, and define $e_\sigma = \underline{f}_\alpha \underline{g}_\beta \ldots$ We shall call the $\underline{e}_\sigma$ the <u>resolvents</u> of the commutative matrix set $\underline{s}$. They are indexed by the characteristic sets σ of $\underline{s}$, and have properties shown as follows.

THEOREM VI. <u>If matrices $\underline{e}_\sigma$ are the resolvents of any commutative matrix set $\underline{s}$, with characteristic sets σ, then</u>

$$\underline{e}_\sigma^2 - \underline{e}_\sigma \neq 0, \quad \underline{e}_\sigma \underline{e}_\omega - 0 \ (\omega \neq \sigma), \quad \sum_\sigma \underline{e}_\sigma = I,$$

<u>and $(\underline{z} - \zeta_\sigma \underline{1})\underline{e}_\sigma$ is nullpotent for every matrix $\underline{z}$ which is rationally dependent on the members of $\underline{s}$, and its characteristic values ζ_σ indexed relative to $\underline{s}$.</u>

Since $\underline{s}$ is a commutative matrix set, and the $\underline{f}_\alpha, \underline{g}_\beta, \ldots$ are polynomials in its members, the $\underline{f}_\alpha, \underline{g}_\beta, \ldots$ all commute, so that, if σ is a characteristic set with elements $\alpha,\beta,\ldots$ we have

$$\underline{e}_\sigma^2 = (\underline{f}_\alpha \underline{g}_\beta \ldots)^2 = \underline{f}_\alpha^2 \underline{g}_\beta^2 \ldots = \underline{f}_\alpha \underline{g}_\beta \ldots = \underline{e}_\sigma,$$

since $\underline{f}_\alpha^2 = \underline{f}_\alpha$, $\underline{g}_\beta^2 = \underline{g}_\beta, \ldots$; also $\underline{e}_\sigma \neq 0$, by theorem IV. Again, if ω is a characteristic set with elements $\lambda,\mu,\ldots$, then

$$\underline{e}_\sigma \underline{e}_\omega = \underline{f}_\alpha \underline{f}_\lambda \underline{g}_\beta \underline{g}_\mu \ldots = \delta_{\alpha\lambda} \underline{f}_\alpha \delta_{\beta\mu} \underline{g}_\beta \ldots = \delta_{\sigma\omega} \underline{e}_\sigma,$$

so that $\underline{e}_\sigma \underline{e}_\omega = 0$ if $\omega \neq \sigma$. By theorem IV, further,

$$\sum_\sigma \underline{e}_\sigma = \sum_{f_\alpha g_\beta \neq 0} \underline{f}_\alpha \underline{g}_\beta \ldots = \sum_{\alpha\beta\ldots} \underline{f}_\alpha \underline{g}_\beta \ldots = \left(\sum_\alpha \underline{f}_\alpha\right)\left(\sum_\beta \underline{g}_\beta\right)\ldots = \underline{I},$$

since $\sum_\alpha \underline{f}_\alpha = \underline{I}$, $\sum_\beta \underline{g}_\beta = \underline{I}, \ldots$ Finally, if $\underline{z}$ is rationally

dependent on the members $\underline{a},\underline{b},\ldots$ of $\underline{s}$, it can be expressed as a polynomial $\underline{z}=p(\underline{a},\underline{b},\ldots)$; and its characteristic values indexed relative to $\underline{s}$ then are $\zeta_\sigma=p(\alpha,\beta,\ldots)$, where $\alpha,\beta,\ldots$ form the characteristic set σ. By the remainder theorem for polynomials,

$$\begin{aligned}\underline{z}-\zeta_\sigma\underline{1}&=p(\underline{a},\underline{b},\ldots)-p(\alpha,\beta,\ldots)\underline{1}\\&\quad-p(\underline{a},\underline{b},\ldots)-p(\alpha,\underline{b},\ldots)+p(\alpha,\underline{b},\ldots)-p(\alpha,\beta,\ldots)+\ldots\\&=\underline{a}_\alpha\underline{L}_\alpha+\underline{b}_\beta\underline{M}_\beta+\ldots\ldots,\end{aligned}$$

where $\underline{L}_\alpha,\underline{M}_\beta,\ldots$ are polynomials in $\underline{a},\underline{b},\ldots$ Thus we have

$$(\underline{z}-\zeta_\sigma\underline{1})\underline{e}_\sigma=\underline{L}_\alpha\underline{a}_\alpha\underline{f}_\alpha\underline{g}_\beta\ldots+\underline{M}_\beta\underline{f}_\alpha\underline{b}_\beta\underline{g}_\beta\ldots+\ldots\ldots;$$

and each term in this sum is nullpotent, since $\underline{a}_\alpha\underline{f}_\alpha,\underline{b}_\beta\underline{g}_\beta,\ldots$ are nullpotent; so that $(\underline{z}-\zeta_\sigma\underline{1})\underline{e}_\sigma$ is nullpotent.

10. Resolution of an algebra.

Let K be an algebra of commutative matrices, with an algebraically closed scalar field; so K is closed with respect to rational operations, and every matrix of K has its full complement of characteristic values. An axial space of K is a maximal space of axial vectors of all the matrices of K. Let $\mathfrak{X}_\pi$ be the axial spaces of K, distinguished by some index π; they are the axial spaces of any generating subset of K. By the linearity of $\mathfrak{X}_\pi$, there is a unique characteristic value α_π of any matrix $\underline{a}$ of K corresponding to all its axial vectors forming $\mathfrak{X}_\pi$, that is such that $\underline{a}X=X\alpha_\pi$ for every matrix X of $\mathfrak{X}_\pi$. We shall call α_π the characteristic value of $\underline{a}$ in the characteristic plane π of K. By the extension theorem 4,II, each characteristic value of any matrix of K lies in some characteristic plane of K. The characteristic sets of any set of matrices in K are determined by their characteristic values in the various characteristic planes, and the various characteristic planes of K may be also defined as the various latently associated sets of characteristic values of the matrices of K. Let $\underline{s}$ be any generating subset of K, with characteristic sets σ. Then any matrix $\underline{a}$ of K is a rational combination of the matrices of $\underline{s}$, and by theorem 6, has characteristic values α_σ which are the same combination of the elements of each σ. The characteristic values of the matrices of K which thus derive f

Such a matrix $\underset{\sim}{k}$ exists. For let $\underset{\sim}{s}$ be a generating subset of matrices in K, so that any matrix in K is of the form $p(\underset{\sim}{s})$ where $p(\omega)$ is a polynomial, and the characteristic values are $p(\sigma)$ where σ ranges over the characteristic sets of $\underset{\sim}{s}$ and is an index for the characteristic planes of K. Hence take any polynomial $f(\omega)$ such that the values of $\phi_\sigma = f(\sigma)$ are all distinct. Then $\underset{\sim}{k} = f(\underset{\sim}{s})$ has distinct characteristic values in each characteristic plane of K, as required.

from any σ form a characteristic plane of K. The characteristic sets σ of any generating subset $\underline{s}$ of K are in a one to one correspondence $\sigma \leftrightarrow \pi$ with the characteristic planes n of K, expressed by $\sigma=\sigma_\pi$, and inversely $\pi=\pi_\sigma$. We may define the multiplicity m_π and the index i_π of any characteristic plane π of K as the multiplicity m_σ and the index i_σ of the corresponding characteristic set $\sigma=\sigma_\pi$ of any generating subset $\underline{s}$ of K. The rank r_π of π is defined as the rank of the axial space $\mathfrak{X}_\pi$, or identically, as the rank r_σ of the axial space $\mathfrak{X}_\sigma$ of $\underline{s}$, where $\sigma=\sigma_\pi$. By the theorems of 8:

THEOREM I. The multiplicity m_π and the index i_π for the characteristic planes π of any commutative matrix algebra K are such that $m_\zeta=\sum_{\zeta_\pi=\zeta} m_\pi$ is the multiplicity of ζ as a characteristic value of $\underline{z}$, and the i_π are the least exponents such that $\prod_\pi (\underline{z}-\zeta_\pi \underline{1})^{i_\pi}=\underline{0}$, for every matrix $\underline{z}$ of K and its characteristic values ζ_π in π; also $0<i_\pi \leq m_\pi$ and $0<r_\pi \leq m_\pi$

Let $\underline{k}$ be any matrix in K which has one distinct characteristic value κ_π in each characteristic plane π of K; so there is a unique characteristic plane π_κ of K which contains any characteristic value κ of k, and there is a one to one correspondence between the characteristic planes π of K and the characteristic value κ of k, expressed by $\kappa=\kappa_\pi$, with inverse $\pi=\pi_\kappa$. Any such matrix $\underline{k}$ will

be called a key to the commutative matrix algebra K. It resolves every matrix of K, so that its resolvents $\underline{e}_\kappa$ reduce the resolvents $\underline{e}_\zeta$ of any matrix $\underline{z}$ of K; and $\underline{e}_\zeta = \sum_{\kappa-\zeta} \underline{e}_\kappa$ Define $\underline{e}_\pi = \underline{e}_\kappa$, where π and κ are united planes and values. Then $\underline{e}_\zeta = \sum_{\zeta_\pi=\zeta} \underline{e}_\pi$. We shall call the e_π the resolvents of the algebra K; and we establish their properties as follows:

THEOREM II. In any commutative matrix algebra K there are quantities e_π, indexed relative to the characteristic planes π of K, such that

$$\underline{e}_\pi^2 - \underline{e}_\pi \neq 0, \underline{e}_\pi \underline{e}_\rho = 0 \ (\rho \neq \pi), \sum_\pi \underline{e}_\pi = \underline{I}, \ (\underline{x}-\underline{\zeta}_\pi \underline{1})^r (\underline{y}-\underline{\eta}_\pi \underline{1})^s \ldots \underline{e}_\pi = 0, r+s+\ldots \geq i_\pi,$$

for any matrices $\underline{x},\underline{y},\ldots$ in K, and their characteristic values ζ_π, $\eta_\pi, \ldots$ in π.

The e_π have been defined as the resolvents of a key matrix $\underline{k}$, so that all but the last part of this theorem may be deduced from theorem II,I. For the last part, we note that $i_\zeta \leq i_\pi$, for the index i_ζ of a characteristic value $\zeta = \zeta_\pi$ of any matrix $\underline{z}$ of K. Moreover, since the key matrix $\underline{k}$ resolves $\underline{z}$, so that its resolvents, which are the $\underline{e}_\pi$, reduce the resolvents $\underline{e}_\zeta$ of z, we have that $\underline{e}_\zeta \underline{e}_\pi = \underline{e}_\zeta$ for each π which contains ζ. But, by theorem II,I $(\underline{z}-\zeta \underline{1})^r \underline{e}_\zeta = \underline{0}$, $r \geq i_\zeta$. Therefore, for $\zeta = \zeta_\pi$, we have

$$(\underline{z}-\zeta_\pi \underline{1})^{i_\pi} \underline{e}_\pi = (\underline{z}-\zeta \underline{1})^{i_\pi} \underline{e}_\zeta \underline{e}_\pi = \underline{0}, \text{ since } i_\pi \geq i_\zeta.$$

Now if $\underline{x},\underline{y},\ldots$ belong to K, so does $\underline{z}=\lambda \underline{x}+\mu \underline{y}+\ldots$, for all

scalars $\lambda,\mu,\ldots$, and the characteristic value of $\underline{z}$ in π is $\zeta_\pi=\lambda\xi_\pi+\mu\eta_\pi-\ldots$, where $\xi_\pi,\eta_\pi,\ldots$ are the characteristic values of $\underline{x},\underline{y},\ldots$ respectively in π. Hence we have

$$0=(\underline{z}-\zeta_\pi\underline{1})^{i_\pi}\underline{e}_\pi=(\lambda(\underline{x}-\xi_\pi\underline{1})+\mu(\underline{y}-\eta_\pi\underline{1})+\ldots)^{i_\pi}\underline{e}_\pi\ ,$$

identically in $\lambda,\mu,\ldots$ By expanding in powers of $\lambda,\mu,\ldots$ it follows that

$$\sum_{\lambda+\mu+\ldots=i_\pi}\lambda^r\mu^s\ldots(\underline{x}-\xi_\pi\underline{1})^r(\underline{y}-\eta_\pi\underline{1})^s\ldots\underline{e}_\pi\equiv_{\lambda,\mu,\ldots}0,$$

and hence that $(\underline{x}-\xi_\pi\underline{1})^r(\underline{y}-\eta_\pi\underline{1})^s\ldots e_\pi=0$, $r+s+\ldots\geq i_\pi$.*(9)

Let K_π be the subalgebra of K of matrices of the form $\underline{a}\underline{e}_\pi$, where $\underline{a}$ ranges in K. We call the algebra K the <u>section</u> of the algebra K in its characteristic plane π. If $\underline{a}_\pi$ is any matrix of K_π, we have $\underline{e}_\pi\underline{a}_\pi=\underline{a}_\pi\underline{e}_\pi=a_\pi$, since $\underline{e}_\pi^2=\underline{e}_\pi\neq0$; so that e_π is the unit in the algebra sedtion K_π. Since $\sum_\pi\underline{e}_\pi=I$, the matrices $\underline{a}$ of K are of the form $a=\sum_\pi a_\pi$, where the matrices $\underline{a}_\pi$ are of the algebras K_π respectively, showing the algebra K to be the direct sum of its section algebras K_π, or $K=\bigoplus_\pi K_\pi$. Moreover, since $\underline{e}_\pi\underline{e}_\rho=0$ $(\rho\neq\pi)$, we have $\underline{a}_\pi\underline{a}_\rho=0$, where a_π,a_ρ are matrices in distinct section algebras K_π, K_ρ. Thus we have $K_\pi K_\rho=0$ $(\pi\neq\rho)$; so that the enveloping algebra of the K_π is the same as their direct sum; defining the <u>resolution of K by its characteristic planes</u>.

THEOREM III. <u>Any commutative matrix algebra K, with characteristic planes π, has subalgebras K_π, with the</u>

resolvents e_π as respective units, such that

$$K_\pi . K_\rho = 0 \ (\pi \neq \rho), \quad K = \bigoplus_\pi K_\pi \ .$$

By §11, theorems V and VI, and §12, we have:

THEOREM IV. The resolvents of any commutative matrix algebra are the same as the rsolvents ofany generating subset.

11. Reducing transformation.

THEOREM. If K is a commutative matrix algebra, with an algebraically closed scalar field, then there exists a regular matrix $\underset{\sim}{P}$ which transforms any matrix $\underset{\sim}{z}$ of K into a direct sum $\underset{\sim}{P}^{-1}\underset{\sim}{z}\underset{\sim}{P}=\underset{\pi}{\Theta}\underset{\sim}{z}_{\pi}$ of matrices $\underset{\sim}{z}_{\pi}$, of order m_{π} respectively, where $\underset{\sim}{z}_{\pi}$ has just one distinct characteristic value, which is the characteristic value ζ_{π} of $\underset{\sim}{z}$ in the characteristic plane π of K.

Such a transformation is called a reducing transformation of the algebra.

By theorem I2,II, there are quantities $\underset{\sim}{e}_{\pi}$ in K such that $\underset{\sim}{e}_{\lambda}\underset{\sim}{e}_{\mu}=\underset{\sim}{0}$ $(\lambda\neq\mu)$, and $\underset{\pi}{\Sigma}\underset{\sim}{e}_{\pi}=\underset{\sim}{I}$. Then by theorem IO,III, there exists a regular matrix $\underset{\sim}{P}$ such that $\underset{\sim}{P}^{-1}\underset{\sim}{e}_{\pi}\underset{\sim}{P}=\underset{\lambda}{\Theta}\underset{\sim}{I}_{m_{\lambda}}\delta_{\lambda\pi}$. Now since $\underset{\sim}{e}_{\pi}$ commutes with each $\underset{\sim}{z}$ of K, and is such that $\underset{\sim}{e}_{\pi}^{2}=\underset{\sim}{e}_{\pi}$, we have that $\underset{\sim}{P}^{-1}\underset{\sim}{z}\underset{\sim}{e}_{\pi}\underset{\sim}{P}$ commutes with and has the effect of unity on $\underset{\sim}{P}^{-1}\underset{\sim}{z}\underset{\sim}{e}_{\pi}\underset{\sim}{P}$. It follows directly that $\underset{\sim}{P}^{-1}\underset{\sim}{z}\underset{\sim}{e}_{\pi}\underset{\sim}{P}=\underset{\lambda}{\Theta}\underset{\sim}{z}_{\lambda}\delta_{\lambda\pi}$; and then that $\underset{\sim}{P}^{-1}\underset{\sim}{z}\underset{\sim}{P}=\underset{\pi}{\Theta}\underset{\sim}{z}_{\pi}$, since $\underset{\pi}{\Sigma}\underset{\sim}{e}_{\pi}=\underset{\sim}{I}$.

COROLLARY. If $\underset{\sim}{a},\underset{\sim}{b},\ldots$ are any commutative matrices, whose latently associated characteristic values are $\alpha_i,\beta_i,\ldots,$ then there exists a transformation $\underset{\sim}{P}$ such that

$$\underset{\sim}{P}^{-1}\underset{\sim}{a}\underset{\sim}{P}=\begin{pmatrix}\underset{\sim}{a}_i & \\ & \underset{\sim}{0}\end{pmatrix},\quad \underset{\sim}{P}^{-1}\underset{\sim}{b}\underset{\sim}{P}=\begin{pmatrix}\underset{\sim}{b}_i & \\ & \underset{\sim}{0}\end{pmatrix},\quad \ldots\ldots$$

where $\underset{\sim}{a}_i,\underset{\sim}{b}_i,\ldots$ each have just one distinct characteristic value, $\alpha_i,\beta_i,\ldots$ respectively *(9)

12. Nullpotents.

A matrix is called a nullpotent if it is not null, but one of its poweres, and hence all further higher powers, is null; and the exponent of the lowest of its powers which is null defines its index. Thus if $\underline{x}^{i-1} \neq \underline{0}$, but $\underline{x}^{i} = \underline{0}$, then i is the index of the nullpotent $\underline{x}$. Any matrix of the form $\underline{N} = \begin{pmatrix} \underline{0} & \underline{I} \\ \underline{0} & \underline{0} \end{pmatrix}$, of any order, is called the primitive nullpotent of that order; it is a nullpotent whose index is equal to its order, that is to say which is non-derogatory. The primitive nullpotent of order r is $\underline{N}_r = \begin{pmatrix} 0 & I_{r-1} \\ \underline{0} & 0 \end{pmatrix}$, and is such that $\underline{N}_r^{s} = \begin{pmatrix} \underline{0} & \underline{I}_{r-s} \\ \underline{0} & \underline{0} \end{pmatrix}$; so that $\underline{N}_r^{r-1} \neq \underline{0}$, but $\underline{N}_r^{r} = \underline{0}$.

THEOREM. Every nullpotent is similar to a direct sum of primitive nullpotents.

Let $\underline{x}$ be a nullpotent of index r+1, so that $\underline{x}^{r+1} = \underline{0}$, but $\underline{x}^{r}$ is of rank a>0. Let $\underline{P}_{\alpha r}$ ($\alpha \leq a$) be a independent vectors from the columns of $\underline{x}^{r}$, and let $\underline{P}_{\alpha\rho}$ be the similarly situated vectors of $\underline{x}^{\rho}$ ($\rho < r$). Then $\underline{x}\underline{P}_{\alpha\rho} = \underline{P}_{\alpha,\rho+1}$ ($\rho < r$), and $\underline{x}\underline{P}_{\alpha r} = \underline{0}$. The ar vectors $\underline{P}_{\alpha\rho}$ are independent. For a relation $\underline{L} = \sum_{\rho \geq t} \sum_{\alpha} \underline{P}_{\alpha\rho} \lambda_{\alpha t} = \underline{0}$, $\lambda_{\alpha t} \neq_{\alpha} 0$ would imply $0 = \underline{x}^{r-t}\underline{L} = \sum_{\alpha} \underline{P}_{\alpha r} \lambda_{\alpha t}$, contradicting the independence of the $\underline{P}_{\alpha r}$. Let $\underline{P}_{\alpha} = (\underline{P}_{\alpha r} \ldots \underline{P}_{\alpha 1})$, and $\underline{P} = (\underline{P}_{1} \ldots \underline{P}_{a})$. Then $\underline{x}\underline{P}_{\alpha} = (\underline{0}\underline{P}_{\alpha r} \ldots \underline{P}_{\alpha 2}) = \underline{P}_{\alpha}\underline{N}_r$; and thus $\underline{x}\underline{P} = \underline{P}\begin{pmatrix} \underline{N}_r & & \\ & \ddots & \\ & & \underline{N}_r \end{pmatrix}$, with a classical nullpotents of order r along the diagonal.

Now $\underline{x}^{t-1}\,\underline{P}_{\alpha 1}=\underline{P}_{\alpha t}$; so that the 2a independent vectors $\underline{P}_{\alpha,t-1}$ and $\underline{P}_{\alpha t}$ lie in the range of $\underline{x}^{t-2}$; and thus $\underline{x}^{t-1}$ is of rank at least 2a, say 2a+b. So take b independent vectors $\underline{Q}_{\beta,t-1}(\beta\leq b)$ of $\underline{x}^{t-1}$ whose range is outside the range of the mentioned 2a vectors, and let $\underline{Q}_{\beta\sigma}$ be the similarly situated vectors of $\underline{x}^{\sigma}(\sigma<r-1)$. Let $\underline{Q}_{\beta}=(\underline{Q}_{\beta,t-1}\ldots\underline{Q}_{\beta 2})$ and $Q=(\underline{Q}_1\ldots\underline{Q}_b)$ Then the vectors of $\underline{P}$ and $\underline{Q}$ are independent. For a relation $L\equiv \sum_{\rho\geq t}\sum_{\alpha}\underline{P}_{\alpha\rho}\lambda_{\alpha\rho}+\sum_{\sigma\geq t}\sum_{\beta}\underline{Q}_{\beta\sigma}\mu_{\beta\sigma}=0$, $\mu_{\beta\sigma}\not\equiv_{\beta}0$, would give

$$0=\underline{x}^{t-t-1}\,\underline{L}=\sum_{\alpha}\underline{P}_{\alpha,t-1}\lambda_{\alpha t}+\sum_{\alpha}\underline{P}_{\alpha,t}\lambda_{\alpha,t+1}+\sum_{\beta}\underline{Q}_{\beta,t-1}\mu_{\beta t}\ ,$$

contradicting the hypothesis of the choice of the $\underline{Q}_{\beta,t-1}$. Moreover we have $\underline{x}\underline{Q}=\underline{Q}\begin{pmatrix}\underline{N}_{t-1} & \\ & \underline{N}_{t-1}\end{pmatrix}$, with b nullpotents on the diagonal. Again, we have $\underline{x}^{t-2}\,\underline{P}_{\alpha 2}=\underline{P}_{\alpha t}$, $\underline{x}^{t-2}\,\underline{P}_{\alpha 1}=\underline{P}_{\alpha,t-1}$, $\underline{x}^{t-2}\,\underline{Q}_{\beta 1}=\underline{Q}_{\beta,t-1}$, showing that the 3a+2b independent vectors in $\underline{P}_{\alpha t}$, $\underline{P}_{\alpha,t-1}$, $\underline{P}_{\alpha,t-2}$, $\underline{Q}_{\beta,t-1}$, $\underline{Q}_{\beta t}$ lie in the range of $\underline{x}^{t-2}$, so that the rank of $\underline{x}^{t-2}$ is at least 3a+2b, say 3a+2b+c. Take c vectors $\underline{R}_{\gamma,t-2}$ in $\underline{x}^{t-2}$, whose range is outside the range of the mentioned 3a+2b vectors, and let $\underline{R}_{\gamma\tau}(\gamma\leq c)$ be the sismilarly situated vectors of $\underline{x}^{\tau}(\tau<r-2)$. Then the vectors of $\underline{P},\underline{Q}$ and $\underline{R}$, from $\underline{x}^{t}$, $\underline{x}^{t-1}$ and $\underline{x}^{t-2}$, are independent; and so on, till we have to consider $\underline{x}^{0}=\underline{I}$, which is of rank $n=ra+(r-1)b+\ldots$; and thus we may complete the selected ŝndependent vectors to a set of n. These then form a regular matrix $\underline{T}=(\underline{P}\underline{Q}\underline{R}\ldots)$, which is such that $\underline{x}\underline{T}=\underline{T}\underline{N}$, or $\underline{x}=\underline{T}\underline{N}\underline{T}^{-1}$, where

$$\underline{N} = \begin{pmatrix} \underline{N}_r & & & \\ & \ddots\, \underline{N}_r & & \\ & & \underline{N}_{r-1} & \\ & & & \ddots\, \underline{N}_{r-1} \\ & & & & \ddots \end{pmatrix}$$

is a direct sum of a primitive nullpotents of order r, b of order r-I, and so on.

15. The normal form of a matrix.

A matrix of the form $\underline{x}=\underline{N}+\xi\underline{I}$, where $\underline{N}$ is a classical nullpotent, defines a primitive matrix, of any order, with any characteristic value; it has just one distinct characteristic value , ξ; and it is nonderogatory, having characteristic equation $(\underline{x}-\xi\underline{I})^{r}=\underline{0}$, where r is the order. We may write $C_r(\xi)=\underline{N}_r+\xi\underline{I}_r$, for the primitive matrix of order r, and characteristic value ξ.

THEOREM. Every matrix is similar to a direct sum of primitive matrices.

By §11, we have that there exists a regular matrix $\underline{Q}$ such that $\underline{Q}^{-1}\underline{a}\underline{Q}=\underset{s}{\Theta}\underline{a}_s$, where each component $\underline{a}_s$ in this direct sum has just one characteristic value, say α_s respectively. Thus each $\underline{a}_s-\alpha_s\underline{I}$ is nullpotent, and, by §12, there exists a regular matrix R_s such that $\underline{R}_s^{-1}(\underline{a}_s-\alpha_s\underline{I})\underline{R}_s$ is a direct sum of primitive nullpotents, that is, $\underline{R}_s^{-1}\underline{a}\underline{R}_s$ is a direct sum of classical matrices. If $\underline{R}=\underset{s}{\Theta}\underline{R}_s$, and $\underline{P}=\underline{Q}\underline{R}$, we now have that $\underline{P}^{-1}\underline{a}\underline{P}$ is a direct sum of classical matrices, which defines the classical normal form of the matrix $\underline{a}$.*[10]

* In other words, if ν_x denotes the nullity of any matrix $\underset{\sim}{x}$, we have $\nu_{pq} = \nu_p + \nu_q$. By Sylvesters law of nullity we have $\nu_{pq} \leq \nu_p + \nu_q$. So now it will be sufficient to show that $\nu_{pq} \geq \nu_p + \nu_q$.

14. Projectors and principal spaces.

THEOREM I. *If $p(\tau)$, $q(\tau)$ are polynomials which have no root in common, and $\underline{p}=p(\underline{a})$, $\underline{q}=q(\underline{a})$ where $\underline{a}$ is any matrix, then*

$$\mathfrak{N}_p \cap \mathfrak{N}_q = \mathfrak{O}, \quad \mathfrak{N}_p \oplus \mathfrak{N}_q = \mathfrak{N}_{pq} ,$$

where $\mathfrak{N}_x$ denotes the nullspace of any matrix $\underline{x}$.

* If $\mathfrak{N}_p \cap \mathfrak{N}_q \neq \mathfrak{O}$, then there exist vectors, forming this non-null intersection, which are latent vectors for zero characteristic value of both p and q. Then since $\underline{a}$, $\underline{p}$ and $\underline{q}$ form a commutative set, it follows from the extension theorem of §4, that there exists a vector $\underline{X}$, in the intersection $\mathfrak{N}_p \cap \mathfrak{N}_q$, which is also a latent vector of $\underline{a}$, say with characteristic value α; that is, such that $\underline{pX}=\underline{0}$, $\underline{qX}=\underline{0}$, and $\underline{aX}=\underline{X}\alpha$. But $\underline{aX}=\underline{X}\alpha$ implies that $\underline{0}=\underline{pX}=\underline{X}p(\alpha)$, $\underline{0}=\underline{qX}=\underline{X}q(\alpha)$, or $p(\alpha)=0$, $q(\alpha)=0$; which shows α as a common root of $p(\tau)$, $q(\tau)$, giving a contradiction. Hence $\mathfrak{N}_p \cap \mathfrak{N}_q = 0$; and therefore $\mathfrak{N}_p \oplus \mathfrak{N}_q$ is of rank (at least) $\nu_p + \nu_q$. Now evidently $\mathfrak{N}_p$ and $\mathfrak{N}_q$ are both contained in $\mathfrak{N}_{pq}$, so that $\mathfrak{N}_p \oplus \mathfrak{N}_q \subset \mathfrak{N}_{pq}$ and this implies that $\nu_p + \nu_q = \nu_{pq}$, as was finally required.

THEOREM II. *The principal spaces $\mathfrak{U}_\alpha$, of any matrix $\underline{a}$, form a decomposition of the unit space, $\mathfrak{U}_\alpha$ being the range of the projector $\underline{e}_\alpha$, and also the nullspace of $(\underline{a}-\alpha\underline{I})^{i_\alpha}$.*

We have that $\underline{e}_\alpha$ differs from $\underline{\underline{\Pi}}_{\lambda\neq\alpha}\underline{a}_\lambda^{i_\lambda}$ only by a non-singular matrix factor, so that their nullspaces are the same. Let $\mathcal{N}_\alpha$ denote the nullspace of $\underline{e}_\alpha$, and $\mathfrak{N}_\alpha$ the nullspave of $\underline{a}_\alpha^{i_\alpha}$. Then $\mathfrak{N}_\lambda \cap \mathfrak{N}_\mu = \mathcal{O}$ $(\lambda \neq \mu)$, and $\underset{\alpha}{\oplus}\mathcal{N}_\alpha = 1$, by the previous theorem; and further, $\mathcal{N}_\alpha = \underset{\lambda\neq\alpha}{\oplus}\mathcal{N}_\alpha$ Now $\underline{a}_\alpha^{i_\alpha}\underline{e}_\alpha = 0$, so that $\mathcal{N}_\alpha \supset \mathcal{U}_\alpha$. But the characteristic values of $\underline{e}_\alpha$ are $\delta_{\alpha\lambda}$ of multiplicity m_λ, or the characteristic value zero has multiplicity $\underset{\lambda\neq\alpha}{\Sigma} m_\lambda$, and hence, by §6 the nullity of $\underline{e}_\alpha$ is at most $\underset{\lambda\neq\alpha}{\Sigma} m_\lambda$, or its rank s_α is at least $n - \underset{\lambda\neq\alpha}{\Sigma} m_\lambda = m_\alpha$. Thus $\underset{\alpha}{\Sigma} s_\alpha > \underset{\alpha}{\Sigma} m_\alpha = n$, unless $s_\alpha = m_\alpha$. Thus we have that $\mathcal{N}_\alpha$ and $\mathfrak{U}_\alpha$ are the same rank, m_α; whence it follows that they are the same, since we already have that one contains the other.

NOTES

46 I. Or the grade. See Turnbull and Aitken (I), 47.

48 2. This is immediate, from any matrix having at least one axial vector, when this proposition is viewed as abstractly applying to the endomorphisms of a vector space.

49 3. See Drazin, Dungay and Gruenberg (2) for a proof along different lines.

54 4. This result is immediate when theorem I, § 4, is viewed abstractly as applying to endomorphisms of a vector space.

56 5. Frobenius (3), proved that the characteristic values of commutative matrices could be ordered into sets so that the characteristic values of any rational function of the matrix are given by that function of these sets. Here the sets are identified as consisting of the characteristic values which belong to common axial vectors. See Phillips (4) and MacDuffe(9

58 6. cf. Drazin, Dungay and Gruenberg (2). ⌊Rademacher (6).

62 7. Phillips (4) gives another proof.

80 8. cf. Phillips(4), Wedderburn (5) and Hamberger and Grimshaw (7).

8 9

8 9. Plemelj (8).

84 IO. See MacDuffee (9), Turnbull and Aitken (I), Halmos(IO) Littlewood (II), and Hamberger and Grimshaw (7).

87 II. cf. Phillips (4).

89 I2. It may be noted that the corresponding result of

Phillips (4) requires a certain convergence for all $r+s+,,.<i_{\pi}$, which is more than is asked here.

92 13. The upper modulus is invariant under unitary transformation. If $\underline{\bar{u}}'\underline{u}=\underline{u}\underline{\bar{u}}'=I$, then $|\underline{u}|^{*2}=|\underline{\bar{u}}'\underline{u}|^{*}=I$, so that $|\underline{u}|^{*}=I$. If $\underline{b}=\underline{\bar{u}}'\underline{a}\underline{u}$, then $|\underline{b}|^{*}\leq|\underline{\bar{u}}'|^{*}|\underline{a}|^{*}|\underline{u}|^{*}=|\underline{a}|^{*}$. Similarly $|\underline{a}|^{*}\leq|\underline{b}|^{*}$. Thus $|\underline{a}|^{*}=|\underline{b}|^{*}$; and similarly $|\underline{a}|_{*}=|\underline{b}|_{*}$. See Airiat (12).

REFERENCES

1. H.W. Turnbull and A.C. Aitken, The theory of Canonical Matrices (London 1932).

2. M.P. Drazin, J.W. Dungey and K.W. Gruenberg, Some theorems on commutative matrices, Journ.London Math. Soc.,Vol.26, 1951.

3. G. Frobenius, S.-B. preuss. Akad. Wiss. 1896, I, 601-14.

4. H.B. Phillips, Functions of matrices, Amer.J.Math. Vol.41(1919), 266-278.

5. J.H.M. Wedderburn, Lectures on Matrices (American Math Soc. Colloquium Publ.,XVIII, 1934).

6. H. Rademacher, On a theorem of Frobenius, Studies and Essays Presented to R. Courant, 301-5.

7. H.L. Hamburger and M.E. Grimshaw, Linear Transformation of the n-dimensional vector space (Cambridge 1951).

8. J. Plemelj, Mh. Math. Phys.,Vol.12 (1901), 82-96.

9. C.C. MacDuffee, The Theory of Matrices (New York 1946).

10. P. Halmos, Finite Dimensional Vector Space (Princeton 1948).

11. D.E. Littlewood, A University Algebra (London 1950).

12. S.N. Afriat, Bounds for the characteristic values of matrix functions, Quart.J.Math.Oxford(2),2(1951),81-4

ON AXIAL AND NORMAL MATRIX SETS

1. Introduction.

The matrices of any commutative matrix set have axial vectors in common, which we call the axial vectors of the matrix set. If the axial vectors of a matrix set are sufficient to span the unit space, that is the complete vector space on which the matrices operate, then that set is said to be axially complete, or to be an axial matrix set. It is shown that the axial matrix sets are commutative; that they are those commutative matrix sets which have the rank of each characteristic set equal to its multiplicity, so that, since they are independent, the axial spaces represent a splitting of the unit space; also it is shown that the axial matrix sets are the commutative sets of axial matrices. A matrix set is axially complete if and only if its members can all be expressed as polynomials in some axial matrix.

A commutative matrix set is called normal if it remains commutative when extended by the set of transposed conjugates of its members. Thus a normal matrix is a matrix which commutes with its transposed conjugate. The matrices of a normal set can all be reduced to diagonal form by the same unitary transformation, and each can be expressed as a polynomial in some normal

matrix. The normal matrix sets are the commutative sets of normal matrices. Any matrix is normal if and only if it can be expressed as a polynomial in its transposed conjugate; or, equivalently, if and only if it commutes with every matrix which commutes with its transposed conjugate. The conjugate axial vectors of a normal matrix set are the conjugates of its axial vectors. Any normal matrix set is axially complete, its axial spaces are mutually orthogonal and are identival with the axial spaces of the transposed conjugate set.

2.. *Axiale* matrix sets.

A set of matrices is said to be complete, or alternatively to be an axial set, if its axial vectors together span the unit space, that is to say the complete vector space. Thus any matrix of an axial set is given by its characteristic values for each of the axes of the set.

THEOREM I. The axial sets of matrices are the commutative sets of axial matrices, and they are those commutative matrix sets which have the rank of each characteristic set equal to its multiplicity.*(1)

Let $\underline{s}$ be an axial matrix set, with members $\underline{a},\underline{b},\ldots,$ so that there exist axial vectors $\underline{P}_1,\ldots,\underline{P}_n$ of $\underline{s}$ which form the columns of a regular matrix $\underline{P}=(\underline{P}_1\ldots\underline{P}_n)$. Then we may write $\underline{s}=\underline{P}\sigma\underline{P}^{-1}$, where σ is a set of diagonal matrices $\underline{\alpha},\underline{\beta},\ldots$; and thus $\underline{s}$ is commutative. By theorem 8,III, we have $r_\sigma\leq m_\sigma$ for each characteristic set σ of $\underline{s}$. But, since $\underline{s}$ is complete, we have $\sum_\sigma r_\sigma=n=\sum_\sigma m_\sigma$; and therefore $r_\sigma=m_\sigma$ for each σ. Now suppose $\underline{s}$ is commutative and such that $r_\sigma=m_\sigma$ for each σ. Then the axial spaces of $\underline{s}$, which, by theorem 9, are all independent, span a space of rank $\sum_\sigma r_\sigma=\sum_\sigma m_\sigma=n$, which is the maximum rank. Hence $\underline{s}$ is complete. Finally it will be proved by induction that the commutative sets of complete matrices are complete, by showing the completeness of any commutative

set of complete matrices $\underline{a},\underline{b},\underline{c},\ldots$ on the hypothesis of the completeness of any subset $\underline{b},\underline{c},\ldots$ where one member $\underline{a}$ has been omitted. Thus, by hypothesis, there exists a regular matrix $\underline{P}$ for which $\underline{\beta}=\underline{P}^{-1}\underline{b}\underline{P}$, $\underline{\gamma}=\underline{P}^{-1}\underline{c}\underline{P}$,... are all diagonal; and let σ_i be the set of ith elements $\beta_i, \gamma_i, \ldots$ Now $\underline{A}=\underline{P}^{-1}\underline{a}\underline{P}$ commutes with $\underline{\beta},\underline{\gamma},\ldots$ since a commutesw with $\underline{b},\underline{c},\ldots$; and it follows that $A_{ij}=0$ $(\sigma_i \neq \sigma_j)$. Therefore, since $\underline{a}$ and hence $\underline{A}$ is complete, there exists a regular matrix $\underline{\zeta}$ which is such that $\zeta_{ij}=0$ $(\sigma_i \neq \sigma_j)$, so that $\underline{\zeta}$ commutes with $\underline{\beta},\underline{\gamma},\ldots$, and such that $\underline{\alpha}=\underline{\zeta}^{-1}\underline{A}\underline{\zeta}$ is diagonal. Let $\underline{P}\underline{\zeta}=\underline{R}$. Then, since $\underline{\beta}=\underline{\zeta}^{-1}\underline{\beta}\underline{\zeta}=\underline{R}^{-1}\underline{b}\underline{R}$, $\underline{\gamma}=\underline{\zeta}^{-1}\underline{\gamma}\underline{\zeta}=\underline{R}^{-1}\underline{c}\underline{R}$,... and $\underline{\alpha}=\underline{R}^{-1}\underline{a}\underline{R}$ are all diagonal, we have that the columns of the regular matrix $\underline{R}$ form a eomplete set of axes for the matrix set $\underline{a},\underline{b},\underline{c},\ldots$; that is, $\underline{a},\underline{b},\underline{c},\ldots$ form a complete matrix set.

Any characteristic set is called <u>simple</u> if it has multiplicity unity, and otherwise <u>multiple</u>. Since $r_\sigma>0$ if $m_\sigma>0$, and since $r_\sigma \leq m_\sigma$, it follows that any simple characteristic set is of rank unity, or, equivalently, that the corresponding axial space forms a ray. A commutative matrix set is called simple if it has only simple characteristic sets.

COROLLARY 1. <u>All simple commutative sets are complete</u>.

For the axial spaces are independent rays of the maximum number.

COROLLARY ii. Any axial matrix set has a commutative extensionw which includes a simple matrix.

COROLLARY iii. Matrices which commute with a simple matrix are polynomials in that matrix.

Suppose $\underline{a}$ is simple and commutes with $\underline{b}$. Then t there is a regular matrix $\underline{P}$ such that $\underline{\alpha}=\underline{P}^{-1}\underline{a}\underline{P}$ is simple, diagonal, and commutes with $\underline{\beta}=\underline{P}^{-1}\underline{b}\underline{P}$, so that $\underline{\beta}$ is also diagonal. If p is any polynomial such that $\beta_\nu=p(\alpha_\nu)$ for each ν, then $\beta=p(\alpha)$, and

$$\underline{b}=\underline{P}\underline{\beta}\underline{P}^{-1}=\underline{P}p(\underline{\alpha})\underline{P}^{-1}=p(\underline{P}\underline{\alpha}\underline{P}^{-1})=\underline{p}(\underline{a}).$$

COROLLARY iv. The matrices of an axial set are all polynomials in some axial matrix.

COROLLARY v. Only triangular matrices commute with simple triangular matrices.

If $\underline{a}$ is simple and commutes with $\underline{b}$, then $\underline{b}$ is a polynomial in $\underline{a}$, or $\underline{b}=p(\underline{a})$, by the corollary iii. Hence if $\underline{a}$ is triangular, so is $\underline{b}$.

§F. Normal matrix sets.*(2)

A matrix set s is called normal if the augmented set $\bar{s}'us$, formed by s joined to the set $\bar{s}'$ of the transposed conjugates $\hat{\bar{a}'}$ of its members $\hat{a}$, is commutative.

THEOREM I. The matrices of a normal set can all be reduced to diagonal form by the same unitary transformatio

If s is normal, so that $\bar{s}'us$ is commutative, there is, by theorem 7,I, a unitary matrix U, such that if a is any matrix of s, then the matrices $\alpha=U^{-1}aU$ and $\bar{\alpha}'=U^{-1}\bar{a}'U$ are triangular, in the same sense. But if a matrix and its transposed conjugate are triangular in the same sense, they must be diagonal. Hence α is a diagonal matrix.

It is now shown that if s is a normal matrix set, then $s=U\sigma U^{-1}$, where U is a unitary matrix, and σ a set of diagonal matrices. Conversely, every matrix set of this form is normal. The ith column U_i of U gives an axial vector of s, and σ_i, the set of ith elements of the diagonals in σ, is the corresponding characteristic set. If σ is any characteristic set of s, then a unitary base P_σ for the corresponding axial space $\mathfrak{A}_\sigma$ of s is given by $P_\sigma=(U_i)_{\sigma_i=\sigma}$, where $\bar{P}'_\sigma P_\omega=I_{m_\sigma}$. Further, $\bar{P}'_\sigma P_\omega=0$ $(\sigma\neq\omega)$, so that the axial spaces of s are mutually orthogonal.

A vector is represented by a single column array, and a conjugate vector by a single row. If X is any vector,

then $\bar{X}'$ is a conjugate vector, called the conjugate of X. Just as X is an axial vector of a matrix a if $aX=X\alpha$ for a scalar α, so $\bar{X}'$ is an axial conjugate vector of a if $\bar{X}'a=\alpha\bar{X}'$, where α is scalar.

THEOREM II. The normal matix sets are complete, their axial spaces are mutually orthogonal, and their axial conjugate vectors are the conjugates of their axial vectors.

This restates the last theorem.

THEOREM III. The normal matrix sets are the axial matrix sets for which the axial conjugatev vectors are the conjugates of the axial vectors.

Let s be complete. Then there is a regular matrix P such that $aP=P\alpha$, for any matrix a of s and a diagonal matrix α; and if the conjugacy relation holds, then we also have $\bar{P}'a=\alpha\bar{P}'$. Let σ denote the set of these diagonal matrices α, one for each member of s, and let σ_i denote the set of the ith elements of the matrices of σ. Then $\bar{P}'P\alpha=\bar{P}'aP=\alpha\bar{P}'P$, so that $H\alpha=\alpha H$, where $H=\bar{P}'P$ for every α of σ. Hence $H_{ij}=0$ $(\sigma_i\neq\sigma_j)$, and there is a regular matrix Q such that $Q_{ij}=0$ $(\sigma_i\neq\sigma_j)$, so that Q commutes with every α of σ, and is such that $\bar{Q}'Q=H$. Let $PQ^{-1}=U$. Then $\bar{U}'U=I$, so U is unitary; and $aU=aPQ^{-1}=P\alpha Q^{-1}=PQ^{-1}\alpha=U\alpha$, for

each $\underline{a}$ of $\underline{s}$, and the corresponding $\underline{\alpha}$ of $\underline{\sigma}$. Thus $\underline{s}$ is of the form $\underline{s}=\underline{U}\underline{\sigma}\bar{\underline{U}}'$, where $\underline{U}$ is unitary, and $\underline{\sigma}$ a set of diagonals. Hence $\underline{s}$ is normal. The complementary part of the present theorem is contained in theorem I.

THEOREM IV. *The normal matrix sets are the complete matrix sets whose axial spaces are mutually orthogonal.*

Take a maximal orthogonal set of unit axial vectors in each axial space of $\underline{s}$. Then $\underline{s}$ is complete and its axial spaces are mutually orthogonal, these axial vectors form a complete orthogonal set, forming the columns of a unitary matrix $\underline{U}$. Since the columns of $\underline{U}$ are axial vectors of every matrix $\underline{a}$ of $\underline{s}$, we have $\underline{a}\underline{U}=\underline{U}\underline{\alpha}$, and $\underline{a}=\underline{U}\underline{\alpha}\bar{\underline{U}}'$, where $\underline{\alpha}$ is diagonal. Hence $\underline{s}$ is normal.

THEOREM V. *A matrix is normal if and only if it can be expressed as a polynomial in its transposed conjugate.*

If a matrix is normal, there is a unitary matrix $\underline{U}$ such that $\underline{a}=\underline{U}\underline{\alpha}\bar{\underline{U}}'$, $\bar{\underline{a}}'=\underline{U}\bar{\underline{\alpha}}\bar{\underline{U}}'$, where $\underline{\alpha}$ is diagonal. Let p be any polynomial such that $\alpha_r=p(\bar{\alpha}_r)$. Then $\underline{\alpha}=p(\bar{\underline{\alpha}})$ and

$$\underline{a}=\underline{U}\underline{\alpha}\bar{\underline{U}}'=\underline{U}p(\bar{\alpha})\bar{U}'=p(\underline{U}\bar{\alpha}\bar{\underline{U}}')=p(\bar{\underline{a}}');$$

and, besides, if p denotes the polynomial with conjugate coefficients, then $\bar{\underline{a}}'=p(\underline{a})$. Conversely, if $\underline{a}=p(\bar{\underline{a}}')$, for some polynomial p, then $\underline{a}$ and $\bar{\underline{a}}'$ commute, so that $\underline{a}$ is normal.

COROLLARY i. A matrix is normal if and only if it commutes with every matrix which commutes with is transposed conjugate.

If a matrix $\underset{\sim}{a}$ is normal, then $\underset{\sim}{a}=p(\bar{\underset{\sim}{a}}')$, where p is some polynomial, by the present theorem. Hence any matrix commutes with $\underset{\sim}{a}$ if it commutes with $\bar{\underset{\sim}{a}}'$. Conversely if a commutes with every matrix which commutes with $\bar{\underset{\sim}{a}}'$, then it commutes with $\bar{\underset{\sim}{a}}'$, and is normal.

COROLLARY ii. The normal matrix sets are the commutative sets of normal matrices.

If a matrix set is normal, then so is each of its members. If every matrix of a set $\underset{\sim}{s}$ is normal, and $\underset{\sim}{s}$ is commutative, then $\bar{\underset{\sim}{s}}'u\underset{\sim}{s}$ is commutative, by the above corollary; and hence $\underset{\sim}{s}$ is normal.

COROLLARY iii. If a matrix is is normal, and triangular, then it is diagonal.*(3)

If $\underset{\sim}{a}$ is triangular, then $\bar{\underset{\sim}{a}}'$ is also triangular, but in the opposite sense. If, further, $\underset{\sim}{a}$ is normal, then $\underset{\sim}{a}=p(\bar{\underset{\sim}{a}}')$, where p is some polynomial, by the present theorem, so that a must also be triangular, in the same sense as $\bar{\underset{\sim}{a}}'$. Hence, since $\underset{\sim}{a}$ is triangular in each sense, it must be diagonal.

NOTES

99 I. That commutative sets of axial matrices are axial matrix sets has been shown for pairs of matrices by Jeffreys and Jeffreys (1), abd for any number by Drazin Dungey ang Gruenberg (2).

102 2. Normal matrices have been studied by several authors. It is possible that the term originated with Toeplitz (3) See Hamberger and Grimshaw (4); Drazin (5); Halmos (6), 130, et seq.

103 3. This is oftern taken as a preliminary result for proving the axial completeness of normal matrices. See MacDuffee (7), 76, for example.

REFERENCES

I. H. Jeffereys and M. Jeffreys, Methods of Mathematical Physics (Cambridge I94).

2. M.P. Drazin, J.W. Dungay and K.W. Gruenberg, Some theorems on commutative matrices, Journ.London Math. Soc.,Vol.26, I95I.

3. O. Toeplitz, Math. Z. Vol.2(I9I8),I90.

4. H.L. Hamberger and M.E. Grimshaw, Linear transformations of the n-dimensional vector space (Cambridge I95I).

5. M.P. Drazin, On diagonable and normal matrices, Quart. J.Math.Oxford(2),2(I95I),I89-98.

6. R. Halmos, Finite Dimensional Vector Space (Princeton I948), I00, et seq.

7. CC. MacDuffee, The Theory of Matrices (New York I946).

COMPOSITE MATRICES

I. Introduction.

It has been proved by Williamson[1] that if a matrix $\underset{\sim}{F}$, of order mn, has a partitioning into m^2 submatrices $\underset{\sim}{f}_{ij}$ of order n, each of which is given as a rational function $\underset{\sim}{f}_{ij}=f_{ij}(\underset{\sim}{a})$, of non-singular denominator, in some matrix $\underset{\sim}{a}$ with characteristic values $\alpha_1,\ldots,\alpha_n$, then the mn characteristic values of the matrix $\underset{\sim}{F}$ are the mn characteristic values of the n mth order matrices $\underset{\sim}{f}^{(k)}$, where the matrix $\underset{\sim}{f}^{(k)}$ has i,jth element $f_{ij}(\alpha_k)$. When m=1 the assertion is vacuous; when m=1 it becomes the theorem of Frobenius[2] on the characteristic values of a rational function of a matrix. Our investigation will show a proposition of this form, where the conclusion depends only on the commutivity of the partitioned components, which in the case here is implied by their all being rational functions in a single matrix. To state this more explicitly, Williamson's result will appear as a consequence of the theorem that if $\underset{\sim}{A}$ is a matrix of order mn, partitioned into m^2 submatrices $\underset{\sim}{a}_{ij}$ of order n, then the mn characteristic values of $\underset{\sim}{A}$ are the mn characteristic value of the n mth order matrices $\underset{\sim}{\alpha}^{(k)}$, where the matrix $\underset{\sim}{\alpha}^{(k)}$ has i,jth element $\alpha_{ij}^{(k)}$, the kth characteristic value of $\underset{\sim}{a}_{ij}$ in some ordering.

Elsewhere[3] there has been occasion to use the identity

$$\left|-\begin{pmatrix} \underset{\sim}{0} & \underset{\sim}{I} & \underset{\sim}{x} \\ \underset{\sim}{I} & \underset{\sim}{y} & \underset{\sim}{v} \\ \underset{\sim}{u} & \underset{\sim}{v} & \underset{\sim}{b} \end{pmatrix}\right| = \left|-(-\underset{\sim}{b}+\underset{\sim}{v}\underset{\sim}{x}+\underset{\sim}{y}\underset{\sim}{u}-\underset{\sim}{y}\underset{\sim}{a}\underset{\sim}{x})\right| ,$$

where the partitioned matrices $\mathfrak{a}$, $\mathfrak{C}$ are square and of the same order. The negative signs here have been inserted as a device, so as to avoid making reference to the matrix orders; or as much, to display simply the double determinental form occurring in the relation. When the component matrices are all taken to be of order unity, this identity reduces vacuously to a statement of definition. However, in the general case, the form of a determinant operation involving matrices as elements is apparent. But to make such an interpretation, it is required that the unit matrices involved be regarded as freed of a fixed order, so that the multiplications into which they enter might be regarded as possible. Now, in the case when the component matrices are all of the same order, this determinant may be taken in a more normal way, provided some convention has been decided for the order of the multiplications of the factors in each term, when these factors are not commutative. The considered identity then shows a twice repeated determinant operation on the right as equivalent to a simple determinant operation on the left. This suggests a form of proposition , of which our investigation will show an example.

Thus we are to be concerned with determinants and matrices whose elements are themselves matrices.

2. Definitions and propositions.

A composite matrix of order n and degree m is an nth order matrix whose elements, called its components, are themselves matrices of order m. Let the i,jth component $\underset{\sim}{A}^*_{ij}$ of a composite matrix $\underset{\sim}{A}^*$ of order n and degree m have r,sth element A^{*rs}_{ij}. Then correlative to $\underset{\sim}{A}^*$ there is defined the dual composite matrix $\underset{\sim}{A}_*$, of order m and degree n, whose r,sth component A^{*rs}_{*} has i,jth element A^{*rs}_{ij}.

Any composite matrix has a ground matrix, which is that simple matrix from which it may be derived by a partitioning. A pair of composite matrices of the same order and degree have a product, which is also a composite matrix of that order and degree. The ground matrix of a product of composite matrices is the product of their duals, and its dual is the product of their duals.

A composite matrix is called commutative if its components all commute, and doubly-commutative if its dual is also commutative.

The ground matrices of a dual pair of composite matrices have the same determinant, defining their ground determinant. The determinant of a commutative composite matrix of order n and degree m is a determinant of order n whose value is a matrix of order m, and the determinant of this determinant defines the double-determinant of that composite matrix.

When the components of a commutative composite matrix are replaced by their respective characteristic values which are in a latently associated set,[4] that is to say which form a characteristic set[5] for the commutative set of components, a characteristic dual component of the composite matrix is obtained, with a multiplicity attached to it equal to the multiplicity of the characteristic set formed by its elements. The characteristic dual components are linearly dependent on the dual components. Thus if a composite matrix is doubly commutative, so there is a set of characteristic components defined besides the characteristic dual components, then the characteristic, and the characteristic dual components form commutative matrix sets such that the characteristic array of one, formed by the array of its characteristic sets, is the transpose of that of the other. The characteristic v values or the ground matrix of any commutative composite matrix are the characteristic values of its characteristic dual components, with multiplicities attached correspondingly, and in the case when the composite matrix is doubly commutative, these characteristic values are also determined as the elements of the characteristic array.

The double determinant of any commutative composite matrix is equal to its ground determinant. This shows a twice repeated determinant operation, on one side, as equivalent to a simple determinant operation, on the other.

For any doubly-commutative composite matrix, there is defined a composite, and also a dual composite characteristic polynomial, which are null when the characteristic and the dual characteristic components are substituted for the argument. The determinant of these polynomials, when the argument is taken as a scalar indeterminate, are both equal to the ground characteristic polynomial, for the composite matrix and its dual.

3. Demonstrations.

THEOREM I. The ground matrix of a product of composite matrices is the product of their ground matrices, and the dual is the product of their duals.

If $\underset{\sim}{A}^*$, $\underset{\sim}{B}^*$ is any pair of composite matrices, of the same order and degree, their product is a similar composite matrix $\underset{\sim}{C}^* = \underset{\sim}{A}^* \underset{\sim}{B}^*$. If $\underset{\sim}{a}^*$, $\underset{\sim}{b}^*$ and $\underset{\sim}{c}^*$ denote the ground matrices of $\underset{\sim}{A}^*$, $\underset{\sim}{B}^*$ and $\underset{\sim}{C}^*$, respectively, then the theorem on the multiplication of partitioned matrices [6] gives $\underset{\sim}{c}^* = \underset{\sim}{a}^* \underset{\sim}{b}^*$

Now the i,jth component of $\underset{\sim}{C}^*$ is $\underset{\sim}{C}^*_{ij} = \sum_{\lambda} \underset{\sim}{A}^*_{i\lambda} \underset{\sim}{B}^*_{\lambda j}$, with r,sth element $C^{rs}_{ij} = \sum_{t,\lambda} A^{rt}_{i\lambda} B^{ts}_{\lambda j}$ which is the i,jth element of the r,sth component $\underset{\sim}{C}^{rs}_*$ of the dual $\underset{\sim}{C}_*$ of $\underset{\sim}{C}^*$. Thus $\underset{\sim}{C}^{rs}_* = \sum_{t} \underset{\sim}{A}^{rt}_* \underset{\sim}{B}^{ts}_*$, which is the r,sth component of the product $\underset{\sim}{A}_* \underset{\sim}{B}_*$, so that $\underset{\sim}{C}_* = \underset{\sim}{A}_* \underset{\sim}{B}_*$.

These results can now be extended by induction to products with any finite number of factors.

Let $\underset{\sim}{A}^*$ be a commutative composite matrix, and let $\underset{\sim}{\alpha}^{\mu}_{ij}$ (μ=I,...,m) be the characteristic sets of the commutative set of components $\underset{\sim}{A}^*_{ij}$ of $\underset{\sim}{A}^*$, enumerated by μ with repetitions equal to their multiplicities as characteristic sets. Then the characteristic dual components of $\underset{\sim}{A}^*$ are the m matrices $\underset{\sim}{\alpha}^{\mu}$ (μ=I,...,m) of order n, where the matrix $\underset{\sim}{\alpha}^{\mu}$ has i,jth element α^{μ}_{ij}. A latent vector $\underset{\sim}{X}^*_{\mu}$ of the components of $\underset{\sim}{A}^*$ corresponding to the charac-

teristic set formed by the elements of $\underset{*}{\alpha}^{\mu}$, that is to say such that $\underset{ij}{A^{*}}\,\underset{\mu}{X^{*}} = \underset{\mu}{X^{*}}\,\alpha^{\mu}_{ij}$, we shall call a component latent vector of $\underset{}{A}^{*}$ corresponding to the characteristic dual component $\underset{*}{\alpha}^{\mu}$.

THEOREM II. If $\underset{}{A}^{*}$ is a commutative composite matrix, and $\underset{\mu}{X^{*}}$ is a component latent vector corresponding to the characteristic dual component $\underset{*}{\alpha}^{\mu}$, then $\sum_{s} \underset{*}{A}^{rs}\,X^{*}_{\mu,s} = \underset{*}{\alpha}^{\mu}\,\underset{\mu,t}{X^{*}}$.

For we have $\underset{ij}{A^{*}}\,\underset{\mu}{X^{*}} = \underset{\mu}{X^{*}}\,\alpha^{\mu}_{ij}$, by definition of $\underset{\mu}{X^{*}}$; which, more explicitly, is $\sum_{s} \underset{ij}{A^{rs}}\,X^{*}_{\mu,s} = \underset{\mu,r}{X^{*}}\,\alpha^{\mu}_{ij}$; and this may be written $\sum_{s} \underset{*}{A}^{rs}\,\underset{\mu,s}{X^{*}} = \underset{\mu,r}{X^{*}}\,\underset{*}{\alpha}^{\mu}$.

COROLLARY i. The characteristic dual components of a commutative composite matrix are linear combinations of the dual components.

COROLLARY ii. The characteristic, and the characteristic dual components of a doubly commutative composite matrix form commutative matrix sets.

THEOREM III. The characteristic value of the ground matrix of a commutative composite matrix are the characteristic values of its characteristic dual components.

It is required to prove that if $\underset{}{A}^{*}$ is a commutative composite matrix of order n and degree m, with ground matrix $\underset{}{a}^{*}$ and characteristic dual components $\underset{*}{\alpha}^{\mu}(\mu=1,\ldots,m)$

then the mn characteristic values of $\underset{*}{a}^*$ are those of the m nth order matrices $\underset{*}{\alpha}^\mu$. Since $\underset{*}{A}^*$ is commutative, its components all have triangular transforms by some unitary matrix $\underset{*}{U}$, such that α_{ν}^{μ} appears as the μth diagonal element in the transform of $\underset{\nu}{A}^*$. Let $\underset{*}{U}^*$ denote the diagonal composite matrix, of order n, whose components are $\underset{\nu}{U}^*=U\delta_{\nu}$, and let $\underset{*}{U}_*$ be the composite matrix dual to $\underset{*}{U}^*$, with components $\underset{*}{U}^{rs}=U_{rs}\underset{=}{I}$, all of which are scalar matrices of order n. Let composite matrices $(\underset{*}{U}^{-1})^*$ and $(\underset{*}{U}^{-1})_*$ be similarly formed from the matrix U^{-1}. Then $\underset{*}{B}^* = \underset{*}{U}^*\underset{*}{A}^*(\underset{*}{U}^{-1})^*$ is a composite matrix whose components are all triangular matrices, and hence its dual, which, by theorem I, is $\underset{*}{B}_* = \underset{*}{U}_*\underset{*}{A}_*(\underset{*}{U}^{-1})_*$, is a triangular composite matrix whose μth diagonal component is the characteristic dual component α_*^μ of $\underset{*}{A}^*$. Now the ground matrices $\underset{*}{a}^*$ of $\underset{*}{A}^*$ and $\underset{*}{b}^*$ of $\underset{*}{B}^*$ have the same characteristic values; for the ground matrix of $(\underset{*}{U}^{-1})^*$ is the reciprocal $\underset{*}{u}^{*-1}$ of the ground matrix $\underset{*}{u}^*$ of $\underset{*}{U}^*$, showing that $\underset{*}{b}^* = \underset{*}{u}^*\underset{*}{a}^*\underset{*}{u}^{*-1}$, by theorem I, or that $\underset{*}{a}^*$ and $\underset{*}{b}^*$ are similar. Moreover the ground matrix b_* of $\underset{*}{B}_*$ has the same characteristic values as the ground matrix $\underset{*}{b}^*$ of $\underset{*}{B}^*$. But the characteristic values of the ground matrix of a triangular composite matrix are the characteristic values of its diagonal components; for the characteristic polynomial of the ground matrix is seen to be the product of the characteristic polynomials of the diagonal components. Thus the characteristic

values of $\underset{*}{b}$, or equivalently the characteristic values of $\underset{}{b}^*$ and hence those of $\underset{}{g}^*$, are the characteristic values of the diagonal components of $\underset{*}{B}$, that is to say the matrices $\underset{*}{\alpha}^{\mu}$; and thus the theorem is proved.

For the next theorem the following lemma will be required:

LEMMA. *If $|x\ y|$ denotes a determinant partitioned in its columns, and a is a square matrix, with the same number of columns as x, then $|xa\ y| = |a||x\ y|$.*

If x is replaced by xa, the minors on the columns of x are multiplied by the determinant of a; and since, by Laplace's theorem, a determinant is a linear function of the minors on any set of its columns, it follows that $|x\ y|$ is replaced by $|a||x\ y|$.

THEOREM IV. *The double determinant of a commutative composite matrix is equal to its ground determinant.*

The proof is by induction on the order of the composite matrix. As the initial proposition, we show that if a, b, c and d are all square and of the same order, and $dc=cd$, then

$$\left|\begin{matrix} a & b \\ c & d \end{matrix}\right| = |ad-bc| \quad .$$

Thus, by the lemma,

$$|\underline{c}|\begin{vmatrix}\underline{a} & \underline{b}\\ \underline{c} & \underline{d}\end{vmatrix} = \begin{vmatrix}\underline{ac} & \underline{b}\\ \underline{cc} & \underline{d}\end{vmatrix};$$

and then, by elementary operations on the columns, which leave the value of any determinant unchanged, this becomes

$$\begin{vmatrix}\underline{ac}-\underline{bc} & \underline{b}\\ \underline{cd}-\underline{dc} & \underline{d}\end{vmatrix} = \begin{vmatrix}\underline{ad}-\underline{bc} & \underline{b}\\ \underline{0} & \underline{d}\end{vmatrix}$$

$$= |\underline{c}|\,|\underline{ad}-\underline{bc}| \quad .$$

Thus we have the required result, provided $|\underline{c}| \neq 0$; but then this identity must also hold when $|\underline{c}| = 0$, by the principle of the irrelevance of algebraic inequalities. [8]

Now take, as the inductive hypothesis, the hypothesis of the validity of the theorem for composite matrices of order n-I. and let $\underline{A}$ be any commutative composite matrix of order n, with ground matrix $\underline{a}$. Then it is required to prove that $\|\underline{A}\| = |\underline{a}|$, under the commutivity conditions of the theorem, and on the inductive hypothesis.

By definition,

$$|\underline{A}| = \underline{A}_{11}|\underline{B}_1| - \underline{A}_{12}|\underline{B}_2| + \ldots\ldots + (-1)^{n-1}\underline{A}_{1n}|\underline{B}_n|,$$

this being the expansion of the determinant of $\underline{A}$ on the components in its first row, where $\underline{B}_r$ denotes the composite matrix of order n-I formed from all the rows of $\underline{A}$ excepting the first, and all the columns of $\underline{A}$ excepting the rth. If $\underline{b}_r$ denotes the ground matrix of $\underline{B}_r$, then, by the inductive hypothesis, we have $\|\underline{B}_r\| = |\underline{b}_r|$.

Now, by the lemma, and the theorem on the multiplication of partitioned matrices, we have

$$\|B_1\|\,|a| = \begin{vmatrix} A_{11}|B_1| & A_{12} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1}|B_1| & A_{n2} & \cdots & A_{nn} \end{vmatrix};$$

and then, by elementary operations on the columns, leaving the value unchanged, this becomes

$$\begin{vmatrix} A_{11}|B_1| - \cdots + (-1)^{n-1} A_{1n}|B_n| & A_{12} & \cdots & A_{1n} \\ A_{21}|B_1| - \cdots + (-1)^{n-1} A_{2n}|B_n| & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1}|B_1| - \cdots + (-1)^{n-1} A_{nn}|B_n| & A_{n2} & \cdots & A_{nn} \end{vmatrix}.$$

The commutations conditions of the theorem now give that all the components in the first column here are null, excepting the first, which is $|A|$; for they are all the determinants of composite matrices with two identical rows, and vanish in consequence of the com utation relations. Hence the expression for the determinant reduces to

$$\left|\begin{array}{c|ccc} |A| & A_{11} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & b_1 & & \end{array}\right| = |b_1|\,\|A\| .$$

Thus, since $|b_1| = \|B_1\|$, it is shown that

$$|b_1|\,\|A\| = |b_1|\,|a| ;$$

and thus we have $\|A\| = |a|$, provided $|b_1| \neq 0$. But the inequality required here is irrelevant to this conclusion. Hence we have proved $\|A\| = |a|$, as required, and the induction is complete.

If $\underset{\sim}{a}$ is any matrix, with elements $a_{\mu\nu}$, write $\underset{\sim}{a}(\alpha)$ for the matrix with elements $\alpha\delta_{\mu\nu} - a_{\mu\nu}$. Then $|\underset{\sim}{a}(\alpha)|$ is the characteristic polynomial of $\underset{\sim}{a}$.

THEOREM V. *If $\underset{\sim}{A}^*$, $\underset{\sim}{A}_*$ is a dual pair of commutative composite matrices, $\underset{\sim}{a}^*$ and $\underset{\sim}{a}_*$ their ground matrices, and $\underset{\sim}{\alpha}^\mu_*$, $\underset{\sim}{\alpha}^*_\nu$ their characteristic components, then*

$$\prod_{\mu=1}^{m} |\underset{\sim}{\alpha}^\mu_*(\alpha)| = |\underset{\sim}{a}^*(\alpha)| = |\underset{\sim}{a}_*(\alpha)| = \prod_{\nu=1}^{n} |\underset{\sim}{\alpha}^*_\nu(\alpha)| \ .$$

For $\underset{\sim}{a}^*$ and $\underset{\sim}{a}_*$ have the same characteristic values, and, by theorem III, these are given as the characteristic values of the matrices $\underset{\sim}{\alpha}^\mu_*(\mu=1,\ldots,m)$ and $\underset{\sim}{\alpha}^*_\nu(\nu=1,\ldots,n)$. Hence, by considering the roots of the characteristic polynomials, we have the required result.

The characteristic values of $\underset{\sim}{a}^*$ or $\underset{\sim}{a}_*$ define the ground characteristic values for the dual pair.

Let $\underset{\sim}{X}^*_\mu$ and $\underset{\sim}{X}^\nu_*$ be component latent vectors of $\underset{\sim}{A}^*$ and $\underset{\sim}{A}_*$, such that

$$\underset{\sim}{A}^*_{\mu\nu}\underset{\sim}{X}^*_\mu = \underset{\sim}{X}^*_\mu \alpha^\mu_{\nu\nu} \ , \text{ and } \ \underset{\sim}{A}^{rs}_*\underset{\sim}{X}^\nu_* = \underset{\sim}{X}^\nu_* \alpha^{rs}_\nu \ .$$

Then, as in theorem II, we have

$$\sum_s \underset{\sim}{A}^{rs}_*\underset{\sim}{X}^*_{\mu,s} = \alpha^\mu_{*\mu,r}\underset{\sim}{X}^*_{\mu,r} \ , \text{ and } \ \sum_j \underset{\sim}{A}^*_{\mu\nu}\underset{\sim}{X}^{\nu,j}_* = \alpha^*_\nu \underset{\sim}{X}^{\nu,i}_* \ ;$$

from which it may be deduced that the characteristic values of $\underset{\sim}{\alpha}^\mu_*$ and $\underset{\sim}{\alpha}^*_\nu$ are the values $\alpha^\mu_{*\nu}$ $(\nu=1,\ldots,n)$ and $\alpha^{*\mu}_\nu$ $(\mu=1,\ldots,m)$ which satisfy the relations [9]

$$\sum_s \alpha^{rs}_\nu X^*_{\mu,s} = \alpha^\mu_{*\nu} X^*_{\mu,r} \text{ and } \sum_j \alpha^\mu_{\nu\nu} X^{\nu,j}_* = \alpha^{*\mu}_\nu X^{\nu,i}_* \ .$$

We further have that the corresponding latent vectors are

$\underset{*}{\overset{\nu}{X}}$ and $\underset{\mu}{\overset{*}{X}}$, that is

$$\underset{*}{\overset{\mu}{\alpha}}\,\underset{*}{\overset{\nu}{X}}=\underset{*}{\overset{\nu}{X}}\,\underset{*\nu}{\overset{\mu}{\alpha}} \quad \text{and} \quad \underset{\nu}{\overset{*}{\alpha}}\,\underset{\mu}{\overset{*}{X}}=\underset{\mu}{\overset{*}{X}}\,\underset{\nu}{\overset{*\mu}{\alpha}} ,$$

or equivalently,

$$\sum_j \underset{ij}{\overset{\mu}{\alpha}}\,\underset{*}{\overset{\nu,j}{X}} = \underset{*\nu}{\overset{\mu}{\alpha}}\,\underset{*}{\overset{\nu,i}{X}} \quad \text{and} \quad \sum_s \underset{\nu}{\overset{*s}{\alpha}}\,\underset{\nu,s}{\overset{*}{X}} = \underset{\nu,t}{\overset{*}{X}}\,\underset{\nu}{\overset{*\mu}{\alpha}} .$$

By comparison with the previous relation, this gives

$$\underset{*\nu}{\overset{\mu}{\alpha}} = \underset{\nu}{\overset{*\mu}{\alpha}} .$$

Now $\underset{*\nu}{\overset{\mu}{\alpha}}$ ($\nu=1,\ldots,n;\mu=1,\ldots,m$) and $\underset{\nu}{\overset{*\mu}{\alpha}}$ ($\mu=1,\ldots,m;\nu=1,\ldots,n$) display the sets of characteristic sets [10] of the commutative matrix sets $\underset{*}{\overset{\mu}{\alpha}}$ ($\mu=1,\ldots,m$) and $\underset{\nu}{\overset{*}{\alpha}}$ ($\nu=1,\ldots,n$), these being commutative by theorem II, corollary ii, and they define the characteristic arrays for the doubly commutative dual pair $\underset{}{\overset{*}{A}}$ and $\underset{*}{A}$. Thus we have the following result:

THEOREM VI. The characteristic arrays of each of a dual pair of doubly-commutative composite matrices are the transposes of each other.

Now we may write

$$\underset{*\nu}{\overset{\mu}{\alpha}} = \underset{\nu}{\overset{\mu}{\alpha}} = \underset{\nu}{\overset{*\mu}{\alpha}} ,$$

and say that $\underset{\nu}{\overset{\mu}{\alpha}}$ defines the characteristic array for the dual pair, whose elements give the ground characteristic values; and give priority to either index, according as the pair is viewed from one member or the other; and we have:

THEOREM VII. The ground characteristic values of a doubly commutative dual pair of composite matrices are the elements of its characteristic array.

The section of the array by the index ν form the characteristic sets of the commutative matrix set $\underset{*}{\alpha}^{\mu}$ $(\mu=1,\ldots,m)$, and the section by μ, of the $\underset{*}{\alpha}^{*}_{\nu}$ $(\nu=1,\ldots,n)$. Thus we further have:

THEOREM VIII. *The complementary sections of the characteristic array of a doubly-commutative dual pair of composite matrices are the characteristic sets of the commutative matrix sets formed by the complementary system of characteristic components.*

Now consider again the sets of relations

$$\sum_{s} \underset{*}{A}^{rs} \, \underset{\mu,s}{X^{*}} = \underset{*}{\alpha}^{\mu} \underset{\mu,r}{X^{*}} \quad \text{and} \quad \sum_{j} \underset{*}{A}^{*}_{\nu,j} \, \underset{*}{X}^{\nu,j} = \underset{\nu}{\alpha}^{*} \underset{*}{X}^{\nu,i} ,$$

obtained from theorem II. Since all the components $\underset{*}{A}^{rs}$ commute, for each μ we may eliminate the scalar quantities $\underset{\mu,r}{X^{*}}$, not all of which are zero, from the first set, and obtain the eleiminant $|\underset{*}{A}(\underset{*}{\alpha}^{\mu})| = 0$, where $\underset{*}{A}(\underset{*}{\alpha})$ denotes the composite matrix with components $\underset{*}{A}^{rs}(\underset{*}{\alpha}) = \underset{*}{A}^{rs} - \underset{*}{\alpha}\,\delta_{rs}$. Similarly there may be obtained the eliminant $|\underset{}{A}^{*}(\underset{}{\alpha}^{*})| = 0$ from the second set, where $\underset{}{A}^{*}(\underset{}{\alpha}^{*})$ denotes the composite matrix with components $\underset{ij}{A^{*}}(\underset{}{\alpha}^{*}) = \underset{ij}{A^{*}} - \underset{}{\alpha}^{*}\,\delta_{ij}$. Let $\underset{}{\alpha}^{*}$, $\underset{*}{\alpha}$ be indeterminates in the algebras generated by the components of $\underset{}{A}^{*}$ and $\underset{*}{A}$ respectively, and define polynomials $|\underset{}{A}^{*}(\underset{}{\alpha}^{*})|$ and $|\underset{*}{A}(\underset{*}{\alpha})|$, mutually dual, as the composite characteristic polynomials of $\underset{}{A}^{*}$ and $\underset{*}{A}$, respectively. We have proved that the matrices $\underset{}{\alpha}^{*}_{\nu}$ and $\underset{*}{\alpha}^{\mu}$ satisfy the equations

$$|\underset{}{A}^{*}(\underset{}{\alpha}^{*})| = 0 \quad \text{and} \quad |\underset{*}{A}(\underset{*}{\alpha})| = 0$$

respectively, and this we may express by the following:

THEOREM IX. The characteristic components of a doubly commutative dual pair of composite matrices give null values for their respective composite characteristic polynomials.

Let $\underset{\sim}{a}$ be any matrix, with elements a_{ij}, and let $C_{ij}^{rs} = a_{rs}\delta_{ij} - a_{ij}\delta_{rs}$. Then C_{ij}^{rs} defines the ground elements of a doubly-commutative dual pair of composite matrices $\underset{\sim}{\mathfrak{C}}^{*}, \underset{\sim}{\mathfrak{C}}_{*}$ each of order n and degree n, with components $\underset{\sim}{C}_{ij}^{*} = \underset{\sim}{a}\delta_{ij} - a_{ij}\underset{\sim}{I}$ and $\underset{\sim}{C}_{*}^{rs} = a_{rs}\underset{\sim}{I} - \underset{\sim}{a}\delta_{rs}$ respectively, such that $\underset{\sim}{\mathfrak{C}}^{*} = -\underset{\sim}{\mathfrak{C}}_{*}$. The Cayley-Hamilton theorem shows that $|\underset{\sim}{\mathfrak{C}}^{*}| = 0$. It may be noted as a consequence that the ground matrix $\underset{\sim}{c}^{*}$ of $\underset{\sim}{\mathfrak{C}}^{*}$ is singular. For, by theorem IV, we now have $|\underset{\sim}{c}^{*}| = |\underset{\sim}{C}^{*}| = |\underset{\sim}{0}| = 0$

The Cayley-Hamilton theorem supplies a system of algebraic identities applied to the elements of a matrix, and hence it remains valid when the components of any commutative composite matrix are substituted.

Now let $\underset{\sim}{A}^{*}$, $\underset{\sim}{A}_{*}$ be any dual commutative composite matrices. There exists a matrix $\underset{\sim}{T}$ which transforms all the components of $\underset{\sim}{A}^{*}$ into direct sums $\underset{\sim}{T}^{-1}\underset{\sim}{A}_{ij}^{*}\underset{\sim}{T} = \underset{\lambda}{\Phi}\underset{\sim}{A}_{ij}^{*\lambda}$ of triangular matrices $\underset{\sim}{A}_{ij}^{*\lambda}$ whose characteristic values, which appear on their diagonals, are the same, and equal to $\alpha_{ij}^{*\lambda}$ respectively, where the sets of elements $\alpha_{ij}^{*\lambda}$ enumerate the sets of elements of the distinct matrices α_{*}^{μ}. Let $\underset{\sim}{T}^{*}$ be the composite matrix with components $\underset{\sim}{T}_{ij}^{*} = \underset{\sim}{T}\delta_{ij}$. ¶

Now $\underset{\sim}{B}^* = \underset{\sim}{T}^{*-1}\,\underset{\sim}{A}^*\,\underset{\sim}{T}^*$ has components $\underset{\sim}{B}^*_{ij} = \sum_\lambda \underset{\sim}{\epsilon}^{*\lambda}_{ij}$, and the dual $\underset{\sim}{B}_*$ of $\underset{\sim}{B}^*$ is a composite matrix which is similar, taking component to element, with each of the components of B^*, that is to say in respect to the distribution of null elements; more explicitly, $\underset{\sim}{B}_*$ is triangular, and the matrices which form its diagonal components occur with suitable multiplicities to give the characteristic matrices $\underset{\sim}{\alpha}^\mu_*$ in some order. It follows that

$$|B_*(\lambda)| = \prod_{\mu=1}^{m} (\lambda \underset{\sim}{I}_* - \underset{\sim}{\alpha}^\mu_*).$$

But it follows from theorem I that the dual of the reciprocal composite matrix $\underset{\sim}{T}^{*-1}$ of $\underset{\sim}{T}^*$ is the reciprocal of the dual $\underset{\sim}{T}_*$ of $\underset{\sim}{T}^*$, and then that $\underset{\sim}{B}_* = (T_*)^{-1}\,\underset{\sim}{A}_*\,\underset{\sim}{T}_*$; so that $|\underset{\sim}{A}_*(\lambda)| = |\underset{\sim}{B}_*(\lambda)|$. Hence

$$|\underset{\sim}{A}_*(\alpha_*)| = \prod_{\mu=1}^{m} (\alpha_* - \underset{\sim}{\alpha}^\mu_*) ;$$

and similarly for $\underset{\sim}{A}^*$. Thus we have

THEOREM X. *If $\underset{\sim}{A}^*$, $\underset{\sim}{A}_*$ are dual commutative composite matrices, with characteristic components $\underset{\sim}{\alpha}^*_\nu$, $\underset{\sim}{\alpha}^\mu_*$, then*

$$|\underset{\sim}{A}^*(\alpha^*)| = \prod_{\nu=1}^{n} (\alpha^* - \underset{\sim}{\alpha}^*_\nu), \quad |\underset{\sim}{A}_*(\alpha_*)| = \prod_{\mu=1}^{m} (\alpha_* - \underset{\sim}{\alpha}^\mu_*)$$

Now, by an application of the Cayley-Hamilton theorem, which gives $C^*(\underset{\sim}{A}^*) = O$ if $C(\alpha^*) = |\underset{\sim}{A}^*(\alpha^*)|$, we deduce:

THEOREM XI. *If $\underset{\sim}{A}^*$, $\underset{\sim}{A}_*$ are dual commutative composite matrices, and $\underset{\sim}{\alpha}^*_\nu$, $\underset{\sim}{\alpha}^\mu_*$ are their characteristic components then*

$$\prod_{\nu=1}^{n} (\underset{\sim}{A}^* - \underset{\sim}{\alpha}^*_\nu \underset{\sim}{I}^*) = \underset{\sim}{O}^* \text{ and } \prod_{\mu=1}^{m} (\underset{\sim}{A}_* - \underset{\sim}{\alpha}^\mu_* \underset{\sim}{I}_*) = \underset{\sim}{O}_* .$$

L Let K be an algebra of matrices of order n, say a matrix algebra of degree n; and let L be another matrix algebra, of degree m. Then we may form the two composite product matrix algebras $C^* = K^* \times L$ and $C_* = K \times^* L$, mutually dual to each other, composed of mutually dual composite matrices formed from ground elements $X_{ij}^{rs} = a_{rs}^* a_*^{ij} + b_{rs}^* b_*^{ij} + \ldots$, with components $\underset{\sim}{X}_{ij}^* = \underset{\sim}{a}^* a_*^{ij} + \underset{\sim}{b}^* b_*^{ij} + \ldots$, $\underset{\sim}{X}_*^{rs} = a_{rs}^* \underset{\sim}{a}_* + b_{rs}^* \underset{\sim}{b}_* + \ldots$ respectively, where $\underset{\sim}{a}^*, \underset{\sim}{b}^*, \ldots$ are any matrices of K, and $\underset{\sim}{a}_*, \underset{\sim}{b}_*, \ldots$ are any of L. The ground matrices form a matrix algebra C, of degree mn, generated by matrices which are direct products of a matrix from K and a matrix from L. The algebras C, C^* and C_* all represent the same abstract algebra, that is to say they are isomorphic. When both K and L are commutative, then the product algebra C is commutative, and each of the composite matrices which belong to the algebras C^* and C_* is doubly-commutative.

For when K and L are of the same degree, let $\underset{\sim}{X}$, $\underset{\sim}{X}^*$ and $\underset{\sim}{X}_*$ be any related members of the related algebras C, C^* and C_*. Then there is a fourth matrix associated with these, called their key matrix, formed by taking the expression $\underset{\sim}{k} = \underset{\sim}{a}^* \times \underset{\sim}{a}_* + \underset{\sim}{b}^* \times \underset{\sim}{b}_* + \ldots$ and then the matrix $\underset{\sim}{k} = \underset{\sim}{a}^* \underset{\sim}{a}_* + \underset{\sim}{b}^* \underset{\sim}{b}_* + \ldots$ When K is commutative, $\underset{\sim}{X}^*$ is a commutative composite matrix, the determinant $|\underset{\sim}{X}^*|$ is defined, and it will appear in the next theorem that $|\underset{\sim}{X}^*| = 0$ whenever $\underset{\sim}{k} = 0$.

We first note that $\underline{x}=\sum_{ij} x_{ij}\underline{e}_{ij}$, for any matrix $\underline{x}$ with elements x_{ij}, where $\underline{e}_{ij}$ is the matrix with r,sth element $\delta_{ri}\delta_{sj}$. If, for anyh r we have $\underline{e}_{rs}\underline{x}=\underline{0}$ for each s, then, since $\underline{e}_{rs}\underline{x}=\sum_{ij}\delta_{si}\underline{e}_{rj}x_{ij}=\sum_{j}\underline{e}_{rj}x_{sj}$, we have $\underline{x}_{sj}=0$ for all j and all s; that is, $\underline{x}=\underline{0}$.

THEOREM XII. *If* $\underline{x}=u\underline{a}+v\underline{b}+\ldots$ *where* $\underline{a},\underline{b},\ldots$ *commute, and if* $\underline{X}_{ij}=u_{ij}\underline{a}+v_{ij}\underline{b}+\ldots$, *then* $|\underline{X}|=\underline{0}$ *if* $\underline{x}=\underline{0}$.

Since $\underline{a}$, $\underline{b}$, ... all commute, so do the matrices $\underline{X}_{ij}$, and the composite matrix $\underline{X}$ with these components has an adjugate composite matrix $\underline{X}^{*}$, with components $\underline{X}^{*}_{ij}$ which are such that $\sum_{s}\underline{X}_{sj}\underline{X}^{*}_{is}=|\underline{X}|\delta_{si}$. Now $\underline{x}=\sum_{ij}\underline{e}_{ij}\underline{X}_{ij}$, and $\underline{e}_{rt}\underline{x}=\sum_{j}\underline{e}_{rj}\underline{X}_{sj}$ Thus, if $\underline{x}=\underline{0}$, for any t, we then have

$$\underline{0}=\sum_{s}\Big(\sum_{j}\underline{e}_{rj}\underline{X}_{sj}\Big)\underline{X}^{*}_{is}=\sum_{j}\underline{e}_{rj}\delta_{tj}|\underline{X}|=\underline{e}_{rt}|\underline{X}|\ ,$$

for all r, which implies $|\underline{X}|=\underline{0}$.

Notes

(96) 1. Williamson(1).

(96) 2. Frobenius(2).

(96) 3. Afriat(3).

(99) 4. We define $\alpha,\beta,\ldots$ as latently associated characteristic values of commutative matrices $\underline{a},\underline{b},\ldots$ if there exists a vector $\underline{X}$ which is a latent vector of $\underline{a}$ corresponding to α, of $\underline{b}$ corresponding to $\beta,\ldots\ldots$, that is to say such that $\underline{a}\underline{X}=\underline{X}\alpha$, $\underline{b}\underline{X}=\underline{X}\beta,\ldots\ldots$

(99) 5. The characteristic sets of a commutative matrix set are the latently associated sets of the characteristic values, taken either with certain repetitions, or with certain multiplicities attached.

(101) 6. See Aitken(4), p.24.

(103) 7. See Drazin, Dungey and Gruenberg(5).

(105) 8. Weyl(6), p.4.

(107) 9. Here we use the result that the characteristic values of any rational function $f(\underline{s})$, of a set $\underline{s}$ of commutative matrices, are the values of $\phi=f(\sigma)$ for when σ ranges over the characteristic sets of $\underline{s}$. See Frobenius(7), and Drazin, Dungey and Gruenberg, loc.cit.

(108) 10. For the characteristic sets of any set $F=(f_\iota(\underline{s}))_{\iota\in I}$ of rational functions of some commutative matrix set $\underline{s}$ are the sets $\overline{\varphi_\sigma}=(\varphi_\iota(\sigma))_{\iota\in I}$, where σ ranges over the characteristic sets of $\underline{s}$.

REFERENCES

I. J. Williamson,'The latent roots of a matrix of special type', Bull.Amer.Math.Soc.,Vol.37(I93I),585-90.

2. G. Frobenius,'Uber lineare Substitutionen und bilineare Formen',J.reine angew.Math.,Vol.84(I878),I-63.

3. S.N. Afriat,'The quadratic form positive definite on a linear manifold', Pro.Camb.Phil.Soc.,Vol.47,Pt.I (I95I),I-6.

4. A.C. Aitken, Determinantsaand Matrices(Edinburgh I942)

5. M.P. Drazin, J.W. Dungey and K.W. Gruenberg,'Some theorems on commutative matrices', J.Lond.Math.Soc., Vol.26(I95I),22I-8.

6. H. Weyl, The Classical Groups(Princeton I939),60I-I4.

7. G. Frobenius, S.-B. preuß.Akad.Wiss.,I(I896),60I-I4. Uber vertauschbare Matrizen'.

FUNCTIONS OF A COMPLEX MATRIX VARIABLE

contents

1 Introduction

2 The resolvent

3 Cauchy integrals

4' Idempotents

4 The partial resolvents

5 Rational functions

6 The resolvent coefficients

7' The principal and axial spaces

7 The principal polynomials

8 Analytic functions

9 Regularity and derivatives

10 Infinite powers and geometric sums

11 Latent association of characteristic values

12 Taylor expansions

13 Power series

I. Introduction.

Let Δ_o be any open connected domain, in the complex plane or on a Riemann surface, and $\underline{a}$ any complex matrix whose characteristic values lie in Δ_o. Let Δ_a denote the domain Δ_o with the characteristic values of $\underline{a}$ removed.

The resolvent $\underline{R}(\zeta)=\frac{\underline{I}}{\zeta\underline{I}-\underline{a}}$ of $\underline{a}$ is defined in Δ_a, and for every ζ which is not a characteristic value of $\underline{a}$. It is bounded and continuous in any closed domain Δ in Δ_a, and satisfies the equation $\underline{R}(\zeta)-\underline{R}(\mu)=-(\lambda-\mu)\underline{R}(\lambda)\underline{R}(\mu)$. Its elements are functions of ζ which are regular in Δ_a and whose derivatives are the elements of the matrix $\underline{R}'(\zeta)=-\underline{R}(\zeta)^2$

If Γ_α denotes any simple contour in Δ_a which encloses just the characteristic value α of $\underline{a}$ in its interior, the matrices

$$\underline{R}_\alpha = \frac{\underline{I}}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\zeta}{\zeta\underline{I}-\underline{a}}$$

are uniquely defined, and such that $\underline{R}_\alpha^2=\underline{R}_\alpha\neq\underline{0}$, $\underline{R}_\alpha\underline{R}_\lambda=\underline{0}(\lambda\neq\alpha)$, $\sum_\alpha\underline{R}_\alpha=\underline{I}$ and $(\underline{a}-\alpha\underline{I})\underline{R}_\alpha$ is nullpotent. They form an independent idempotent decomposition of the unit matrix, and are called the partial units relative to $\underline{a}$. They are defined also as the unique matrices $\underline{R}_\alpha$, indexed by the characteristic values α of $\underline{a}$, suchthat $\underline{a}\underline{R}_\alpha=\underline{R}_\alpha\underline{a}$, $\underline{R}_\alpha^2=\underline{R}_\alpha\neq\underline{0}$, $\sum_\alpha\underline{R}_\alpha=\underline{I}$ and $(\underline{a}-\alpha\underline{I})\underline{R}_\alpha$ is nullpotent; and also as certain polynomials in $\underline{a}$, called the principal polynomials in a. If m_o denotes th

in $\underline{a}$. If m_α denotes the multiplicity of any α as a characteristic value of $\underline{a}$, then they are of rank m_α respectively.

Contours which can be continuously deformed into each other in Δ_a are said to be of the same class in Δ_a. If $f(\zeta)$ is an analytic function, and Δ_o a Riemann surface on which $f(\zeta)$ is represented as single valued and regualr, and if Γ is any contour in Δ_a, then the value of the Cauchy integral

$$\oint_\Gamma \frac{f(\zeta)d\zeta}{\zeta I - \underline{a}}$$

depends only on the class of Γ in Δ_a. If Γ_a denotes any contour in Δ_a which encloses each characteristic value of $\underline{a}$ just once in its interior, then the value of

$$f(\underline{a}) = \frac{I}{2\pi i} \oint_{\Gamma_a} \frac{F(\zeta)d\zeta}{\zeta I - \underline{a}}$$

is specified when the sheet of Δ_o on which each characteristic value of $\underline{a}$ is enclosed in Γ_a is specified, or equivalently when a determination of $f(\zeta)$ at each characteristic value is assigned; and it defines a general determination of a function which extends the analytic function $f(\zeta)$, and which is represented as single valued on the Riemann matrix domain D_o consisting of the matrices $\underline{a}$ whose characteristic values are described on the Riemann surface Δ_o. If Γ_α denotes any simple contour on which $f(\zeta)$ is regular, and which encloses in its interior just the characteristic value α of a, and no singularities of $f(\zeta)$, then any determination of $f(\underline{a})$ is given by

$$f(\underline{a}) = \frac{I}{2\pi i} \sum_{\alpha} \oint_{\Gamma_\alpha} \frac{f_\alpha(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} \quad ,$$

where, for each α, the function $f_\alpha(\zeta)$ represents any determination of $f(\zeta)$, single valued and regular on and in the interior of Γ_α.

Let $f(\zeta)$ be single value and regular in Δ_0, and let Γ be any contour in Δ_a. Then there is the formula

$$f(\underline{a}) \oint_{\Gamma} \frac{d\zeta}{\zeta\underline{I}-\underline{a}} = \oint_{\Gamma} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} \quad ,$$

which reduces to the definition of $f(\underline{a})$, since

$$\oint_{\Gamma_a} \frac{d\zeta}{\zeta\underline{I}-\underline{a}} = 2\pi i\underline{I},$$

where Γ_a is any contour which contains all the characteristic values of $\underline{a}$ simply in its interior.

It is shown further that if $r(\zeta)$ is any rational function, with poles ouside Γ_a, then

$$r(\underline{a}) = \frac{I}{2\pi i} \oint_{\Gamma} \frac{r(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} \quad ,$$

and this shows that our definition of the extension $f(\underline{a})$ of any analytic function $f(\zeta)$ comprehends the natural extension of any rational function $r(\zeta)$.

For any polynomial $p(\zeta)$, a necessary and sufficient condition that $p(\underline{a})=\underline{0}$ is that $\frac{p(\zeta\underline{I})}{\zeta\underline{I}-\underline{a}}$ should be a polynomial in ζ. Thus if $C(\zeta)=|\zeta\underline{I}-\underline{a}|$, then $|\frac{\zeta\underline{I}-\underline{a}}{\zeta\underline{I}-\underline{a}}| = (\zeta\underline{I}-\underline{a})^*$, the adjugate of $\zeta\underline{I}-\underline{a}$, is a polynomial, and hence $C(\underline{a})=\underline{0}$. It follows that the resolvent of $\underline{a}$ can be expressed as an irreducible proper rational function $\frac{I}{\zeta\underline{I}-\underline{a}} = \frac{P(\zeta\underline{I},\underline{a})}{\phi(\zeta)}$, where $P(\zeta\underline{I},\underline{a})$ is a polynomial in $\underline{a}$ and ζ, and $\phi(\zeta)$ is the minimum polynomial in $\underline{a}$. Moreover, $\phi(\zeta) = \prod_{\alpha} (\zeta-\alpha)^{i_\alpha}$,

where $C(\zeta) = \prod_{\alpha} (\zeta - \alpha)^{m_\alpha}$ and $0 < i_\alpha \leq m_\alpha$. The exponents m_α and i_α define the multiplicity and the index, respectively, of any characteristic value α. The rational expression for the resolvent decomposes into its partial fractions to give

$$\underline{R}(\zeta) = \sum_{\alpha} \sum_{t<i_\alpha} \frac{\underline{R}_{\alpha,t}}{(\zeta-\alpha)^{t+1}} ,$$

where the coefficient matrices $\underline{R}_{\alpha,t}$ are all polynomials in $\underline{a}$, and $\underline{R}_{\alpha,i_\alpha} \neq \underline{0}$. It is deduced that

$$\underline{R}_{\alpha,t} = (\underline{a}-\alpha\underline{I})^r \underline{R}_\alpha, \quad \text{and } \underline{R}_{\alpha,t} = 0 \text{ if and only if } r \geq i_\alpha$$

so that

$$\frac{\underline{I}}{\zeta\underline{I}-\underline{a}} = \sum_{\alpha} \underline{R}_\alpha \sum_{t<i_\alpha} \frac{(\underline{a}-\alpha\underline{I})^r}{(\zeta-\alpha)^{t+1}} .$$

Also it is shown that the matrices $(\underline{a}-\alpha\underline{I})^r \underline{R}_\alpha$ $(r<i_\alpha, \alpha)$ form a linearly independent set.

The determinations of $f(\underline{a})$ now appear according to the formula

$$f(\underline{a}) = \sum_{\alpha} \underline{R}_\alpha \sum_{t<i_\alpha} \frac{I}{r!} f_\alpha^{(t)}(\alpha)(\underline{a}-\alpha\underline{I})^r ;$$

and a corollary is that $f(\underline{a})=\underline{0}$ if and only if $f_\alpha^{(t)}(\alpha)=0$ $(r<i_\alpha, \alpha)$.

With the regularity of $f(\underline{z})$ at $\underline{a}$ defined in the sense that the quotient $\frac{f(\underline{b})-f(\underline{a})}{\underline{b}-\underline{a}}$ has a uniquely defined limit when $\underline{b}$, commuting with $\underline{a}$, tends to $\underline{a}$ in an arbitrary manner such that $\frac{|\underline{b}-\underline{a}|^*}{|\underline{b}-\underline{a}|_*}$ remains bounded, and with this limit, when it exists, defining the derivative of $f(\underline{z})$ at $\underline{a}$, it is shown that $f(\underline{z})$ is regular in $\underline{D}_0$, and has derivative $f'(\underline{z})$, where $f'(\underline{z})$ denotes the function which

extends the definition of the derivative $f'(\zeta)$ of $f(\zeta)$ to the domain $\underline{I}_0$ of matrices whose characteristic values all lie in the domain of regularity Δ_0 of $f(\zeta)$.

If $\underline{a},\underline{b}$ is any commutative matrix pair, and $\underline{R}_\alpha$, $\underline{S}_\beta$ are the partial units relative to $\underline{a},\underline{b}$ respectively, then α, β are called <u>latently associated</u> characteristic values of $\underline{a}$ and $\underline{b}$ if $\underline{R}_\alpha \underline{S}_\beta \neq 0$. It is shown that

$$f(\underline{b}) = \sum_{r<\infty} \frac{I}{r!} f^{(r)}(\underline{a})(\underline{b}-\underline{a})^r$$

where $\underline{b}$ is any matrix which commutes with $\underline{a}$ and is such that each characteristic value of $\underline{a}$ is nearer to every latently associated characteristic value of $\underline{b}$ than the nearest singularity.

Finally power series are studied. If $\mathcal{P}(\zeta)$ denotes a power series, with derived series $\mathcal{P}^{(1)}(\zeta)$, then the power series $\mathcal{P}(\underline{a})$ is convergent if and only if the series $\mathcal{P}^{(r)}(\alpha)$ $(r<i_\alpha, \alpha)$ are all convergent. Also there is the following theorem: If $\mathcal{P}(\zeta)$ denotes a power series, Δ the interior of its circle of convergence, $f(\zeta)$ its sum function, defined and regular in Δ, $\mathcal{P}^{(1)}(\zeta)$ its derived series, whose sums in Δ are the derivatives $f^{(r)}(\zeta)$ of $f(\zeta)$, and $\underline{D}$ the domain of matrices whose characteristic values all lie in Δ, then the power series $\mathcal{P}(\underline{a})$ and its derived series $\mathcal{P}^{(1)}(\underline{a})$ are convergent for all $\underline{a}$ in $\underline{D}$, and their sums, representing a regular function in $\underline{D}$ and its derivatives, are given by

$$f(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a} \frac{f(\underline{\zeta})d\underline{\zeta}}{\underline{\zeta I}-\underline{a}}$$

and

$$f^{(r)}(\underline{a}) = \frac{1}{2\pi i}\oint_{\Gamma_a} \frac{f^{(r)}(\underline{\zeta})d\underline{\zeta}}{\underline{\zeta I}-\underline{a}} = \frac{r!}{2\pi i}\oint_{\Gamma_a} \frac{f(\underline{\zeta})d\underline{\zeta}}{(\underline{\zeta I}-\underline{a})^{r+1}},$$

where Γ_a is any contour lying in Δ and containing all the characteristic values of $\underline{a}$ in its interior.

2. The resolvent.

Let $\underline{a}$ be any matrix, with complex elements, and characteristic values in some open connected domain Δ_o; and let Δ_a be the domain Δ_o with the characteristic values of $\underline{a}$ removed.

The lower distance of any domain Δ from $\underline{a}$ is defined as $|\Delta-\underline{a}|_* = \inf_{\Delta} |\zeta\underline{I}-\underline{a}|_*$. [1]

THEOREM I. If Δ is closed in Δ_a, then $|\Delta-\underline{a}|_*>0$.

For then $|\zeta\underline{I}-\underline{a}|_*$, which is a continuous function of ζ in Δ, attains its lower bound at a point ζ_o in Δ and hence in Δ_a. Further, any point in Δ_a is of non-zero distance from each characteristic value α of $\underline{a}$, so that $\rho_o = \inf_{\alpha} |\zeta_o-\alpha|>0$. Since $0<\rho_o^n \leq |\zeta_o\underline{I}-\underline{a}|_* |\zeta_o\underline{I}-\underline{a}|^{*n-1}$, and since $|\zeta_o\underline{I}-\underline{a}|^*<\infty$, with Δ closed, and hence bounded, it follows that $|\zeta_o\underline{I}-\underline{a}|_*>0$.

The resolvent of $\underline{a}$ is given by $\underline{R}(\zeta) = \frac{\underline{I}}{\zeta\underline{I}-\underline{a}}$. [2] Its elements are functions of ζ which are defined for every ζ which is not a characteristic value of $\underline{a}$, and hence for all ζ in Δ_a.

THEOREM II. $\underline{R}(\zeta)$ is bounded in every closed domain Δ in Δ_a.

The bound of $\underline{R}(\zeta)$ in Δ is defined by $|\underline{R}(\Delta)|^* = \sup_{\Delta} |\underline{R}(\zeta)|^*$

By theorem 2.I,

$$|R(\Delta)|^* = \sup_{\Delta} \left|\frac{\underline{I}}{\zeta\underline{I}-\underline{a}}\right|^* = \sup_{\Delta} \left(\frac{1}{|\zeta\underline{I}-\underline{a}|_*}\right) = \frac{\underline{I}}{\inf_{\Delta} |\zeta\underline{I}-\underline{a}|_*} = \frac{1}{|\Delta-\underline{a}|_*} < \infty$$

[3]
THEOREM III. $\underline{R}(\lambda)-\underline{R}(\mu)=-(\lambda-\mu)\underline{R}(\lambda)R(\mu)$ for λ,μ in Δ_a.

Thus, for all λ,μ in Δ_a,

$$\underline{R}(\lambda)-\underline{R}(\mu)=\frac{\underline{I}}{\lambda\underline{I}-\underline{a}}-\frac{\underline{I}}{\mu\underline{I}-\underline{a}}=-\frac{\lambda-\mu}{(\lambda\underline{I}-\underline{a})(\mu\underline{I}-\underline{a})}=-(\lambda-\mu)R(\lambda)R(\mu)$$

THEOREM IV. $\underline{R}(\zeta)$ is continuous in any closed domain Δ in Δ_a

By theorem 2.II,

$$|\underline{R}(\zeta)|^*\leq|R(\Delta)|^*<\infty, \text{ for } \zeta \text{ in } \Delta.$$

Then, by theorem 2.III,

$$|\underline{R}(\lambda)-\underline{R}(\mu)|^*\leq|\lambda-\mu||\underline{R}(\lambda)|^*|\underline{R}(\mu)|^*$$
$$\leq|\lambda-\mu||\underline{R}(\Delta')|^{*2},$$

for all λ,μ in Δ'. Accordingly, for any $\epsilon>0$, with $\delta=\epsilon/|R(\Delta')|^{*2}$, we have

$$|\underline{R}(\lambda)-\underline{R}(\mu)|^*<\epsilon \quad \text{if } |\lambda-\mu|<\delta,$$

for all λ,μ in Δ'; and this establishes the continuity of $\underline{R}(\zeta)$ in Δ'.

THEOREM V. The elements of $\underline{R}(\zeta)$ are regular functions of ζ in Δ_a whose derivatives are the elements of the matrix $\underline{R}'(\zeta)=-\underline{R}(\zeta)^2$.

If ζ is in Δ_a there exists a closed neighbourhood Δ' of ζ in Δ_a. By theorem 2.IV, $\underline{R}(\omega)$ is a continuous function of ω in Δ', so that $R(\omega)\to R(\zeta)$, $(\omega\to\zeta)$. Hence, by theorem 2.III,

$$\frac{\underline{R}(\zeta)-\underline{R}(\omega)}{\zeta-\omega}=-\underline{R}(\zeta)\underline{R}(\omega)\to-\underline{R}(\zeta)^2, \quad (\omega\to\zeta)$$

which is the desired result.

3. Contour integrals.

Any contours Γ, Γ' in $\Delta_{\underline{a}}$ are said to be equivalent, or of the same class, or to have the relation $\Gamma \backsim \Gamma'$, if they can be continuously deformed into each other in $\Delta_{\underline{a}}$; and a necessary and sufficient condition for this is that they should enclose the same characteristic values of $\underline{a}$ in their interiors. If Γ', Γ'' are any contours in $\Delta_{\underline{a}}$, write $\Gamma \backsim \Gamma'.\Gamma''$ for any contour Γ which contains in its interior the characteristic values of $\underline{a}$ which are contained in the interiors of both Γ' and Γ''. Write $\Gamma \backsim 0, \alpha$ or $\underline{a}$, respectively, if Γ encloses none, just α, or all the characteristic values of $\underline{a}$ in its interior; also write $\alpha < \Gamma$ or $\alpha > \Gamma$ according as an element α lies in the interior or exterior of a contour Γ in $\Delta_{\underline{a}}$, and define $\delta_{\alpha\Gamma} = 0, 1$ according as $\alpha > \Gamma$ or $\alpha < \Gamma$.

THEOREM I. *If $f(\zeta)$ is an analytic function regular in Δ, and Γ_0 is any contour such that $\Gamma_0 \backsim 0$ in $\Delta_{\underline{a}}$, then*

$$\oint_{\Gamma_0} \frac{f(\zeta)d\zeta}{\zeta\underline{1}-\underline{a}} = \underline{0}.$$

For then, by theorem 2.V, each element of $\dfrac{f(\zeta)}{\zeta\underline{1}-\underline{a}}$ is an analytic function of ζ which is regular throughout Γ; and hence the considered integral is null by Cauchy's theorem applied to each of the element integrals.

THEOREM II. *If $f(\zeta)$ is an analytic function regular in Δ and Γ, Γ' are any contours such that $\Gamma \backsim \Gamma'$ in $\Delta_{\underline{a}}$ then*

$$\oint_{\Gamma} \frac{f(\zeta)d\zeta}{\zeta\underline{1}-\underline{a}} = \oint_{\Gamma'} \frac{f(\zeta)d\zeta}{\zeta\underline{1}-\underline{a}}$$

Since Γ may be continuously deformed into Γ' in Δ_ε, let γ be the trajectory of one of the points of Γ in such a deformation. Then a contour Γ_0 can be formed which describes Γ positively, Γ' negatively, and γ in both directions, and which contains no characteristic values of $\underline{a}$ in its interior, so that $\Gamma_0 \sim 0$. The contribution to the integral round Γ_0 by the part of the path which falls on γ is null. Hence, by theorem 3.I,

$$\underline{0} = \oint_{\Gamma_0} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} = \oint_{\Gamma} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} - \oint_{\Gamma'} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} .$$

THEOREM III. *If $f(\zeta)$ is regular in Δ, Γ is any contour in Δ_ε, and for each $\alpha < \Gamma$ Γ_α is any contour in Δ_ε such that $\Gamma_\alpha \sim \alpha$, then*

$$\oint_{\Gamma} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} = \sum_{\alpha < \Gamma} \oint_{\Gamma_\alpha} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

First take contours $\Gamma_\alpha \sim \alpha$ which lie in the interior of Γ and are exterior to each other, and join a point on each Γ_α to a point on Γ by non-intersecting arcs γ_α in the interior of Γ and in the exteriors of the Γ_α. Then a contour Γ_0 is formed which describes Γ positively, each Γ_α negatively, and each γ_α in two directions. The interior of Γ_0 consists of the interior of Γ with the contours γ_0 and their interiors and the arcs γ_α removed. Thus the interior of Γ_0 contains no characteristic values of $\underline{a}$, so that $\Gamma_0 \sim 0$; and hence, by theorem 3.I, the integral round Γ_0 is null.. But the contribution to this integral by the parts of the path which fall on the γ_α is zero.

Hence we have

$$Q = \oint_{\Gamma_0} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} = \oint_{\Gamma} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} - \sum_{\alpha<\Gamma} \oint_{\Gamma_\alpha} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

But, by theorem 3.II, the integrals round any Γ_α depend only on the class of Γ_α in Δ_a, so that this result is true when Γ_α is any contour in Δ_a such that $\Gamma_\alpha \frown \alpha$.

THEOREM IV. *If Γ is any contour in Δ_a, $f(\zeta)$ is regular on and in the interior of Γ, and ζ is any point of Δ_a which does not lie on Γ, then*

$$\frac{I}{\zeta\underline{I}-\underline{a}} \cdot \frac{I}{2\pi i} \oint_{\Gamma} \frac{f(\omega)d\omega}{\omega\underline{I}-\underline{a}} = \frac{I}{2\pi i} \oint_{\Gamma} \frac{f(\omega)d\omega}{(\zeta-\omega)(\omega\underline{I}-\underline{a})} + \frac{f(\zeta)\delta_{\zeta\Gamma}}{\zeta\underline{I}-\underline{a}}$$

For, with ζ and ω in Δ_a,

$$\frac{I}{(\zeta-\omega)(\omega\underline{I}-\underline{a})} = \frac{I}{\zeta\underline{I}-\underline{a}} \left(\frac{I}{\zeta-\omega} + \frac{I}{\omega\underline{I}-\underline{a}} \right).$$

Hence

$$\oint_{\Gamma} \frac{f(\omega)d\omega}{(\zeta-\omega)(\omega\underline{I}-\underline{a})} = \frac{I}{\zeta\underline{I}-\underline{a}} \left(\oint_{\Gamma} \frac{f(\omega)d\omega}{\zeta-\omega} + \oint_{\Gamma} \frac{f(\omega)d\omega}{\omega\underline{I}-\underline{a}} \right)$$

$$= \frac{I}{\zeta\underline{I}-\underline{a}} \left(-2\pi i f(\zeta)\delta_{\zeta\Gamma} + \oint_{\Gamma} \frac{f(\omega)d\omega}{\omega\underline{I}-\underline{a}} \right),$$

since $\frac{I}{2\pi i} \oint_{\Gamma} \frac{f(\omega)d\omega}{\omega-\zeta} = f(\zeta)\delta_{\zeta\Gamma}$.

COROLLARY. *If ζ lies in the exterior of Γ, then*

$$\frac{I}{(\zeta\underline{I}-\underline{a})^{+}} \oint_{\Gamma} \frac{f(\omega)d\omega}{\omega\underline{I}-\underline{a}} = \oint_{\Gamma} \frac{f(\omega)d\omega}{(\zeta-\omega)^{+}(\omega\underline{I}-\underline{a})}$$

Since $\delta_{\zeta\Gamma}=0$, with ζ in the exterior of Γ, this formula for $+=I$ is an immediate consequence of the present theorem. By taking its validity for r=s as hypothesis, and applying the case with r=I to the function

$\phi(\omega) = \frac{f(\omega)}{(\zeta-\omega)^r}$, which is regular on and in the interior of Γ, the formula for r=s+1 is deduced. It follows, by induction, that the formula is true for each r.

THEOREM V. *If $f(\zeta)$, $g(\zeta)$ are analytic functions regular in Δ, and Γ_f, Γ_g are any contours in Δ_a, and if Γ_{fg} denotes any contour such that $\Gamma_{fg} \backsim \Gamma_f . \Gamma_g$ in Δ_a, then*

$$\frac{I}{2\pi i}\oint_{\Gamma_f}\frac{f(\xi)d\xi}{\xi\underline{I}-\underline{a}} \cdot \frac{I}{2\pi i}\oint_{\Gamma_g}\frac{g(\eta)d\eta}{\eta\underline{I}-\underline{a}} = \frac{I}{2\pi i}\oint_{\Gamma_{fg}}\frac{f(\zeta)g(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

Take mutually disjoint contours Γ_α such that $\Gamma_\alpha \backsim \alpha$ for each characteristic value α. By theorem 3.III, each of the considered integrals, round any contour Γ, may be decomposed into a sum of integrals round various contours Γ_α, for the various α in the interior of Γ. Hence it will be sufficient to show that

$$\frac{I}{2\pi i}\oint_{\Gamma_\lambda}\frac{f(\xi)d\xi}{\xi\underline{I}-\underline{a}} \cdot \frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{g(\eta)d\eta}{\eta\underline{I}-\underline{a}} = \frac{\delta_{\alpha\lambda}}{2\pi i}\oint_{\Gamma_\alpha}\frac{f(\zeta)g(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} .$$

By theorem 3.IV,

$$\frac{I}{\xi\underline{I}-\underline{a}}\oint_{\Gamma_\alpha}\frac{g(\eta)d\eta}{\eta\underline{I}-\underline{a}} = \oint_{\Gamma_\alpha}\frac{g(\zeta)d\zeta}{(\xi-\zeta)(\zeta\underline{I}-\underline{a})}$$

for ξ in the exterior of Γ_α. For $\lambda \neq \alpha$, let $\Gamma_\lambda' = \Gamma_\lambda$; and let Γ_α' be any contour, disjoint from the Γ_λ $(\lambda \neq \alpha)$, lying in the exterior of Γ_α, such that $\Gamma_\alpha' \backsim \Gamma_\alpha$. Then it follows that

$$\oint_{\Gamma_\lambda}\frac{f(\xi)d\xi}{\xi\underline{I}-\underline{a}} \cdot \oint_{\Gamma_\alpha}\frac{g(\eta)d\eta}{\eta\underline{I}-\underline{a}} = \oint_{\Gamma_\lambda'} f(\xi)\left(\oint_{\Gamma_\alpha}\frac{g(\zeta)d\zeta}{(\xi-\zeta)(\zeta\underline{I}-\underline{a})}\right);$$

and the order of integration here may be changed, to give

$$\oint_{\Gamma_\alpha}\frac{g(\zeta)}{\zeta\underline{I}-\underline{a}}\left(\oint_{\Gamma_\lambda'}\frac{f(\xi)d\xi}{\xi-\zeta}\right)d\zeta = 2\pi i\delta_{\alpha\lambda}\oint_{\Gamma_\alpha}\frac{f(\zeta)g(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

4'. Idempotents.

A matrix $\underline{e}$ which is such that $\underline{e}^2=\underline{e}$ is called an idempotent. Thus if $\underline{e}$ is idempotent, then $\underline{e}^r=\underline{e}$ for any $r>0$. Two idempotents $\underline{f}$, $\underline{g}$ are called independent if $\underline{fg}=\underline{gf}=\underline{0}$.

THEOREM I. If $\underline{f},\underline{g},\ldots$ are independent idempotents, then

$$R_f \cap R_g = 0, \ldots\ldots, R_f \oplus R_g \oplus \ldots = R_{f+g+\ldots} .$$

where R_x denotes the range of any matrix x

This theorem will be shown for two idempotent, and the proof will then be completed by induction for an arbitrary finite number.

Firstly, we have $R_f \cap R_g = 0$, for $\underline{f}$ has the value I on its range R_f, since it is idempotent, and has the value 0 on R_g, since $\underline{f}$ and $\underline{g}$ are independent.

Let $\underline{Z}$ belong to R_{f+g}; then $\underline{Z}=(\underline{f}+\underline{g})\underline{Y}$ for some vector $\underline{Y}$, that is $\underline{Z}=\underline{fY}+\underline{gY}$; and thus $\underline{Z}$ belongs to $R_f \oplus R_g$. Hence $R_{f+g} \subset R_f \oplus R_g$. Also, if $\underline{Z}$ belongs to $R_f \oplus R_g$, then $\underline{Z}=\underline{fX}+\underline{gY}$ for some vectors $\underline{X}$ and $\underline{Y}$, so that

$$(\underline{f}+\underline{g})\underline{Z}=(\underline{f}+\underline{g})(\underline{fX}+\underline{gY})=\underline{fX}+\underline{gY}=\underline{Z},$$

since $\underline{f}$, $\underline{g}$ are independent idempotents; and thus $\underline{Z}$ belongs to the range of $\underline{f}+\underline{g}$. Hence $R_f \oplus R_g \subset R_{f+g}$; and, accordingly $R_f \oplus R_g = R_{f+g}$.

We now deduce the theorem for an arbitrary finite number of idempotents $\underline{f},\underline{g},\ldots$ on the hypothesis of the theorem for one less, and the proof by induction will then be complete.

Thus, we have

$$E_f \dotplus E_g \dotplus \ldots = E_f \dotplus (E_g \dotplus \ldots) = E_f \dotplus R_{g+\ldots} = R_{f+(g+\ldots)} = R_{f+g+\ldots}$$

by the associative law, the inductive hypothesis, the th theorem for two idempotents, and the associative law, respectively.

THEOREM II. *If $\sum_\lambda e_\lambda = I$ is an independent idempotent decomposition of the unit matrix, then there exists a regular matrix P such that*

$$P^{-1} e_\lambda P = \begin{pmatrix} 0 & & 0 \\ & I_{r_\lambda} & \\ 0 & & 0 \end{pmatrix}$$

where r_λ is the rank of e_λ.

Let P_λ be a base for the range $R_{e_\lambda} = e_\lambda \mathfrak{J}$. Then, since by theorem I we have $\dotplus_\lambda R_{e_\lambda} = R_{\sum_\lambda e_\lambda} = R_I = \mathfrak{J}$, it appears that the matrix $P = (P_1 \ldots P_\lambda \ldots P_l)$ is regular. Moreover e_λ has the value I on R_{e_λ}, and the value 0 on R_{e_μ} $(\mu \neq \lambda)$. Hence

$$e_\lambda P = (0 \ldots I_{r_\lambda} \ldots 0) = P \begin{pmatrix} 0 & & 0 \\ & I_{r_\lambda} & \\ 0 & & 0 \end{pmatrix},$$

which proves the theorem.

An idempotent f is said to *reduce* an idempotent e if there exists an idempotent g such that $e = f + g$.

THEOREM III. *If e, f are commuting idempotents, then f reduces e if and only if $ef = f$.*

If $\underline{f}$ reduces $\underline{e}$, then $\underline{e}-\underline{f}$ is idempotent, that is

$$\underline{e}-\underline{f} = (\underline{e}-\underline{f})^2 = \underline{e}-2\underline{e}\underline{f}+\underline{f} \quad ,$$

so that $\underline{e}\underline{f}=\underline{f}$. Conversely, if $\underline{e}\underline{f}=\underline{f}$, then

$$(\underline{e}-\underline{f})^2 = \underline{e}-2\underline{e}\underline{f}+\underline{f} = \underline{e}-\underline{f} \quad .$$

THEOREM IV. *If $\underline{f}$ and $\underline{g}$ are idempotents, then $\underline{f}+\underline{g}$ is idempotent if and only if $\underline{f}$ and $\underline{g}$ are independent.*

If $\underline{f}$, $\underline{g}$ are independent idempotents, then

$$(\underline{f}+\underline{g})^2 = \underline{f}^2+2\underline{f}\underline{g}+\underline{g}^2 = \underline{f}+\underline{g} \quad ,$$

so that $\underline{f}+\underline{g}$ is idempotent. Conversely, if $\underline{f}+\underline{g}$ is idempotent, then

$$\underline{f}+\underline{g} = (\underline{f}+\underline{g})^2 = \underline{e}+2\underline{f}\underline{g}+\underline{g} \quad ,$$

so that $\underline{g}\underline{f}=\underline{f}\underline{g}=0$.

4. The partial units.

THEOREM I. If $\Gamma_\alpha \sim \alpha$ in Δ_a, and $\underline{R}_\alpha = \frac{I}{2\pi i}\oint_{\Gamma_\alpha} \frac{d\zeta}{\zeta I - \underline{a}}$, then $\underline{R}_\alpha^2 = \underline{R}_\alpha \neq \underline{0}$, $\underline{R}_\alpha \underline{R}_\lambda = \underline{0}$ $(\lambda \neq \alpha)$, $\sum_\alpha \underline{R}_\alpha = \underline{I}$ and $(\underline{a} - \alpha \underline{I})\underline{R}_\alpha$ is nullpotent

By theorem 3.II, the set of matrices $\underline{R}_\alpha$ is uniquely defined. By theorem 3.V we have that $\underline{R}_\alpha^2 = \underline{R}_\alpha$ and $\underline{R}_\alpha \underline{R}_\lambda = \underline{0} (\lambda \neq \alpha)$ or that the $\underline{R}_\alpha$ are independent idempotent matrices. Let $\underline{R} = \sum_\alpha \underline{R}_\alpha$, so that $\underline{R}^2 = \underline{R}$. Then $\underline{R}$ will be shown to be the unit matrix when its characteristic values are shown to be all non-zero; also $(\underline{a} - \alpha \underline{I})\underline{R}_\alpha$ will be shown to be nullpotent when all its characteristic values are shown to be zero.

There exists a regular matrix $\underline{T}$ which transforms $\underline{a}$ into a triangular matrix $\underline{t} = \underline{T}^{-1} \underline{a} \underline{T}$, whose diagonal elements are the characteristic values α of $\underline{a}$, with repetitions equal to their multiplicities, m_α respectively, as characteristic values of $\underline{a}$. This matrix also transforms $\underline{R}_\alpha$ into the triangular matrix $\underline{S}_\alpha = \frac{I}{2\pi i}\oint_{\Gamma_\alpha} \frac{d\zeta}{\zeta I - \underline{t}}$, whose diagonal elements are $\frac{I}{2\pi i}\oint_{\Gamma_\alpha} \frac{d\zeta}{\zeta - \lambda} = \delta_{\alpha\lambda}$, with repetitions m_λ respectively. Thus it transforms $\underline{R}$ into the triangular matrix $\underline{S} = \sum_\alpha \underline{S}_\alpha$ whose diagonal elements are $\sum_\alpha \delta_{\alpha\lambda}$, all of which are unity, so that $\underline{R}$ is regular; and it transforms $(\underline{a} - \alpha \underline{I})\underline{R}_\alpha$ into $(\underline{t} - \alpha \underline{I})\underline{S}_\alpha$, whose diagonal elements are $(\lambda - \alpha)\delta_{\alpha\lambda}$, all of which are zero, so that $(\underline{a} - \alpha \underline{I})\underline{R}_\alpha$ is nullpotent..

We say that the $\underline{R}_\alpha$ define the independent idempotent decomposition of the unit matrix relative to $\underline{a}$, and that they constitute the partial units, relative to $\underline{a}$. [4]

THEOREM II. If $\Gamma_a \frown \underline{a}$ in Δ_a, then $\frac{I}{2\pi i}\oint_{\Gamma_a}\frac{d\zeta}{\zeta I-\underline{a}} = \underline{I}$

Let $\underline{R}_a = \frac{I}{2\pi i}\oint_{\Gamma_a}\frac{d\zeta}{\zeta I-\underline{a}}$, where $\Gamma_a \frown a$. Then, by theorem 3.III, $\underline{R}_a = \sum_\alpha \frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\zeta}{\zeta I-\underline{a}}$, where $\Gamma_\alpha \frown \alpha$ for each α; and hence by theorem 4.I, $\underline{R}_a = \sum_\alpha \underline{R}_\alpha = I$.

THEOREM III. The partial units $\underline{R}_\alpha$ are of rank m_α respectively, where m_α is the multiplicity of any α as a characteristic value of $\underline{a}$.

By transforming $\underline{a}$ to a triangular matrix it is seen that the zero characteristic value of $\underline{R}_\alpha$ is of multiplicity $\sum_{\lambda\neq\alpha} m_\lambda = n-m_\alpha$. Hence, since the rank of a characteristic value is at most its multiplicity, we have $\nu_\alpha \leq n-m_\alpha$, where ν_α denotes the nullity of $\underline{R}_\alpha$; or, if ρ_α is the rank of $\underline{R}_\alpha$, then $\rho_\alpha \geq m_\alpha$. Now the rank of the sum of any independent idempotents is the sum of their ranks. But the sum of the $\underline{R}_\alpha$ is the unit matrix, of rank n. Hence we would have $n=\sum_\alpha \rho_\alpha > \sum_\alpha m_\alpha = n$, providing a contradiction, unless $\rho_\alpha = m_\alpha$ for each α.

5. Representation of rational functions.

THEOREM I. If $p(\zeta)$ is a polynomial in ζ, and Γ a contour lying in Δ_a, then

$$p(a)\oint_{\Gamma}\frac{d\zeta}{\zeta I-a}=\oint_{\Gamma}\frac{p(\zeta)d\zeta}{\zeta I-a}$$

For, by the remainder theorem,

$$p(\zeta)-p(\omega)=(\zeta-\omega)q(\zeta,\omega),$$

where $q(\zeta,\omega)$ is a polynomial in ζ and ω. Therefore, for any ζ on Γ,

$$\frac{p(\zeta I)-p(a)}{\zeta I-a}=q(\zeta I,a)\quad,$$

and hence

$$\oint_{\Gamma}\frac{p(\zeta)d\zeta}{\zeta I-a}-p(a)\oint_{\Gamma}\frac{d\zeta}{\zeta I-a}=\oint_{\Gamma}q(\zeta I,a)d\zeta\quad.$$

But each element of $q(\zeta I,a)$ is a polynomial in ζ. Hence the integral of $q(\zeta I,a)$ round Γ is null, by Cauchy's theorem applied to each of the element integrals.

COROLLARY. If Γ_a is any contour which encloses all the characteristic values of a in its interior, then

$$p(a)=\frac{I}{2\pi i}\oint_{\Gamma_a}\frac{p(\zeta)d\zeta}{\zeta I-a}$$

For, if $\Gamma_a\sim a$, then $\oint_{\Gamma_a}\frac{d\zeta}{\zeta I-a}=2\pi i I$, by theorem 4.III.

THEOREM II. For any polynomial $p(\zeta)$ and a matrix a, a necessary and sufficient condition for $p(a)=0$ is that $\frac{p(\zeta I)}{\zeta I-a}$ should be a polynomial in ζ.

If $p(\zeta)$ is a polynomial in ζ, then $p(\zeta)-p(\omega)=(\zeta-\omega)q(\zeta,\omega)$ where $q(\zeta,\omega)$ is a polynomial in ζ and ω. Hence if

$p(\underline{a})=\underline{0}$ we have $\frac{p(\zeta\underline{I})}{\zeta\underline{I}-\underline{a}} = q(\zeta\underline{I},\underline{a})$, which is a polynomial in ζ. Conversely if $\frac{p(\zeta\underline{I})}{\zeta\underline{I}-\underline{a}}$ is a polynomial in ζ, let $\Gamma_a \sim \underline{a}$; then $p(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a}\frac{p(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} = \underline{0}$, by theorem 5.I, corollary, and then by Cauchy's theorem applied to each of the element integrals.

COROLLARY. *If* $C(\zeta)=|\zeta\underline{I}-\underline{a}|$ *then* $C(\underline{a})=\underline{0}$.

For $\frac{|\zeta\underline{I}-\underline{a}|}{\zeta\underline{I}-\underline{a}} = (\zeta\underline{I}-\underline{a})^*$ is the adjugate of $\zeta\underline{I}-\underline{a}$, and is a polynomial in ζ. The polynomial $C(\zeta)$ defines the *characteristic polynomial* of $\underline{a}$. The present result is the Cayley-Hamilton theorem, that every matrix satisfies its characteristic equation.[5]

THEOREM III. *If* $r(\zeta)$ *is a rational function of* ζ *whose poles lie outside* Δ_0, *and* Γ *is any contour in* Δ_a, *then*

$$r(\underline{a})\oint_\Gamma \frac{d\zeta}{\zeta\underline{I}-\underline{a}} = \oint_\Gamma \frac{r(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

Let $r(\zeta)$ have poles ρ of order m_ρ respectively. Then $r(\zeta)$, decomposed into its partial fractions, is of the form

$$r(\zeta) = p(\zeta)+\sum_\rho\sum_{\tau\le m_\rho}\frac{p_{\rho,\tau}(\zeta)}{(\rho-\zeta)^\tau},$$

where $p(\zeta)$ and the $p_{\rho,\tau}(\zeta)$ are polynomials in ζ. By theorem 5.I,

$$r_{\rho,\tau}(\underline{a})\oint_\Gamma\frac{d\zeta}{\zeta\underline{I}-\underline{a}} = \oint_\Gamma\frac{p_{\rho,\tau}(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}, \quad p(\underline{a})\oint_\Gamma\frac{d\zeta}{\zeta\underline{I}-\underline{a}} = \oint_\Gamma\frac{p_{\rho,\tau}(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

and then by theorem 3.IV, corollary,

$$\frac{I}{(\zeta\underline{I}-\underline{a})^\tau}\oint_\Gamma\frac{p_{\rho,\tau}(\zeta)d\zeta}{\zeta\underline{1}-\underline{a}} = \oint_\Gamma\frac{p_{\rho,\tau}(\zeta)d\zeta}{(\rho-\zeta)^\tau(\zeta\underline{I}-\underline{a})}$$

thus

$$\frac{p_{\rho,t}(\underline{a})}{(\rho\underline{I}-\underline{a})^t}\oint_\Gamma \frac{d\underline{\zeta}}{\underline{\zeta I}-\underline{a}} = \oint_\Gamma \frac{p_{\rho,t}(\underline{\zeta})d\underline{\zeta}}{(\rho-\underline{\zeta})^t(\underline{\zeta I}-\underline{a})} \quad ;$$

and accordingly,

$$r(\underline{a})\oint_\Gamma \frac{d\underline{\zeta}}{\underline{\zeta I}-\underline{a}} = p(\underline{a})\oint_\Gamma \frac{d\underline{\zeta}}{\underline{\zeta I}-\underline{a}} + \sum_\rho \sum_{t\le m_\rho} \frac{p_{\rho,t}(\underline{a})}{(\rho\underline{I}-\underline{a})^t}\oint_\Gamma \frac{d\underline{\zeta}}{\underline{\zeta I}-\underline{a}}$$

$$= \oint_\Gamma \frac{p(\underline{\zeta})d\underline{\zeta}}{\underline{\zeta I}-\underline{a}} + \sum_\rho \sum_{t\le m_\rho} \oint_\Gamma \frac{p_{\rho,t}(\underline{\zeta})d\underline{\zeta}}{(\rho-\underline{\zeta})^t(\underline{\zeta I}-\underline{a})}$$

$$= \oint_\Gamma \frac{r(\underline{\zeta})d\underline{\zeta}}{\underline{\zeta I}-\underline{a}} \quad .$$

6. The resolvent coefficients.

The monic polynomial $\phi(\zeta)$ of least degree such that $\phi(\underline{a})=\underline{0}$ defines the minimum polynomial of $\underline{a}$, and the degree i_a of this polynomial defines the index of $\underline{a}$. By theorem 5.II, corollary, $i_a \leq n$.

THEOREM I. The resolvent has the irreducible proper rational expression $\frac{\underline{I}}{\zeta\underline{I}-\underline{a}} = \frac{P(\zeta\underline{I},\underline{a})}{\phi(\zeta)}$, where $P(\zeta\underline{I},\underline{a})$ is a polynomial in $\underline{a}$ and ζ, and $\phi(\zeta)$ is the minimum polynomial of $\underline{a}$.

That the rational expression exists follows from theorem 5.II, since $\phi(\underline{a})=\underline{0}$. It is the expression $\frac{(\zeta\underline{I}-\underline{a})^*}{|\zeta\underline{I}-\underline{a}|}$ reduced, so that it is a proper rational expression; and moreover it is in its lowest terms. For were it reducible further, say to an expression with denominator $\psi(\zeta)$ of degree less than $\phi(\zeta)$, then it would follow from theorem 5.II that $\psi(\underline{a})=\underline{0}$, making a contradiction.

It appears that $\phi(\zeta)$ divides every polynomial $p(\zeta)$ such that $p(\underline{a})=\underline{0}$; for, by theorem 5.II, $p(\zeta)$ is then the denominator of a rational expression which reduces to that given in theorem 6.I with denominator $\phi(\zeta)$. Hence, also, $\phi(\zeta)$ is uniquely defined.

Thus $\phi(\zeta)$ divides $C(\zeta)$; moreover $P(\zeta\underline{I},\underline{a})$ divides $(\zeta\underline{I}-\underline{a})^*$, with the same quotient. Also, by taking determinants on either side of the expression given in theorem 6.I, $(\phi(\zeta))^n = C(\zeta)/P(\zeta\underline{I},\underline{a})/$, so that any root of

$C(\zeta)$ is a root of $\phi(\zeta)$. [6] Hence we have the following:

THEOREM II. $\phi(\zeta) = \prod_{\alpha}(\zeta-\alpha)^{i_\alpha}$ where $C(\zeta) = \prod_{\alpha}(\zeta-\alpha)^{m_\alpha}$ and $0<i_\alpha \leqslant m_\alpha$

The exponents m_α and i_α define the multiplicity and the index, respectively, of any characteristic value α. By theorem I and II, since the expression in theorem I is an irreducible proper polynomial fraction, this expression may be decomposed into its partial fractions, and the following fundamental expression for the resolvent is obtained:

THEOREM III. $\underline{R}(\zeta) = \sum_{\alpha} \sum_{0\leq r<i_\alpha} \frac{\underline{R}_{\alpha,r}}{(\zeta-\alpha)^{r+1}}$, where the coefficient matrices $\underline{R}_{\alpha,r}$ are all polynomials in $\underline{a}$, and $\underline{R}_{\alpha,i_\alpha} \neq \underline{0}$.

The matrices $\underline{R}_{\alpha,r}$ in this expression define the resolvent coefficients of $\underline{a}$.

Take any contour $\Gamma_\alpha - \alpha$ in Δ_a. Then, by theorem 6.II

$$\frac{I}{2\pi i}\oint_{\Gamma_\alpha} \frac{(\zeta-\alpha)^r}{\zeta\underline{I}-\underline{a}}\,d\zeta = \sum_{\lambda}\sum_{0\leq s<r_\lambda} \frac{\underline{R}_{\lambda,s}}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\zeta}{(\zeta-\lambda)^{s-r+1}}.$$

But Cauchy's formulae give

$$\frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\zeta}{(\zeta-\lambda)^{s-r+1}} = 0, \text{ unless } \lambda=\alpha \text{ and } s=r;$$

and

$$\frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\zeta}{\zeta-\alpha} = I.$$

Hence we have

$$\underline{R}_{\alpha,r} = \frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{(\zeta-\alpha)^r}{\zeta\underline{I}-\underline{a}}\,d\zeta$$

where $\underline{R}_{\alpha,r} = \underline{0}$, $r \geqslant i_\alpha$. But, by theorem 6.I,

$$\frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{(\zeta-\alpha)^r}{\zeta\underline{1}-\underline{a}}\,d\zeta = (\underline{a}-\alpha\underline{1})^r \oint_{\Gamma_\alpha}\frac{d\zeta}{\zeta\underline{1}-\underline{a}}.$$

Thus we deduce that

$$\underset{\sim}{R}_{\alpha,r} = (\underset{\sim}{a}-\alpha\underset{\sim}{I})^{r}\underset{\sim}{R}_{\alpha} ,$$

and, further, that

$$(\underset{\sim}{a}-\alpha\underset{\sim}{I})^{i_\alpha-1}\underset{\sim}{R}_{\alpha} \neq \underset{\sim}{0}, \text{ but } (\underset{\sim}{a}-\alpha\underset{\sim}{I})^{i_\alpha}\underset{\sim}{R}_{\alpha} = \underset{\sim}{0}.$$

The resolvent coefficients of $\underset{\sim}{a}$ are thus expressed in terms of the partial units relative to $\underset{\sim}{a}$, as follows:

THEOREM IV. *The resolvent coefficients for a matrix $\underset{\sim}{a}$ are given by*

$$\underset{\sim}{R}_{\alpha,r} = (\underset{\sim}{a}-\alpha\underset{\sim}{I})^{r}\underset{\sim}{R}_{\alpha}, \text{ where } \underset{\sim}{R}_{\alpha} = \frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\varsigma}{\varsigma\underset{\sim}{I}-\underset{\sim}{a}}$$

and Γ_α is any contour in Δ_a containing just the characteristic value α of $\underset{\sim}{a}$ in its interior, and where

$$(\underset{\sim}{a}-\alpha\underset{\sim}{I})^{r}\underset{\sim}{R}_{\alpha} = \underset{\sim}{0} \text{ if and only if } r \geqslant i_\alpha.$$

Consequently the following formula may be stated:

THEOREM V. $$\frac{I}{\varsigma\underset{\sim}{I}-\underset{\sim}{a}} = \sum_{\alpha}\underset{\sim}{R}_{\alpha}\sum_{0\leqslant r<i_\alpha}\frac{(\underset{\sim}{a}-\alpha\underset{\sim}{I})^{r}}{(\varsigma-\alpha)^{r+1}}, \text{ where } \underset{\sim}{R}_{\alpha} = \frac{I}{2\pi i}\oint_{\Gamma_\alpha}\frac{d\varsigma}{\varsigma\underset{\sim}{I}-\underset{\sim}{a}}.$$

We now prove the fundamental linear independence theorem for the resolvent coefficients.

THEOREM VI. *The resolvent coefficients $\underset{\sim}{R}_{\alpha,r} = (\underset{\sim}{a}-\alpha\underset{\sim}{I})^{r}\underset{\sim}{R}_{\alpha}$ $(r<i_\alpha,\alpha)$, form a linearly independent matrix set.*

Suppose they are linearly dependent, and there exist scalars $\lambda_{\alpha,r}$, not all zero, such that

$$\underset{\sim}{\phi} \equiv \sum_{\alpha}\sum_{r<i_\alpha}(\underset{\sim}{a}-\alpha\underset{\sim}{I})^{r}\underset{\sim}{R}_{\alpha} = \underset{\sim}{0}.$$

Then, since $\underset{\sim}{R}_{\alpha}^{2}=\underset{\sim}{R}_{\alpha}$ for any α, and $\underset{\sim}{R}_{\alpha}\underset{\sim}{R}_{\omega}=\underset{\sim}{0}$ $(\omega\neq\alpha)$, by theorem 4.I, we have

$$\underline{0}=\underline{R}_\alpha\underline{\phi} = \sum_{r<i_\alpha} \lambda_{\alpha,r}(\underline{a}-\alpha\underline{I})^r\underline{R}_\alpha = \phi_\alpha(\underline{a})\underline{R}_\alpha,$$

where $\phi_\alpha(\zeta) = \sum_{r<i_\alpha} \lambda_{\alpha,r}(\zeta-\alpha)^r$. Now

$$(\underline{a}-\alpha\underline{I})^s\phi_\alpha(\underline{a})\underline{R}_\alpha = \sum_{r+s<i_\alpha} \lambda_{\alpha,s}(\underline{a}-\alpha\underline{I})^{r+s}\underline{R}_\alpha,$$

since $(\underline{a}-\alpha\underline{I})^{r+s}\underline{R}_\alpha=\underline{0}$, $r+s>i_\alpha$, by theorem 6.V. Hence take $s=i_\alpha-I$, $i_\alpha-2,\ldots,0$ in succession. Then we successively get $0=\lambda_{\alpha,0}=\lambda_{\alpha,1}=\ldots=\lambda_{\alpha,i_\alpha-1}$. Thus we have $\lambda_{\alpha,r}=0$ $(r<i_\alpha,\alpha)$, contradicting the hypothesis.

COROLLARY. *For a polynomial $\phi(\zeta)$, $\phi(\underline{a})\underline{R}_\alpha=\underline{0}$ if and only if $(\zeta-\alpha)^{i_\alpha}$ divides $\phi(\zeta)$.*

The matrices $\underline{R}_\alpha$ are all polynomials in $\underline{a}$. So let $R_\alpha(\zeta)$ be a polynomial such that $\underline{R}_\alpha=R_\alpha(\underline{a})$. For any polynomial $\psi(\zeta)$, $\psi(\underline{a})=\underline{0}$ if and only if $\psi(\zeta)$ is divided by the minimum polynimial $\prod_\alpha (\zeta-\alpha)^{\gamma_\alpha}$ of $\underline{a}$. Hence $\phi(\underline{a})\underline{R}_\alpha=0$ if and only if $\prod_\alpha (\zeta-\alpha)^{\gamma_\alpha}$ divices $\phi(\zeta)R_\alpha(\zeta)$. But, by theorem V, $(\underline{a}-\alpha\underline{I})^r\underline{R}_\alpha=\underline{0}$ if and only if $r\geq i_\alpha$. Hence it follows that $\phi(\underline{a})\underline{R}_\alpha=0$ if and only if $(\zeta-\alpha)^{\gamma_\alpha}$ divides $\phi(\zeta)$.

7'. The principal and axial spaces.

LEMMA. If $p(\zeta)$, $q(\zeta)$ are polynomials which have no root in common, $\underline{p}=p(\underline{a})$ and $\underline{q}=q(\underline{a})$, where $\underline{a}$ is any matrix, then

$$\mathcal{N}_p \cap \mathcal{N}_q \quad \& \quad \mathcal{N}_p \oplus \mathcal{N}_q = \mathcal{N}_{pq}$$

where $\mathcal{N}_x$ denotes the null-space of any matrix $\underline{x}$.

In other words, if ν_x denotes the nullity of $\underline{x}$, then $\nu_{pq}=\nu_p+\nu_q$. By Sylvester's law of nullity, $\nu_{pq} \leq \nu_p+\nu_q$. So it remains to prove that $\nu_{pq} \geq \nu_p+\nu_q$, and since $\mathcal{N}_{pq}$ contains both $\mathcal{N}_p$ and $\mathcal{N}_q$, it will be sufficient to show that $\mathcal{N}_p \cap \mathcal{N}_q = \mathcal{O}$.

Since $p(\zeta)$, $q(\zeta)$ have no root in common, by Euclid's algorithm there exist polynomials $h(\zeta)$, $k(\zeta)$ such that $h(\zeta)p(\zeta)+h(\zeta)q(\zeta)=1$; and then, if $\underline{h}=h(\underline{a})$, $\underline{k}=k(\underline{a})$, we have $\underline{hp}+\underline{kq}=\underline{I}$. Now if there exists a vector $\underline{X} \neq \underline{0}$ common to $\mathcal{N}_p$ and $\mathcal{N}_q$, that is such that $\underline{pX}=\underline{0}$ and $\underline{qX}=\underline{0}$, we would have $\underline{X}=\underline{hpX}+\underline{kqX}=\underline{0}$, making a contradiction. Hence $\mathcal{N}_p \cap \mathcal{N}_q = \mathcal{O}$.

THEOREM I. The range of $\underline{R}_\alpha$ is the null-space of $(\underline{a}-\alpha\underline{I})^{i_\alpha}$.

Let $\mathcal{U}_\alpha$ denote the range of $\underline{R}_\alpha$, and $\mathcal{N}_\alpha$ the null-space of $(\underline{a}-\alpha\underline{I})^{i_\alpha}$. By theorem 6.IV, we have $(\underline{a}-\alpha\underline{I})^{i_\alpha}\underline{R}_\alpha=\underline{0}$, so we conclude that $\mathcal{U}_\alpha \subset \mathcal{N}_\alpha$. By theorem 4.III, the space $\mathcal{U}_\alpha$ is of rank m_α. Let ν_α be the rank of $\mathcal{N}_\alpha$, that is to say the nullity of $(\underline{a}-\alpha\underline{I})^{i_\alpha}$, so that we have $\nu_\alpha \geq m_\alpha$. If then follows from theorem above that the nullity of $\prod_\alpha (\underline{a}-\alpha\underline{I})^{i_\alpha}=\underline{0}$ is $\sum_\alpha \nu_\alpha = n$. But $\sum_\alpha \nu_\alpha > \sum_\alpha m_\alpha = n$, making a

contradiction, unless $\nu_\alpha = m_\alpha$, for each α. Hence $\mathcal{U}_\alpha = \mathcal{M}_\alpha$.

The range $\mathcal{U}_\alpha$ of the partial resolvent R_α we call the principal space relative to a for the characteristic value α. Thus the principal spaces relative to any matrix a constitute a splitting of the unit space $\mathcal{I}$ into independent subspaces $\mathcal{U}_\alpha$, of rank m_α respectively:

$$\mathcal{U}_\alpha \cap \mathcal{U}_\lambda = \mathcal{O}, (\lambda \neq \alpha); \quad \bigoplus_\alpha \mathcal{U}_\alpha = \mathcal{I}$$

Also we define the axial space $\mathcal{X}_\alpha$ of a corresponding to α as the space of axial or latent vectors of a with characteristic value α, that is to say the space of vectors X such that $aX = X\alpha$; and the rank r_α of $\mathcal{X}_\alpha$ defines the rank of any α as a characteristic value of a. Since $\mathcal{X}_\alpha$ is the same as the null-space of $(a-\alpha I)$, and since $(a-\alpha I)^{\gamma_\alpha}\mathcal{U}_\alpha = \mathcal{O}$ by theorem I, we deduce that $\mathcal{X}_\alpha \subset \mathcal{U}_\alpha$, for each α, or:

THEOREM II. *Any axial space of a matrix is contained in the corresponding principal space relative to that matrix.*

From this we immediately deduce the following:

COROLLARY. *The rank of a characteristic value is at most its multiplicity.* [18]

7. The principal polynomials.

Take $\underline{p}_\alpha = \prod_{\lambda \neq \alpha} (\underline{a}-\lambda \underline{I})^{i_\lambda}$ and $\underline{p} = \sum_\alpha \underline{p}_\alpha$. The characteristic values of $\underline{p}$ are $p(\alpha) = \sum_\lambda \sum_{\mu \neq \lambda} (\alpha-\mu)^{i_\mu}$, none of which is zero. Hence $\underline{p}$ is regular, and thus has an inverse $\underline{p}^{-1}$ which is a polynomial in $\underline{p}$ and hence also in $\underline{a}$. Thus there are matrices $\underline{e}_\alpha = \underline{p}_\alpha/\underline{p}$ defined, and all are polynomials in $\underline{a}$. They are called the principal polynomials in $\underline{a}$, and have the following properties: [7]

THEOREM I. *The principal polynomials $\underline{e}_\alpha$ in a matrix $\underline{a}$ are such that $\underline{e}_\alpha^2 = \underline{e}_\alpha \neq 0$, $\underline{e}_\alpha \underline{e}_\lambda = 0$ $(\lambda \neq \alpha)$, $\sum_\alpha \underline{e}_\alpha = \underline{I}$, and for any polynomial $\phi(\gamma)$, $\phi(\underline{a})\underline{e}_\alpha = 0$ if and only if $(\gamma-\alpha)^{i_\alpha}$ divides $\phi(\gamma)$.*

These results follow directly from the definition of the $\underline{e}_\alpha$, and the properties of the minimum polynomial of a matrix.

The principal polynomials in a matrix are uniquely defined with respect to certain of their properties, as follows:

THEOREM II. *The principal polynomials in any matrix $\underline{a}$ are the unique matrices $\underline{e}_\alpha$, indexed by the characteristic values α of $\underline{a}$, such that $\underline{a}\underline{e}_\alpha = \underline{e}_\alpha \underline{a}$, $\underline{e}_\alpha^2 = \underline{e}_\alpha \neq 0$, $\sum_\alpha \underline{e}_\alpha = \underline{I}$ and $(\underline{a}-\alpha \underline{I})\underline{e}_\alpha$ is nullpotent.*

Suppose that $\underline{f}_\alpha$ could be another set of such matrices. Then, since $\underline{a}\underline{f}_\alpha = \underline{f}_\alpha \underline{a}$, and the $\underline{e}_\alpha$ are

polynomials in $\underline{a}$, we have $\underline{e}_\mu \underline{f}_\lambda = \underline{f}_\lambda \underline{e}_\mu$; so if $\underline{g}_{\lambda\mu} = \underline{f}_\lambda \underline{e}_\mu$, then $\underline{g}_{\lambda\mu}^2 = \underline{g}_{\lambda\mu}$; and thus $\underline{g}_{\lambda\mu}$ is nullpotent only if it is null. The matrices $\underline{p}_\mu = (\underline{a} - \mu \underline{I})\underline{e}_\mu$, $\underline{q}_\lambda = (\underline{a} - \lambda \underline{I})\underline{f}_\lambda$ are nullpotent. Hence the matrix

$$(\lambda - \mu)\underline{g}_{\lambda\mu} = \underline{p}_\mu \underline{f}_\lambda - \underline{q}_\lambda \underline{e}_\mu$$

is nullpotent. But, if $\lambda \neq \mu$, this implies that $\underline{g}_{\lambda\mu}$ is nullpotent, and hence null. That is, $\underline{f}_\lambda \underline{e}_\mu = 0$ $(\lambda \neq \mu)$. Thus we have

$$\underline{e}_\alpha = \underline{e}_\alpha \sum_\lambda \underline{f}_\lambda = \underline{e}_\alpha \underline{f}_\alpha = \underline{f}_\alpha \sum_\mu \underline{e}_\mu = \underline{f}_\alpha .$$

THEOREM III. <u>The partial units relative to any matrix $\underline{a}$ and the principal polynomials in that matrix define the same quantities</u>: $\underline{e}_\alpha = \underline{R}_\alpha$.

For, by theorem 4.1, the partial units $\underline{R}_\alpha$ relative to a conditions which, by theorem 7.II, are sufficient to define the principal polynomials $\underline{e}_\alpha$.

8. Analytic functions.

A multi-valued analytic function $f(\zeta)$ of a complex variable ζ may be extended to a multivalued function $f(\underline{a})$ of a complex matrix variable $\underline{a}$ by the following definition: [8]

DEFINITION. If $f(\zeta)$ is any analytic function of a complex variable, $\underline{a}$ any matrix over the complex field whose characteristic values lie in the domain of regularity of $f(\zeta)$, $\bar{\Phi}$ a Riemann surface on which $f(\zeta)$ is represented as single valued and regular, $\bar{\Phi}_a$ the Riemann surface $\bar{\Phi}$ with the characteristic values of $\underline{a}$ removed, and Γ_a an arbitrary contour in $\bar{\Phi}_a$ which encloses all the characteristic values of $\underline{a}$ simply in its interior, then an arbitrary determination of $f(\underline{a})$ is defined by

$$f(\underline{a}) = \frac{I}{2\pi i} \int_{\Gamma_a} \frac{f(\zeta)d\zeta}{\zeta I - \underline{a}}$$

Since the contour Γ_a encloses each characteristic value α of $\underline{a}$ simply in its interior, it will associate with each characteristic value α a definite sheet of the Riemann surface $\bar{\Phi}$, and hence a definite determination of the analytic function $f(\zeta)$ at and in the neighbourhood of α, which will be called the branch of $f(\zeta)$ determined by Γ_a about α, and denoted by $f_\alpha(\zeta)$. The determinations of $f(\underline{a})$ are in a one-to-one correspondence with the combinations of the determinations of $f(\zeta)$ at the characteristic values of $\underline{a}$; and with any determination of $f(\underline{a})$ there is thus associated a set of characteristic deter-

minations of $f(\zeta)$, one for each characteristic value of $\underline{a}$. This is shown explicitly as follows:

[9]
THEOREM I. *The various determinations of* $f(\underline{a})$ *are given by*

$$f(\underline{a}) = \frac{I}{2\pi i} \sum_{\alpha} \oint_{\Gamma_\alpha} \frac{f_\alpha(\zeta)d\zeta}{\zeta I - \underline{a}}$$

where Γ_α *is any simple contour which encloses no singularities of* $f(\zeta)$, *and just the characteristic value* α *of* a, *in its interior, and* $f_\alpha(\zeta)$ *is any single valued and regular determination of* $f(\zeta)$ *throughout* Γ_α.

The proof follows those given in § 3.

COROLLARY. $$f(\underline{a}) = \sum_{\alpha} \frac{R_\alpha}{2\pi i} \oint_{\Gamma'_\alpha} \frac{f_\alpha(\zeta)d\zeta}{\zeta I - \underline{a}}$$

where Γ'_α *is any contour which contains* α *simply in its interior*.

For we have $\Gamma_\alpha \cdot \Gamma'_\alpha \sim \Gamma_\alpha$, and $R_\alpha = \frac{I}{2\pi i} \oint_{\Gamma_\alpha} \frac{d\zeta}{\zeta I - \underline{a}}$, and hence, by theorem 3.V,

$$\frac{R_\alpha}{2\pi i} \oint_{\Gamma'_\alpha} \frac{f_\alpha(\zeta)d\zeta}{\zeta I - \underline{a}} = \frac{I}{2\pi i} \oint_{\Gamma_\alpha} \frac{f_\alpha(\zeta)d\zeta}{\zeta I - \underline{a}} .$$

By our definition we thus have $f(\underline{a})$ single valued when $f(\zeta)$ is single valued.

Theorem I expresses $f(\underline{a})$ in terms of the values of $f(\zeta)$ on contours surrounding each characteristic value of $\underline{a}$. The formula in the next theorem expresses $f(\underline{a})$ in t terms of $f(\zeta)$ and its derivatives up to certain orders

at each of the characteristic values of $\underline{a}$.

THEOREM II.[10] $$f(\underline{a}) = \sum_{\alpha} \underline{R}_{\alpha} \sum_{r<i_{\alpha}} \frac{I}{r!} f_{\alpha}^{(r)}(\alpha)(\underline{a}-\alpha\underline{I})^{r}$$

By substitution from theorem 6.I, we have

$$\frac{I}{2\pi i}\oint_{\Gamma_{\alpha}} \frac{f_{\alpha}(\zeta)}{\zeta\underline{I}-\underline{a}}d\zeta = \frac{I}{2\pi i}\oint_{\Gamma_{\alpha}} \sum_{\lambda} \underline{R}_{\lambda} \sum_{r<i_{\lambda}} \frac{(\underline{a}-\lambda\underline{I})^{r}}{(\zeta-\lambda)^{r+1}} f_{\alpha}(\zeta)d\zeta$$

$$= \sum_{\lambda} \underline{R}_{\lambda} \sum_{r<i_{\lambda}} (\underline{a}-\lambda\underline{I})^{r} \cdot \frac{I}{r!}\oint_{\Gamma_{\alpha}} \frac{f_{\alpha}(\zeta)}{(\zeta-\lambda)^{r+1}}d\zeta$$

Since Γ_{α} contains just the characteristic value α of $\underline{a}$ in its interior, it follows from Cauchy's formulae that

$$\frac{I}{2\pi i}\oint_{\Gamma_{\alpha}} \frac{f_{\alpha}(\zeta)}{(\zeta-\lambda)}d\zeta = \frac{\delta_{\alpha\lambda}}{r!} f_{\alpha}^{(r)}(\alpha)$$

Hence

$$\frac{I}{2\pi i}\oint_{\Gamma_{\alpha}} f_{\alpha}(\zeta)d\zeta = \underline{R}_{\alpha} \sum_{r<i_{\alpha}} (\underline{a}-\alpha\underline{I})^{r} \frac{I}{r!} f_{\alpha}^{(r)}(\alpha) ,$$

and, by summation with respect to α, the required formula is obtained from theorem I.

COROLLARY. $f(\underline{a})=\underline{0}$ *if and only if* $f_{\alpha}^{(r)}(\alpha)=0$ $(r<i_{\alpha}, \alpha)$.

For, by theorem 6.VI, the matrices $(\underline{a}-\alpha\underline{I})^{r}\underline{R}_{\alpha}$ $(r<i_{\alpha}, \alpha)$ form a linearly independent set.

COROLLARY. If $f_m(\zeta)$ is a sequence of analytic functions, and $\underline{a}$ any matrix, then the sequence $f_m^{(r)}(\underline{a})$ is convergent if and only if the sequences $f_m^{(r)}(\alpha), (r<i_\alpha, \alpha)$, are all convergent, and, moreover, $f_m(\underline{a}) \to f(\underline{a}), (m\to\infty)$, if and only if $f_m^{(r)}(\alpha) \to f^{(r)}(\alpha), (m\to\infty), (r<i_\alpha, \alpha)$.

In the following theorems, $f(\zeta)$, $g(\zeta)$,... are any analytic functions regular in some domain Δ and $\underline{a}$ is any matrix whose characteristic values lie in Δ.

THEOREM III. If $u(\zeta)=\lambda f(\zeta)$, $v(\zeta)=f(\zeta)+g(\zeta)$ and $w(\zeta)=f(\zeta)g(\zeta)$ then $u(\underline{a})=\lambda f(\underline{a})$, $v(\underline{a})=f(\underline{a})+g(\underline{a})$ and $w(\underline{a})=f(\underline{a})g(\underline{a})$.

The first two relations here express the linearity of the operation of integration. The third expresses a corollary to theorem 3.V.

THEOREM IV. If $h(\zeta) = f(g(\zeta))$, then $h(\underline{a}) = f(g(\underline{a}))$.

By definition,

$$f(g(\underline{a})) = \frac{I}{2\pi i} \oint_{\Phi_g} \frac{f(\lambda)d\lambda}{\lambda I - g(\underline{a})} ,$$

where Φ_g is any contour which encloses all the characteristic values $g(\alpha)$ of $g(\underline{a})$ in its interior. Let $\lambda = g(\zeta)$ so if ζ describes any contour Ψ_a which encloses in its interior all the characteristic values of $\underline{a}$, then λ describes a contour Φ_g which encloses all the characteristic values of $g(\underline{a})$. Thus we have

$$f(\phi(\underline{a})) = \frac{I}{2\pi i}\oint_{\Psi_a}\frac{f(g(\zeta))g'(\zeta)d\zeta}{g(\zeta)-g(\underline{a})}$$

$$= \frac{I}{2\pi i}\oint_{\Psi_a}\frac{h(\zeta)g'(\zeta)d\zeta}{g(\zeta)-g(\underline{a})}$$

$$= \frac{I}{2\pi i}\oint_{\Psi_a}\frac{h(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} + \frac{I}{2\pi i}\oint_{\Psi_a}\left\{\frac{1}{2\pi i}\oint_{\Gamma_a}\left(\frac{g'(\zeta)}{g(\zeta)-g(\omega)}-\frac{1}{\zeta-\omega}\right)\frac{d\omega}{\omega\underline{I}-\underline{a}}\right\}h(\zeta)d\zeta$$

$$= h(\underline{a}) + \frac{1}{2\pi i}\oint_{\Gamma_a}\left\{\frac{1}{2\pi i}\oint_{\Psi_a}\left(\frac{g'(\zeta)}{g(\zeta)-g(\omega)}-\frac{1}{\zeta-\omega}\right)h(\zeta)d\zeta\right\}\frac{d\omega}{\omega\underline{I}-\underline{a}}$$

$$= h(\underline{a}) \quad ,$$

since each ω on Γ_a lies in the exterior of the contour Ψ_a, so that the function $F(\zeta)=\left(\frac{g'(\zeta)}{g(\zeta)-g(\omega)}-\frac{1}{\zeta-\omega}\right)h(\zeta)$ is regular on and in the interior of the contour Ψ_a, and hence the integral of $F(\zeta)$ round Ψ_a is zero for each ω on Γ_ω, by Cauchy's theorem.

THEOREM V. *If $\zeta = g(\omega)$ is the inverse of $\omega = f(\zeta)$, then $\underline{a} = g(\underline{b})$ is the inverse of $\underline{b} = f(\underline{a})$.*

Thus we have $f(\phi(\omega)) = \omega$; and hence, by theorem IV, $f(g(\underline{b})) = \underline{b}$.

9. Regularity.

An extension has been given of any analytic function $f(\zeta)$, of a complex variable ζ, to a function $f(\underline{z})$ of a complex matrix variable $\underline{z}$, defined in the domain $\underline{D}$ of matrices whose characteristic values all belong to the domain of regularity Δ of the function $f(\zeta)$; and this extension has been shown in §5 to comprehend the natural extension of polynomials and rational functions. Let the regularity of the function $f(\underline{z})$, for any matrix $\underline{a}$ in D, be defined in the sense of the quotient $\frac{f(\underline{b})-f(\underline{a})}{\underline{b}-\underline{a}}$ having a uniquely defined limit for when $\underline{b}$, always commuting with $\underline{a}$, tends to $\underline{a}$ along an arbitrary path on which $\frac{|\underline{b}-\underline{a}|^*}{|\underline{b}-\underline{a}|_*}$ is bounded. If $f(\underline{z})$ is regular at $\underline{a}$, let this limit define the derivative of $f(\underline{z})$ at $\underline{a}$.

The function $f(\zeta)$ is regular in Δ, and has as its derivative a function $f'(\zeta)$, which is an analytic function which has Δ for its domain of regularity. Moreover, the function $f'(\zeta)$, defined and regular just in Δ, extends, by §3, to a function $f'(\underline{z})$ which is defined in $\underline{D}$. [11] We now show the following:

THEOREM I. *$f(\underline{z})$ is regular in $\underline{D}$, and $f'(\underline{z})$ is its derivative.*

Let

$$f(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{\zeta I-\underline{a}}$$

according to the definition of §8; let $\underline{b}$ commute with

$\underline{a}$, and be such that $|\underline{b}-\underline{a}|_* \neq 0$, or $\underline{b}-\underline{a}$ is regular; and let $|\underline{b}-\underline{a}|^*$ be suitably small, making $\underline{b}$ suitably close to $\underline{a}$, and thus the contour Γ_a such that we can take

$$f(\underline{b}) = \frac{I}{2\pi i} \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} \quad .$$

Then we have

$$\frac{f(\underline{b})-f(\underline{a})}{\underline{b}-\underline{a}} = \frac{I}{2\pi i} \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{(\zeta\underline{I}-\underline{b})(\zeta\underline{I}-\underline{a})} \quad ,$$

and

$$\oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{(\zeta\underline{I}-\underline{b})(\zeta\underline{I}-\underline{a})} - \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{(\zeta\underline{I}-\underline{a})^2} = (\underline{b}-\underline{a}) \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{(\zeta\underline{I}-\underline{b})(\zeta\underline{I}-\underline{a})} \to 0,$$

$(\underline{b} \to \underline{a})$.

It follows that $f(\underline{a})$ is regular in $\underline{D}$, and has derivative

$$\frac{I}{2\pi i} \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{(\zeta\underline{I}-\underline{a})^2} \quad .$$

But it follows from theorem 3.IV, corollary, and theorem 4.II, that

$$\frac{I}{(\zeta\underline{I}-\underline{a})^2} = \frac{I}{2\pi i} \oint_{\Gamma_a'} \frac{d\omega}{(\zeta-\omega)^2(\zeta\underline{I}-\underline{a})} \quad ,$$

and thus another expression for the derivative is

$$\frac{I}{2\pi i} \oint_{\Gamma_a'} \left(\frac{I}{2\pi i} \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{(\zeta-\omega)^2} \right) \frac{d\omega}{(\omega\underline{I}-\underline{a})}$$

$$= \frac{I}{2\pi i} \oint_{\Gamma_a} \frac{f'(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

$$= f'(\underline{a})$$

where Γ_a is taken surrounding Γ_a', the order of the integrations involved is interchanged, and Cauchy's formula is used for the derivative $f'(\zeta)$.

Proceeding further, since all the derivatives $f^{(r)}(\zeta)$, $r>0$, of the function $f(\zeta)$, are regular in Δ, we also have the regularity of the function $f^{(r)}(\underline{a})$ in $\underline{D}$, and that its derivative is $f^{(r+1)}(\underline{a})$. By a function $f(\underline{a})$ being <u>differentiable</u> in $\underline{D}$, we mean it has derivatives in $\underline{D}$ of all orders ~~in~~. Now, by recursion, we have that $f(\underline{a})$ is differentiable in $\underline{D}$, and that $f^{(r)}(\underline{a})$ is its rth derivative. Here

$$f^{(r)}(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a}\frac{f^{(r)}(\zeta)d\zeta}{\zeta I-\underline{a}} ,$$

and Cauchy's formula gives

$$f^{(r)}(\zeta) = \frac{r!}{2\pi i}\oint_{\Gamma_\zeta}\frac{f(\omega)d\omega}{(\omega-\zeta)^{r+1}} ,$$

where Γ_ζ is any contour in Δ whose interior contains ζ. Thus, since ζ ranges on Γ_a, take Φ_a as any contour in Δ_a such that $\Phi_a \sim a$, and which encloses Γ_a in its interior. Then we have

$$f^{(r)}(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a}\left(\frac{r!}{2\pi i}\oint_{\Phi_a}\frac{f(\omega)d\omega}{(\omega-\zeta)^{r+1}}\right)\frac{d\zeta}{\zeta I-\underline{a}}$$

$$= \frac{r!}{2\pi i}\oint_{\Phi_a}\left(\frac{I}{2\pi i}\oint_{\Gamma_a}\frac{d\zeta}{(\omega-\zeta)^{r+1}(\zeta I-\underline{a})}\right)f(\omega)d\omega$$

$$\frac{r!}{2\pi i}\oint_{\Phi_a}\frac{f(\omega)d\omega}{(\omega I-\underline{a})^{r+1}} ,$$

by interchanging the order of the integrations, and using theorem 3.IV, corollary, and theorem 4.II.

We have proved the following theorem:

THEOREM II. $f(\underline{a})$ <u>is differentiable in</u> $\underline{D}$, <u>and its rth is given by the equivalent expressions</u>

$$f^{(r)}(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a}\frac{f^{(r)}(\zeta)d\zeta}{\zeta I-\underline{a}} = \frac{r.'}{2\pi i}\oint_{\Gamma_a}\frac{f(\zeta)d\zeta}{(\zeta I-\underline{a})^{r+1}}$$

10. Infinite powers and geometric sums.

THEOREM I. *If $\underset{\sim}{a}$ is a matrix of order n, and R is the maximum of the moduli of its characteristic values, then*

$$|\underset{\sim}{a}^{m+1}|^* \leqslant n^2 \left(\frac{|\underset{\sim}{a}|^*}{R}\right)^{n-1} \frac{m^n}{n!} R^{m+1},$$

provided $m > n$ *and* $R \neq 0$.

Let $\underset{\sim}{t} = \underset{\sim}{U}^{-1} \underset{\sim}{a} \underset{\sim}{U}$ be any triangular unitary transform of $\underset{\sim}{a}$; so that the diagonal elements of $\underset{\sim}{t}$ are the characteristic values of $\underset{\sim}{a}$, of modulus at most the characteristic radius R, and $|\underset{\sim}{t}|^* = |\underset{\sim}{a}|^*$; [12] and

$$t_{ij}^{m+1} = \sum_{i \leqslant k_1 \leqslant \cdots \leqslant k_m \leqslant j} t_{ik_1} t_{k_1 k_2} \cdots t_{k_{m-1} k_m} t_{k_m j},$$

since $t_{rs} = 0$ $(r>s)$. Since the m indices $k_1, \ldots, k_m$ have a range of $r = j-i+1$ values, if $m>n$ they have coincidences which involve there being at least $m-r+2$ diagonal elements and accordingly at most $r-1$ diagonal elements forming the $m+1$ factors in each of the $\frac{m^r}{m!}$ terms of t_{ij}^{m+1}. Hence

$$|t_{ij}^{m+1}| \leqslant \frac{m^r}{m!} (|\underset{\sim}{a}|^*)^{r-1} R^{m-r+2} \leqslant \frac{m^n}{m!} \left(\frac{|\underset{\sim}{a}|^*}{R}\right)^{n-1} R^{m+1},$$

since $|t_{ij}| \leqslant |\underset{\sim}{t}|^*$, and $R \leqslant |\underset{\sim}{a}|^*$ and now the required formula is obtained from $|\underset{\sim}{a}^{m+1}|^* = |\underset{\sim}{t}^{m+1}|^* \leqslant \sum_{ij} |t_{ij}^{m+1}|$.

COROLLARY. $|\underset{\sim}{a}^m|^* = o(R^m)$ $(m \to \infty)$.

THEOREM II. [13] *$\underset{\sim}{a}^m \to 0$ $(m \to \infty)$ if and only if $\alpha^m \to 0$ $(m \to \infty ; \alpha)$.*

The characteristic values of $\underset{\sim}{a}^m$ are α^m. Hence

$\alpha^m \to 0$ for each α if $\underset{\sim}{a}^m \to \underset{\sim}{0}$ $(m\to\infty)$. Conversely if $\alpha^m \to 0$ $(m\to\infty;\alpha)$, then $|\alpha|<1$ for each α, and hence $R<1$, so that $R^m \to 0$ $(m\to\infty)$. Accordingly $\underset{\sim}{a}^m \to \underset{\sim}{0}$ $(m\to\infty)$, by theorem I, corollary.

THEOREM III. $\sum_{t<m} \underset{\sim}{a}^t \to \frac{\underset{\sim}{I}}{\underset{\sim}{I}-\underset{\sim}{a}}$ *if and only if* $|\alpha|<1$ *for each* α.

If $|\underset{\sim}{I}-\underset{\sim}{a}| \neq 0$, so that $\frac{\underset{\sim}{I}}{\underset{\sim}{I}-\underset{\sim}{a}}$ exists, then $\alpha \neq 1$ for each α, and

$$\frac{\underset{\sim}{I}}{\underset{\sim}{I}-\underset{\sim}{a}} - \sum_{t<m} \underset{\sim}{a}^t = \frac{\underset{\sim}{a}^m}{\underset{\sim}{I}-\underset{\sim}{a}} ;$$

where $\frac{\underset{\sim}{a}^m}{\underset{\sim}{I}-\underset{\sim}{a}} \to \underset{\sim}{0}$ if and only if $|\alpha|<1$ for each α, by theorem II.

THEOREM IV. $\sum_{t<m} \underset{\sim}{a}^t$ *is uniformly convergent in any bounded domain of matrices where the characteristic radius is less than unity.*

Let $\underset{\sim}{D}$ be such a domain, so that $\delta = \inf_{D} |\underset{\sim}{I}-\underset{\sim}{a}|_* \neq 0$, and $K_0 = \sup_{D} |\underset{\sim}{a}|^* < \infty$, and $R \leq R_0 < 1$. By theorem I, for all $\underset{\sim}{a}$ in $\underset{\sim}{D}$, and $m > n$,

$$\left|\frac{\underset{\sim}{I}}{\underset{\sim}{I}-\underset{\sim}{a}} - \sum_{t<m} \underset{\sim}{a}^t\right|^* = \left|\frac{\underset{\sim}{a}^{m+1}}{\underset{\sim}{I}-\underset{\sim}{a}}\right|^* \leq \frac{|\underset{\sim}{a}^{m+1}|^*}{|\underset{\sim}{I}-\underset{\sim}{a}|_*}$$

$$\leq \frac{n^2}{\delta}\left(\frac{|\underset{\sim}{a}|^*}{R}\right)^{n-1} \frac{m^n}{m!} R^{m+1}$$

$$\leq \frac{n^2}{\delta}\left(\frac{K_0}{R_0}\right)^{n-1} R_0^m \frac{m^n}{m!}$$

Then, since $R_0^m \frac{m^n}{m!} \to 0$ $(m\to\infty)$, there exists an m_0 such that for all $\underset{\sim}{a}$ in D,

$$\left|\frac{\underset{\sim}{I}}{\underset{\sim}{I}-\underset{\sim}{a}} - \sum_{t<m} \underset{\sim}{a}^t\right|^* < \varepsilon \quad \text{if } m > m_0 .$$

II. Characteristic values.

THEOREM I. *If $\underline{a}$ and $\underline{b}$ commute, and $\underline{b}-\underline{a}$ is nullpotent, then $\underline{a}$ and $\underline{b}$ have the same characteristic values.*

Let $p(\rho)$, $q(\rho)$ be the minimum polynomials of $\underline{a}$ and $\underline{b}$. Since $\underline{a}$ and $\underline{b}$ commute,

$$p(\underline{b})-p(\underline{a}) = (\underline{b}-\underline{a})P(\underline{b},\underline{a}),$$

where $P(\underline{a},\underline{b})$ is a polynomial in $\underline{a}$ and $\underline{b}$; and hence, since $p(\underline{a})=0$, and $\underline{b}-\underline{a}$ is nullpotent, say of order r, we have

$$(p(\underline{b}))^r = (\underline{b}-\underline{a})^r(P(\underline{b},\underline{a}))^r = \underline{0}.$$

But $q(\rho)$ is the minimum polynomial of $\underline{b}$, and hence $q(\rho)$ divides $(p(\rho))^r$. Similarly $p(\rho)$ divides $(q(\rho))^r$, and accordingly $p(\rho)$, $q(\rho)$ have the same roots, and $\underline{a},\underline{b}$ the same characteristic values.

If any matrices $\underline{a},\underline{b}$ commute, and the partial units relative to $\underline{a}$ and $\underline{b}$ are $\underline{R}_\alpha$, and $\underline{S}_\beta$ respectively, we say [14] α, β are *latently associated* characteristic values of $\underline{a}$ and $\underline{b}$, and write $\alpha\frown\beta$, if $\underline{R}_\alpha\underline{S}_\beta\neq\underline{0}$.

If $p(\lambda,\mu)$ is any polynomial in λ,μ then

$$p(\lambda,\mu)-p(\alpha,\beta) = (\lambda-\alpha)L + (\mu-\beta)M,$$

where L and M are polynomial in λ,α,μ and α,μ,β respectively. Hence, if $\underline{a}$ and $\underline{b}$ commute, we have

$$(p(\underline{a},\underline{b})-p(\alpha,\beta)I)\underline{R}_\alpha\underline{S}_\beta$$

is nullpotent. It follows, from theorem I, that $p(\underline{a},\underline{b})\underline{R}_\alpha\underline{S}_\beta$ has just the characteristic values of $\underline{J}=p(\alpha,\beta)\underline{R}_\alpha\underline{S}_\beta$. But

$J^2=p(\alpha,\beta)J$, since $E_\alpha S_\beta$ is idempotent. Hence the characteristic values of J are 0 and $p(\alpha,\beta)$, and we have the following:

THEOREM II. *If $p(\lambda,\mu)$ is a polynomial, and E_α, S_β are partial units relative to commutative matrices a, b, and $E_\alpha S_\beta \neq 0$, then the characteristic values of $p(a,b)E_\alpha S_\beta$ are 0 and $p(\alpha,\beta)$.*

Proceeding further, since $\sum_{\alpha,\beta} E_\alpha S_\beta = (\sum_\alpha E_\alpha)(\sum_\beta S_\beta)=I$, we have

$$p(a,b) = \sum_{\alpha,\beta} p(\alpha,\beta)E_\alpha S_\beta + N$$

where N is a nullpotent polynomial in a and b. Thus, by theorem I, $p(a,b)$ has the characteristic values of $\sum_{\alpha,\beta} p(\alpha,\beta)E_\alpha S_\beta$; and, since the non-null matrices $E_\alpha S_\beta$ form an independent idempotent decomposition of the unit matrix, these are the values of $p(\alpha,\beta)$ for when $E_\alpha S_\beta \neq 0$.

THEOREM III. *The characteristic values of any polynomial $p(a,b)$ of commutative matrices a,b are the values of $p(\alpha,\beta)$ for when α,β are latent, associated characteristic values of a, b.*[15]

12. Taylor expansions.

Let $f(\zeta)$ be an analytic function, and let

$$f(\underline{a}) = \frac{I}{2\pi i} \int_{\Gamma_a} f(\zeta)\, d\zeta \,(\zeta \underline{I} - \underline{a})^{-1}$$

be the extension of $f(\zeta)$, according to the definition of § 8, to the domain $\underline{D}$ of matrices $\underline{a}$ whose characteristic values all belong to the domain of regularity Δ of $f(\zeta)$. Let $\delta_\alpha = |F-\alpha|$ be the distance of any characteristic value α of $\underline{a}$ to the frontier of regularity F of $f(\zeta)$. Denote by Δ_α the open circular domain with centre α and radius δ_α. Let $\underline{D}_\alpha$ be the domain of matrices $\underline{b}$ which commute with $\underline{a}$ and are such that Δ_α contains every characteristic value β of $\underline{b}$ which is latently associated with α. Then for any matrix $\underline{b}$ of $\underline{D}_\alpha$ we can take, for each α, a circle Γ_α with centre α which contained in Δ_α and contains in its interior all the characteristic values of $\underline{b}$ which are latently associated with α. Thus if α, β are any latently associated characteristic values of $\underline{a}$ and $\underline{b}$, and ζ any point of Γ_α, then $|\beta-\alpha| < |\zeta-\alpha|$. Now, by theorem 8.I, $f(\underline{a})$ has the expression

$$f(\underline{a}) = \frac{I}{2\pi i} \sum_\alpha \int_{\Gamma_\alpha} f_\alpha(\zeta)\, d\zeta \,(\zeta \underline{I} - \underline{a})^{-1}$$

where $f_\alpha(\zeta)$ denotes a single valued regular determination of $f(\zeta)$ in the neighbourhood of a characteristic value α. Let $\underline{R}_\alpha$, $\underline{S}_\beta$ be the partial units relative to $\underline{a}$ and $\underline{b}$ respectively, and consider the matrix $\underline{R}_\alpha \underline{S}_\beta \frac{\underline{b}-\underline{a}}{\zeta \underline{I}-\underline{a}}$, $\underline{T}_{\alpha\beta} = \underline{R}_\alpha \underline{S}_\beta$

when $\underline{E}_\alpha \underline{S}_\beta \neq \underline{0}$. For ζ ranging in Δ_α this matrix is bounded; and, by § II, has only one non-zero characteristic value $\frac{\beta-\alpha}{\zeta-\alpha}$ the modulus of which we have seen to be less than unity. Hence, by theorem IO, IV, $\frac{\underline{I}_{\alpha\beta} f_\alpha(\zeta)}{\zeta \underline{I}-\underline{b}}$ can be represented as the sum of a geometric series, which is uniformly convergent for ζ in Δ_α, and of which the mth partial sum is

$$\frac{\underline{T}_{\alpha\beta}}{\zeta \underline{I}-\underline{a}} \sum_{r<m} \left(\underline{T}_{\alpha\beta} \frac{\underline{b}-\underline{a}}{\zeta \underline{I}-\underline{a}}\right)^r f_\alpha(\zeta) = \underline{T}_{\alpha\beta} \left(\sum_{r<m} \frac{(\underline{b}-\underline{a})^r}{(\zeta \underline{I}-\underline{a})^{r+1}} \right) f_\alpha(\zeta),$$

since $\underline{T}^*_{\alpha\beta} = \underline{T}_{\alpha\beta}$. The corresponding elements of this matrix sequence of partial sums form a sequences of functions of ζ which are regular and uniformly convergent in any closed domain in Δ_α. Now the determination of $f(\underline{b})$ of the same branch as $f(\underline{a})$ is given by

$$f(\underline{b}) = \sum_{\alpha,\beta} f_{\alpha\beta}(\underline{b}),$$

where

$$f_{\alpha\beta}(\underline{b}) = \frac{\underline{T}_{\alpha\beta}}{2\pi i} \oint_{\Gamma_\alpha} \frac{f_\alpha(\zeta)}{\zeta \underline{I}-\underline{a}} d\zeta \quad ;$$

and, by theorem IO, III, each of the parts $f_{\alpha\beta}(\underline{b})$ can be expressed as the limit, as m tends to infinity, of the sequence of integrals

$$\oint_{\Gamma_\alpha} \frac{\underline{T}_{\alpha\beta}}{2\pi i} \left(\sum_{r<\infty} \frac{(\underline{b}-\underline{a})^r}{(\zeta \underline{I}-\underline{a})^{r+1}} \right) f_\alpha(\zeta) d\zeta$$

A theorem of Weierstrass is that the integrals of the terms of a uniformly convergent sequence of regular functions form a convergent sequence whose limit is the integral of their limit; and by this theorem, applied to the element integrals, we may replace the expression for $f_{\alpha\beta}(\underline{b})$ by

the expression

$$\sum_{r<\infty} (\underline{b}-\underline{a})^r \frac{\underline{T}_{\alpha\beta}}{2\pi i} \oint_{\Gamma_\alpha} \frac{f_\alpha(\zeta)\,d\zeta}{(\zeta\underline{I}-\underline{a})^{r+1}}$$

Now, since

$$f^{(r)}(\underline{a}) = \frac{r!}{2\pi i} \sum_{\alpha,\beta} \underline{T}_{\alpha\beta} \oint_{\Gamma_\alpha} \frac{f_\alpha(\zeta)\,d\zeta}{(\zeta\underline{I}-\underline{a})^{r+1}},$$

by §9, we have established the following result:

THEOREM. *If $f(\zeta)$ is regular at each characteristic value of $\underline{a}$, then*

$$\mathbf{f}(\underline{b}) = \sum_{r<\infty} \frac{I}{r!} f^{(r)}(\underline{a})(\underline{b}-\underline{a})^r$$

where $\underline{b}$ is any matrix which commutes with $\underline{a}$ and is such that each characteristic value of $\underline{a}$ is nearer to every latently associated characteristic value of $\underline{b}$ than the nearest singularity. [16]

13. Power series.

If $p(\zeta)$ is any polynomial, and $\underline{a}$ any matrix, by theorem 5.I, corollary, and theorem 8.I, we have [19]

$$p(\underline{a}) = \sum_{\alpha} \sum_{r<i_\alpha} \frac{I}{r!} p^{(r)}(\alpha)(\underline{a}-\alpha\underline{I})^r \underline{R}_\alpha ;$$

and since, by theorem 6.VI, the matrices $(\underline{a}-\alpha\underline{I})^r \underline{E}_\alpha (r<i_\alpha, \alpha)$ are linearly independent, we may conclude the following:

THEOREM I. *If $f_m(\zeta)$, $(m\to\infty)$, is a sequence of polynomials and $\underline{a}$ is any matrix, then the sequence $f_m(\underline{a})$, $(m\to\infty)$, is convergent if and only if the sequences $f_m^{(r)}(\alpha)$, $(m\to\infty)$, $(r<i_\alpha, \alpha)$, are convergent,* [17] *and then the limit is given by*

$$f_\infty(\underline{a}) = \sum_{\alpha} \underline{E}_\alpha \sum_{r<i_\alpha} \frac{I}{r!} f_\infty^{(r)}(\alpha)(\underline{a}-\alpha\underline{I})^r .$$

Any given power series $\mathcal{P}(\zeta)$ and its derived series $\mathcal{P}^{(r)}(\zeta)$ represents a regular function and its derivatives in the interior Δ of its circle of convergence. Thus the partial sums form a sequence of polynomials $f_m(\zeta)$ such that

$$f_m(\zeta) \to f(\zeta), \quad (m\to\infty),$$

for ζ in the interior Δ of the circle of convergence, where $f(\zeta)$ thus defined is a function regular in Δ, and

$$f_m^{(r)}(\zeta) \to f^{(r)}(\zeta), \quad (m\to\infty),$$

for the derivatives in Δ.

Let $\underline{a}$ be an the domain $\underline{D}$ of matrices whose characteristic values lie in Δ. Then the sequences $f_m^{(r)}(\alpha)$, $(m\to\infty)$, for all $r<i_\alpha$ and all α, are convergent:

$$f_\infty^{(r)}(\alpha) = f^{(r)}(\alpha), \quad (r<i_\alpha, \alpha).$$

Hence, by theorem I, the matrix sequence $f_m(\underline{a})$ is convergent, and the limit matrix is given by

$$f_\infty(\underline{a}) = \sum_\alpha \underline{R}_\alpha \sum_{r<i_\alpha} \frac{I}{r!} f^{(r)}(\alpha)(\underline{a}-\alpha\underline{I})^r .$$

Now, if Γ_α is any contour lying in Δ which contains α in its interior, then a formula of Cauchy gives

$$f^{(r)}(\alpha) = \frac{r!}{2\pi i} \oint_{\Gamma_\alpha} \frac{f(\zeta)d\zeta}{(\zeta-\alpha)^{r+1}} .$$

Let Γ_a be any contour lying in Δ and enclosing all the characteristic values α of $\underline{a}$ in its interior. Then, by taking $\Gamma_\alpha = \Gamma_a$ for each α, we have

$$f_\infty(\underline{a}) = \sum_\alpha \frac{R_\alpha}{2\pi i} \oint_{\Gamma_a} \left\{ \sum_{r<i_\alpha} \frac{(\underline{a}-\alpha\underline{I})^r}{(\zeta-\alpha)^{r+1}} \right\} f(\zeta)d\zeta .$$

But we have the identity

$$\sum_{r<i_\alpha} \frac{(\underline{a}-\alpha\underline{I})^r}{(\zeta-\alpha)^{r+1}} + \frac{(\underline{a}-\alpha\underline{I})^{i_\alpha}}{\zeta\underline{I}-\underline{a}} = \frac{\underline{I}}{\zeta\underline{I}-\underline{a}} ;$$

and since $(\underline{a}-\alpha\underline{I})^{i_\alpha}\underline{R}_\alpha = \underline{0}$, it follows that

$$f_\infty(\underline{a}) = \frac{\underline{I}}{2\pi i} \oint_{\Gamma_a} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}} ;$$

that is to say $f_\infty(\underline{a}) = f(\underline{a})$, where $f(\underline{a})$ is taken according to the definition given in § 8.

Now, by theorem 9.I, $f(\underline{a})$ is regular in $\underline{D}$, and has derivatives $f^{(r)}(\underline{a})$. But, by what we have now proved, $f_m(\underline{a})$ is convergent in $\underline{D}$, and has limit $f_\infty^{(r)}(\underline{a}) = f^{(r)}(a)$. Thus we may conclude the following:

THEOREM II. *If $\mathcal{P}(\zeta)$ is a power series, Δ the interior of its circle of convergence, $f(\zeta)$ its sum function, defined and regular in Δ, $\mathcal{P}^{(r)}(\zeta)$ its derived series,*

whose sums in Δ are the derivatives $f^{(n)}(s)$ of $f(s)$, and $\underline{D}$ the domain of matrices whose characteristic values all lie in Δ, then the power series $\mathcal{P}(\underline{a})$ and it derived series $\mathcal{P}^{(n)}(\underline{a})$ are convergent for all matrices $\underline{a}$ in $\underline{D}$, and their sums, representing a regular function in $\underline{D}$ and its derivatives, are given by

$$f(\underline{a}) = \frac{I}{2\pi i}\oint_{\Gamma_a} \frac{f(s)ds}{s\underline{I}-\underline{a}}$$

and

$$f^{(n)}(\underline{a}) = \frac{1}{2\pi i}\oint_{\Gamma_a} \frac{f^{(n)}(s)ds}{s\underline{I}-\underline{a}} = \frac{n!}{2\pi i}\oint_{\Gamma_a} \frac{f(s)ds}{(s\underline{I}-\underline{a})^{n+1}},$$

where Γ_a is any contour lying in Δ and enclosing all the characteristic values of $\underline{a}$ in its interior.

Notes

1. The upper and lower moduli $|\underline{a}|^*$ and $|\underline{a}|_*$, of any complex matrix $\underline{a}$, are defined as the non-negative real numbers whose squares are the maximum and minimum characteristic values, respectively, of the non-negative definite Hermitian matrix $\bar{\underline{a}}'\underline{a}$. Some of their properties are to be found in Rella(I) and Afriat(2); other properties which we assume here are easily proved. (124)

2. The resolvent is considered by various authors; see Turnbull and Aitken(3),160, et.seq. (124)

3. This formula is sometimes known as Hilbert's functional equation. (125)

4. Some accounts which have bearing here are to be found in Wedderburn(9), Hamburger and Grimshaw(5) and Halmos(6). (133)

5. c.f. Turnbull and Aitken(3), 1 (136)

6. Phillips(7) gives this proof. Alternatively, if $C(\alpha)=0$ there exists a vector $\underline{X}$ such that $\underline{aX}=\underline{X}\alpha$. Then $\underline{0}=\phi(\underline{a})\underline{X}=\underline{X}\phi(\alpha)$ so that $\phi(\alpha)=0$. (139)

7. cf. Wedderburn((9) (144)

8. See MacDuffee, where a survey is made of theories of functions of matrices. According to Macduffee, the definition

$$f(\underline{a}) = \frac{I}{2\pi i} \int_{\Gamma_a} \frac{f(\zeta)d\zeta}{\zeta\underline{I}-\underline{a}}$$

was suggested to Giorgi(10) in a letter from Cartan, it being taken that $f(\zeta)$ is a single value function whose domain of regularity contains the characteristic values of the matrix $\underline{a}$, and Γ_a lies in this domain and encloses these characteristic values in its interior. (146)

9. This formula has relation to the results of Pxantappie (16); also it connects our definition with the many-valued definition of Cipolla(11). (147)

10. Hamburger and Grimshaw, loc.cit, consider this formula as a definition, and show that it is satisfied by polynomials. (148)

11. With some bearing here, Ferrar(4),III, shows that, for a power series with sum $f(z)$, $f(\underline{z}+h\underline{I})-f(\underline{z})/h \to f'(\underline{z})$, $(h\to 0)$. (151)

12. Rella, loc.cit. (154)

13. This is a theorem of Oldenburger. The proof given by Oldenburger depends on the transformation of a to its classical canonical form. (154)

14. The characteristic values of commutative matrices $\underline{a},\underline{b}$ which are latently associated according to this definition can be shown to be those characteristic values α,β

which are such that there exists a vector X which is a latent vector of a corresponding to o and b corresponding to o, that is to say such that $\underline{a}\underline{X}=\underline{X}\alpha$ and $\underline{b}\underline{X}=\underline{X}\beta$. (156)

I5. Frobenius(I3) has proved a theorem, which has a common implication with ours, that the characteristic values of commutative matrices $\underline{a},\underline{b},\ldots$ can be arranged in sets $\alpha_\iota,\beta_\iota,\ldots$ such that the characteristic values of any rational function $\underline{f}=\rho(\underline{a},\underline{b},\ldots)$ are $\phi_\iota=\rho(\alpha_\iota,\beta_\iota,\ldots)$. (157)

I6. This result has been proved by Phillips in consideration of power series. (160)

I7. The theorem proved by Hensel(I4), that $\mathscr{P}(\underline{a})$ converges if and only if the series $\mathscr{P}^{(r)}(\alpha),(r<i_\alpha,\alpha)$, all converge, is included here. Weyr(I5) previously proved that $\mathscr{P}(\underline{a})$ converges if all the characteristic values of $\underline{a}$ lie in the circle of convergence of $\mathscr{P}(\delta)$. (161)

~~[8. cf. Ferrar, loc.cit.]~~

I8. For an independent proof, we have that t=0 is a root of multiplicity m_α of the equation

$$|(\underline{a}-\alpha\underline{I})-t\underline{I}| = \sum_{r=0}^{n} (-t)^{n-r} \operatorname{trace}(\underline{a}-\alpha\underline{I})^{(r)} .$$

Hence $\operatorname{trace}(\underline{a}-\alpha\underline{I})^{(n-m_\alpha)} \neq 0$; and so $(\underline{a}-\alpha\underline{I})^{(n-m_\alpha)} \neq \underline{0}$. Thus the matrix $\underline{a}-\alpha\underline{I}$ has rank at least $n-m_\alpha$, and the rank r_α of the space $\mathfrak{X}_\alpha$ of vectors $\underline{X}$ such that $(\underline{a}-\alpha\underline{I})\underline{X}=\underline{0}$ is at most m_α. An alternative procedure is to start from the formula

$$\left(\frac{\partial}{\partial\alpha}\right)^{k} |\underline{a}-\alpha\underline{I}| = (-1)^{k} k! \operatorname{trace}(\underline{a}-\alpha\underline{I})^{(n-k)} . \qquad (143)$$

I9. An independent proof is as follows: For any polynomial $p(\zeta)$, of degree d, and its derivatives $p^{(r)}(\zeta)$, there is the identity

$$p(\lambda) = \sum_{r<d} \frac{1}{r!} p^{(r)}(\mu)(\lambda-\mu) ,$$

by Taylor's theorem; so that

$$p(\underline{a})\underline{R}_\alpha = \sum_{r<i_\alpha} \frac{1}{r!} p^{(r)}(\alpha)(\underline{a}-\alpha\underline{I})^{r} \underline{R}_\alpha ,$$

since $(\underline{a}-\alpha\underline{I})^{r}\underline{R}_\alpha=\underline{0}$, $r>i_\alpha$; and then, since $\sum_\alpha \underline{R}_\alpha=\underline{I}$, summation with respect to α gives the required formula. (161)

REFERENCES

I. F. Rella, 'Uber den absoluten Betrag von Matrizen', Internat.Cong.of Mathematicians at Oslo(1936),II,29.

2. S.N. Afriat, 'Bounds for the characteristic values of matrix functions', Quart.J.Math.Oxford(2),2(1951), 81-4.

3. H.W. Turnbull and A.C. Aitken, An Introduction to the Theory of Canonical Matrices (London,1932).

4. W.L. Ferrar, Finite Matrices(Oxford,1951).

5. H.L. Hamburger and M.E. Grimshaw, Linear Transformations in n-dimensional Vector Space (Cambridge,1951).

6. P.R. Halmos, Finite Dimensional Vector Spaces (Annals of Mathematics Studies, Number 7; Princeton,1942).

7. H.B. Phillips, 'Functions of Matrices', Amer.J.Math., Vol.41(1919),266-278.

8. C.C. MacDuffee, Theory of Matrices (Ergebnisse der Mathematic und ihrer Grenzgebiete,Vol.2,part 5;Berlin 1933).

9. J.M.H. Wedderburn, Lectures on Matrices (American Math Soc.Colloquium Publ.,XVIII,1934).

10. G. Giorgi,Atti Accad.naz.Lincei,Rend.VI Vol.8(1928)3-9

11. M. Cipolla, Rend.Circ.mat.Palermo,Vol.56(1932),144-154

12. R. Oldenburger,'Infinite powers of matrices and characteristic roots', Duke Math.Journ.6(1940),257.

13. G. Frobenius, S.-B. preuß.Akad.Wiss.,1896 I,601-614.

I4. K. Hensel, J.reine angew.Math.,Vol.I5t(I926),I07-II0.

I5. E. Weyr, Bull.Sci.Math.II,Vol.ii(I887),205-2I5.

I6. L. Fantappie, C.R.Acad.,Paris, Vol.I86(I928),6i(-6II.

THE MODULI OF A COMPLEX MATRIX

I. Introduction.

For any complex matrix $\underline{a}$, the matrix $\bar{\underline{a}}'\underline{a}$ is non-negative definite Hermitian, so that its characteristic values, trace and determinant are all real and non-negative. The upper modulus $|\underline{a}|^*$ and the lower modulus $|\underline{a}|_*$ of $\underline{a}$ are defined as the non-negative real numbers whose squares are the maximum and the minimum characteristic values respectively of $\bar{\underline{a}}'\underline{a}$; and the absolute trace $|\underline{a}|_{(+)}$ and the absolute determinant $|\underline{a}|_{(\times)}$ are defined as the non-negative real numbers whose squares are the trace and determinant respectively of $\bar{\underline{a}}'\underline{a}$.

Five postulates P^*, to be satisfied by a real valued function defined over the complex matrices, are given by

$$P^*1 \quad \phi^*(k\underline{a}) = |k|\phi^*(\underline{a})$$

$$P^*2 \quad \phi^*(\underline{a}+\underline{b}) \leq \phi^*(\underline{a}) + \phi^*(\underline{b})$$

$$P^*3 \quad \phi^*(\underline{a}\underline{b}) \leq \phi^*(\underline{a})\phi^*(\underline{b})$$

$$P^*4 \quad \phi^*(\underline{e}_{ij}) = I$$

$$P^*5 \quad \phi^*(\bar{\underline{a}}'\underline{a}) = \phi^{*2}(\underline{a}) \ ,$$

where $\underline{e}_{ij}$ denotes the i,jth fundamental matrix, with i,jth element equal to unity and the rest of its elements zero.

Four other postulates P_* are given by

$$P_*1 \quad \phi_*(k\underline{a}) = |k|\phi_*(\underline{a})$$

$$P_*2 \quad \phi_*(\underline{a}\underline{b}) \geq \phi_*(\underline{a})\phi_*(\underline{b})$$

$$P_*3 \quad \phi_*(\underline{a}) > 0 \quad (|\bar{\underline{a}}'\underline{a}| > 0)$$

$$P_*4 \quad \phi_*(\bar{\underline{a}}'\underline{a}) = \phi_*^2(\underline{a}) \ .$$

The function $|\underline{a}|^*$ satisfies the postulates P^*, and if $\phi^*(\underline{a})$ is any other function which satisfies them then $|\underline{a}|^* \leq \phi^*(\underline{a})$; also $|\underline{a}|_*$ satisfies the postulates P_*, and if $\phi_*(\underline{a})$ is any other function which satisfies them, then $\phi_*(\underline{a}) \leq |\underline{a}|_*$. In this way it appears that the upper and lower moduli are the most fitting functions which satisfy the postulates P^*, P_* respectively for the role in which they are required.

The lower modulus may be constructed from the upper modulus by the relation

$$R.\quad |\underline{a}|_* = \begin{cases} (|\underline{a}^{-1}|^*)^{-1} & (|\underline{a}| \neq 0) \\ 0 & (|\underline{a}| = 0), \end{cases}$$

for when $\underline{a}$ is square, and then extended to rectangular matrices by the relation

$$E.\quad |\underline{a}|_*^2 = |\bar{\underline{a}}'\underline{a}|_* .$$

We have the following fundamental properties of the upper and lower moduli which are all as a consequence of their satisfying the postulates P^*, P_* respectively.

$$|\underline{a}|^* > 0 \ (\underline{a} \neq \underline{0}), \quad |\underline{0}|^* = \underline{0}.$$

$$|\underline{a}|_* > 0 \ (|\bar{\underline{a}}'\underline{a}| > 0), \quad |\underline{a}|_* = 0 \ (|\bar{\underline{a}}'\underline{a}| = 0)$$

$|\underline{a}|^* = |\bar{\underline{a}}'|^*$ for all $\underline{a}$; and if $|\underline{a}|_* > 0$ and $|\bar{\underline{a}}'|_* > 0$ then $|\underline{a}|_* = |\bar{\underline{a}}'|_*$.

$|\underline{k}|^* = |\underline{k}|_* = |k|$ for any scalar matrix $\underline{k}(=k\underline{I}$ where k is any scalar).

$|\underline{X}|^* = |\underline{X}|_* = |\underline{X}|$ for any vector $\underline{X}$ of length $|\underline{X}|$.

If $\underline{u}$ is any unitary matrix and $\underline{b} = \bar{\underline{u}}'\underline{a}\underline{u}$, then

$$|\underline{u}|^* = |\underline{u}|_* = I, \quad |\underline{a}|^* = |\underline{b}|^* \text{ and } |\underline{a}|_* = |\underline{b}|_*.$$

$$|\underline{a}|_* \leq |\alpha| \leq |\underline{a}|^*$$

for any matrix $\underline{a}$ with characteristic value α.

For any matrix $\underline{a}$ with elements a_{ij}

$$|a_{rs}| \leq |\underline{a}|^* \leq \sum_v |a_{ij}|$$

$$\left|\sum_k \bar{a}_{ki} a_{kj}\right|^{\frac{1}{2}} \leq |\underline{a}|^*, \text{ for all } i;j;$$

$$|\underline{a}|_* \leq \left(\sum_k |a_{ki}|^2\right)^{\frac{1}{2}} \leq |\underline{a}|^*, \text{ for all } i;$$

$$|\underline{a}-\underline{b}|^* \geq |(|\underline{a}|^* - |\underline{b}|^*)|$$

$$|\underline{a}|^* \to |\underline{b}|^* \quad (\underline{b} \to \underline{a}).$$

The upper and lower moduli, the absolute trace and absolute determinant satisfy the following relations, where n is the number of columns of the matrices:

$$n\,|\underline{a}|_* \leq |\underline{a}|_{(+)} \leq n\,|\underline{a}|^*$$

$$(|\underline{a}|_*)^n \leq |\underline{a}|_{(\times)} \leq (|\underline{a}|^*)^n$$

$|\underline{a}|^* \leq |\underline{a}|_{(+)}$; and if $|\underline{a}|^* < I$ then $|\underline{a}|_{(\times)} \leq |\underline{a}|_*$.

For the absolute trace and determinant, we have

$$|\underline{a}|^2_{(+)} = \sum_j \left(\sum_i |a_{ij}|^2\right), \quad |\underline{a}|^2_{(\times)} \leq \prod_j \left(\sum_i |a_{ij}|^2\right).$$

The absolute trace satisfies the inequalities

$$|\underline{ab}|_{(+)} \leq |\underline{a}|_{(+)}\,|\underline{b}|_{(+)}; \quad |\underline{a}+\underline{b}|_{(+)} \leq |\underline{a}|_{(+)} + |\underline{b}|_{(+)}.$$

The quantities studied here enter into various matrix processes. Rella has considered the five postulates P^* and has shown the first four to be independent. Bowker calls a function satisfying the first four of P^* a norm, and defines one norm as better than another if it is always less than the other. He asserts that no

best possible norm exists, though $\sum_{ij}|a_{ij}|$ represents the worst. Reiia defines a function which satisfies all five postulates P^* an <u>absolute value</u>, and we have that $|a|^*$ is the best possible absolute value, and $\sum_{ij}|a_{ij}|$ the worst.

Applications to numerical iteration methods are indicated.

2. Axiomatic definition.

The upper and lower moduli $|\underline{a}|^*$ and $|\underline{a}|_*$, of any complex matrix $\underline{a}$, are the non-negative real numbers whose squares are the maximum and minimum characteristic values, respectively, of the non-negative definite Hermitian matrix $\bar{\underline{a}}'\underline{a}$. It can be shown that the upper modulus is a function $\phi^*(\underline{a})=|\underline{a}|^*$ which satisfies the following postulates:

$$P^*1. \quad \phi^*(k\underline{a})=|k|\phi^*(\underline{a})$$

$$P^*2. \quad \phi^*(\underline{a}+\underline{b})\leq\phi^*(\underline{a})+\phi^*(\underline{b})$$

$$P^*3. \quad \phi^*(\underline{a}\underline{b})\leq\phi^*(\underline{a})\phi^*(\underline{b})$$

$$P^*4. \quad \phi^*(\underline{e}_{ij})=I$$

$$P^*5. \quad \phi^*(\bar{\underline{a}}'\underline{a})=\phi^{*2}(\underline{a}) \quad ,$$

where $\underline{e}_{ij}$ denotes the i,jth fundamental matrix, with i,jth element equal to unity, and the rest of its elements zero. Also, for the lower modulus $\phi_*(\underline{a})=|\underline{a}|_*$, we have

$$P_*1. \quad \phi_*(k\underline{a})=|k|\phi_*(\underline{a})$$

$$P_*2. \quad \phi_*(\underline{a}\underline{b})\leq\phi_*(\underline{a})\phi_*(\underline{b})$$

$$P_*3. \quad \phi_*(\underline{a})>0 \quad (|\underline{a}|\neq 0)$$

$$P_*4. \quad \phi_*(\underline{a}'\underline{a})=\phi_*^2(\underline{a}) \quad .$$

When $\underline{a}$ is a regular square matrix, $\phi^*(a)$ and $\phi_*(a)$ are connected by the relation

$$R. \quad \phi_*(\underline{a})=(\phi^*(\underline{a}^{-1}))^{-1} \quad .$$

For, since if $\underline{h}$ is positive definite Hermitian, and Θ is its minimum characteristic value, then Θ^{-1} is the maximum characteristic value of $\underline{h}^{-1}$, that is $|\underline{h}|_*=|\underline{h}^{-1}|^{*-1}$, we have

$$|\underline{a}|_*^2 = |\bar{\underline{a}}'\underline{a}|_* = (\,|(\bar{\underline{a}}'\underline{a})^{-1}|^*)^{-1} = (\,|\underline{a}^{-1}(\overline{\underline{a}^{-1}})'|^*)^{-1} = (\,|(\overline{\underline{a}^{-1}})'|^*)^{-2}.$$

A consequence of the postulates P^* is that $\phi^*(\bar{\underline{a}}') = \phi^*(\underline{a})$. Hence we have $|\underline{a}|_* = (\,|\underline{a}^{-1}|^*)^{-1}$, which is the relation R.

First we shall deduce the consequences of the postulates P^*, applied to any function $\phi^*(\underline{a})$. Thus, by P^*I, we have $\phi^*(\underline{o}) = 0$ and $\phi^*(-\underline{a}) = \phi^*(\underline{a})$; and then, by P^*2,

$$0 = \phi^*(\underline{o}) \leq \phi^*(\underline{a}) + \phi^*(-\underline{a}) = 2\phi^*(\underline{a}), \text{ so that we}$$

have

$$\text{(i)} \quad \phi^*(\underline{a}) \geq 0, \quad \phi^*(\underline{o}) = 0.$$

Now if $\underline{X}$ is any vector, and $|\underline{X}|$ is its length, then $\bar{\underline{X}}'\underline{X}$ is a matrix with a single element, the value of which is $|\underline{X}|^2$. Hence, by P^*5,I and 4 successively, we have

$$(\phi^*(\underline{X}))^2 = \phi^*(\bar{\underline{X}}'\underline{X}) = |\underline{X}|^2 \phi^*(I) = |\underline{X}|^2;$$

and this, by (i), gives

$$\text{(ii)} \quad \phi^*(\underline{X}) = |\underline{X}|, \text{ for any vector } \underline{X}.$$

By P^*5 and 3,

$$\phi^{*2}(\underline{a}) = \phi^*(\bar{\underline{a}}'\underline{a}) \leq \phi^*(\bar{\underline{a}}')\phi^*(\underline{a}), \quad \phi^{*2}(\bar{\underline{a}}') = \phi^*(\underline{a}\bar{\underline{a}}') \leq \phi^*(\underline{a})\phi^*(\bar{\underline{a}}')$$

which show that $\phi^*(\underline{a})$ and $\phi^*(\bar{\underline{a}}')$ are zero together. When they are not zero, we get

$$\phi^*(\underline{a}) \leq \phi^*(\bar{\underline{a}}'), \quad \phi^*(\bar{\underline{a}}') \leq \phi^*(\underline{a}).$$

That is :

$$\text{(iii)} \quad \phi^*(\underline{a}) = \phi^*(\bar{\underline{a}}').$$

From the relations

$$\underline{a} = \sum_{ij} a_{ij}\underline{e}_{ij} \qquad a_{rs} = \underline{e}_r'\underline{a}\underline{e}_s,$$

where $\underline{e}_i$ is the ith fundamental unit vector, with ith element unity and the rest zero, by applying P^*I, to, 4, we have

$$\phi^*(\underline{a}) \leq \sum_{ij} \phi^*(\underline{a}\, e_{ij}) = \sum_{ij} |a_{ij}| \phi^*(e_{ij}) = \sum_{ij} |a_{ij}|,$$

and

$$|a_{rs}| = \phi^*(a_{rs}) = \phi^*(e_r' a e_s) \leq \phi^*(e_r') \phi^*(\underline{a}) \phi^*(e_s) = \phi^*(\underline{a});$$

which give

$$(v) \qquad |\underline{a}_{rs}| \leq \phi^*(\underline{a}) \leq \sum_{ij} |a_{ij}| \ .$$

A corollary to this proposition is

$$(v) \quad \phi^*(\underline{a}) > 0 \ (\underline{a} \neq \underline{0}), \quad \phi^*(\underline{a}) \to 0 \ (\underline{a} \to \underline{0}).$$

Now, by P^*5, $\phi^{*2}(\underline{I}) = \phi^*(\underline{I})$, and, with (v), this gives

$$(vi) \quad \phi^*(\underline{I}) = I.$$

Hence, for any regular square matrix $\underline{a}$, we have

$$I = \phi^*(\underline{I}) = \phi^*(\underline{a}\bar{\underline{a}}^{-1}) \leq \phi^*(\underline{a}) \phi^*(\underline{a}^{-1});$$

and, accordingly,

$$(vii) \quad \phi^*(\underline{a}) \geq (\phi^*(\underline{a}^{-1}))^{-1} \ .$$

By P^*2,

$$\phi^*(\underline{a}) = \phi^*(\underline{a} - \underline{b} + \underline{b}) \leq \phi^*(\underline{a} - \underline{b}) + \phi^*(\underline{b});$$

that is,

$$(viii) \quad \phi^*(\underline{a} - \underline{b}) \geq |\phi^*(\underline{a}) - \phi^*(\underline{b})| \ .$$

Taken with (v), this gives

$$(ix) \quad \phi^*(\underline{b}) \to \phi^*(\underline{a}) \quad (\underline{b} \to \underline{a}).$$

The application of the first part of (iv) to $\bar{\underline{a}}'\underline{a}$ gives $(\bar{\underline{a}}'\underline{a})_{ij} \leq \phi^*(\bar{\underline{a}}'\underline{a})$. Let $\bar{\underline{a}}_{(i}'$ denote the ith row of $\bar{\underline{a}}'$, and $\underline{a}_{j)}$ the jth column of $\underline{a}$, so that $\bar{\underline{a}}_{(j}' = (\overline{\underline{a}_{j)}})'$. Then, using P^*5, we have $|\bar{\underline{a}}_{(i}' \underline{a}_{j)}| \leq \phi^{*2}(\underline{a})$; that is

$$(x) \quad \Big| \sum_{k} \bar{a}_{k i} a_{k j} \Big|^{1/2} \leq \phi^*(\underline{a}), \text{ for all } i, j.$$

A corollary to this is $|\underline{a}_{ij}| \leq \phi^*(\underline{a})$; or

$$(xi) \quad \Big(\sum_{k} |a_{ki}|^2 \Big)^{1/2} \leq \phi^*(\underline{a}), \text{ for all } i.$$

This states that $\phi^*(\underline{a})$ is at least the length of each column of $\underline{a}$. By Cauchy's inequality, this again implies (viii); so that (viii) and (ix) are equivalent. Let $\underline{a}$ be square, and α any of its characteristic values; and let $\underline{X}$ be a corresponding axial vector, so that $a\underline{X}=\underline{X}\alpha$. Then, by P^*I and 3, and by (ii), we have

$$|\alpha|\,|\underline{X}|=\phi^*(\underline{X}\alpha)=\phi^*(\underline{a}\underline{X})\leq\phi^*(\underline{a})\phi^*(\underline{X})=\phi^*(\underline{a})\,|\underline{X}|\ ,$$

or

(xii) $|\alpha|\leq\phi^*(\underline{a})$, for any square matrix $\underline{a}$ with characteristic value α.

By (x) applied to $\bar{\underline{a}}'\underline{a}$ we have $|\underline{a}|^2\leq\phi^*(\bar{\underline{a}}'\underline{a})=\phi^*(\underline{a})^2$, using P^*5. Hence we have

(xiii) $|\underline{a}|^*\leq\phi^*(\underline{a})$.

Now since $|\underline{a}|^*$satisfies the postulates P^* , we have $|\alpha|\leq|\underline{a}|^*$. Hence, combining (x) and (xi), we have

(xiv) $|\alpha|\leq|\underline{a}|^*\leq\phi^*(\underline{a})$.

From (iv) we see that $|\underline{a}|^*$, among the functions $\phi^*(\underline{a})$ which satisfy the postulates P^*, gives the smallest upper bound for the moduli of the characteristic values of $\underline{a}$, and hence it is the best of these functions for the role we are considering. Now, finally, we show

(v) $\phi^*(\underline{u})=I$ and $\phi^*(\bar{\underline{u}}'\underline{a}\underline{u})=\phi^*(\underline{a})$ if $\bar{\underline{u}}'\underline{u}=\underline{u}\bar{\underline{u}}'=\underline{1}$

Thus, by P^*5, $\phi^*(\underline{u})=\phi^*(\bar{\underline{u}}'\underline{u})=\phi^*(\underline{I})=I$; and hence, by (v), we have $\phi^*(\underline{u})=\phi^*(\bar{\underline{u}}')=\underline{I}$. Further, if $\underline{b}=\bar{\underline{u}}'\underline{a}\underline{u}$, so that $\underline{a}=\underline{u}\underline{b}\bar{\underline{u}}'$, we have $\phi^*(\underline{b})=\phi^*(\bar{\underline{u}}'\underline{a}\underline{u})\leq\phi^*(\bar{\underline{u}}')\phi^*(\underline{a})\phi^*(\underline{u})=\phi^*(\underline{a})$; hence $\phi^*(\underline{a})=\phi^*(\underline{b})$. Our results may be stated as follows:

THEOREM I. If $\phi^*(\underline{a})$ satisfies the postulates P^*, then

T I. $\phi^*(\underline{a})>0$ $(\underline{a}\neq 0)$, $\phi^*(\underline{0})=0$

T 2. $\phi^*(\underline{a}-\underline{b})\geq|\phi^*(\underline{a})-\phi^*(\underline{b})|$

T 3. $\phi^*(\underline{b})\rightarrow\phi^*(\underline{a})$ $(\underline{b}\rightarrow\underline{a})$

T 4. $\phi^*(\underline{X})=|\underline{X}|$, for any vector $\underline{X}$

T 5. $\phi^*(\underline{k})=|k|$, for any scalar matrix $\underline{k}=\underline{k}\underline{I}$, where k is any scalar

T 6. $\phi^*(\underline{a})=\phi^*(\underline{\bar{a}}')$

T 7. $\phi^*(\underline{a})\geq(\phi^*(\underline{a}^{-1}))^{-1}$, for any regular square matrix $\underline{a}$

T 8. $|a_{rs}|\leq\phi^*(\underline{a})\leq\sum_{ij}|a_{ij}|$, for any matrix $\underline{a}$, with elements a_{ij}.

T 9. $|\alpha|\leq\phi^*(\underline{a})$, for any square matrix $\underline{a}$ with α for characteristic value α.

T* IO. $|\underline{a}|^*\leq\phi^*(\underline{a})$

T II. $|\sum_{k}\bar{a}_{k_1}a_{k_2}|^{\frac{1}{2}}\leq\phi^*(a)$, for all i,j

T I2. $\phi^*(\underline{u})=I$ and $\phi^*(\underline{\bar{u}}'\underline{a}\underline{u})=\phi^*(\underline{a})$ if $\underline{\bar{u}}'\underline{u}=\underline{u}\underline{\bar{u}}'=\underline{I}$

$|a|_* \leq \sum|a_{ij}|$.

We shall now deduce the consequence of the postulates P^* applied to a function $\phi^*(a)$. Thus, if $\underline{X}$ is any vector, then the successive application of P^* 4, I and 3 gives

$$\phi_*^2(\underline{X})=\phi_*(\underline{\bar{X}}'\underline{X})=|\underline{X}|^2\phi_*(I)=|\underline{X}|^2;$$ or

(i) $\phi_*(\underline{X})=|\underline{X}|$, for any vector $\underline{X}$.

If $\underline{a}$ is any square matrix, with α as characteristic value and $\underline{X}$ as a corresponding axial vector, so that $\underline{a}\underline{X}=\underline{X}\alpha$, then, by (i) and P^* I and 2, we have

$$|\alpha||\underline{X}| = \phi_*(\underline{X}\alpha) = \phi_*(\underline{a}\underline{X}) \geq \phi_*(\underline{a})\phi_*(\underline{X}) = \phi_*(\underline{a})|\underline{X}|,$$

so that

$$\text{(ii)}\quad \phi_*(\underline{a}) \leq |\alpha| .$$

As a corollary to (ii), it follows that

$$\text{(iii)}\quad \phi_*(\underline{a}) = 0 \text{ if } |\underline{a}| = 0, \text{ for any square matrix } \underline{a}.$$

Now (iii) may be extended to rectangular matrices by application of the relation P_*4, to give

$$\text{(iv)}\quad \phi_*(\underline{a}) = 0 \text{ if } |\bar{\underline{a}}'\underline{a}| = 0.$$

Also, taking (ii) together with P_*4, we get

$$\text{(v)}\quad \phi_*(\underline{a}) \leq |\underline{a}|_* .$$

By P_*4 and 2,

$$\phi_*^2(\underline{a}) = \phi_*(\bar{\underline{a}}'\underline{a}) \geq \phi_*(\bar{\underline{a}}')\phi_*(\underline{a}),$$

$$\phi_*^2(\bar{\underline{a}}') = \phi_*(\underline{a}\bar{\underline{a}}') \geq \phi_*(\underline{a}')\phi_*(\underline{a});$$

and hence

$$\phi_*(\underline{a}) = \phi_*(\bar{\underline{a}}') \text{ if } \phi_*(\underline{a}) \neq 0 \text{ and } \phi_*(\bar{\underline{a}}') \neq 0.$$

Now if $\bar{\underline{u}}'\underline{u} = \underline{u}\bar{\underline{u}}' = I$, then, using P_*3 and 4, we have $I = \phi_*(\underline{u}'u) = \phi_*^2(\underline{u})$, so that $\phi_*(\underline{u}) = I$, and also $\phi_*(\bar{\underline{u}}') = I$. If $\underline{b} = \bar{\underline{u}}'\underline{a}\underline{u}$, then, using P_*2, we get $\phi_*(\underline{b}) = \phi_*(a)$. Thus we have the following results:

THEOREM II. If $\phi_*(a)$ satisfies the postulates P_*, then

T I. $\phi_*(\underline{a}) = 0$ if $|\bar{\underline{a}}'\underline{a}| = 0$

T 2. $\phi_*(\underline{X}) = |\underline{X}|$, for any vector $\underline{X}$

T 3. $\phi_*(\underline{a}) = \phi_*(\bar{\underline{a}}')$ if $\phi_*(\underline{a}) \neq 0$ and $\phi_*(\bar{\underline{a}}') \neq 0$

T 4. $\phi_*(\underline{a}) \leq |\alpha|$, for any matrix $\underline{a}$ with a characteristic value α

T 5. $\phi(u) = \phi(u')$

T 5. $\phi_*(\underline{u})=\phi(\underline{\bar{u}}')=I$ and $\phi_*(\underline{\bar{u}}'\underline{a}\underline{u})=\phi_*(\underline{a})$ if $\underline{\bar{u}}'\underline{u}=\underline{u}\underline{\bar{u}}'=\underline{I}$

T 6. $\phi_*(\underline{a})\leq|\underline{a}|_*$.

Now the following enquiry may be made. Suppose we have a function $\phi^*(\underline{a})$ to which the five postulates P^* apply. We may then introduce a function $\phi_*(\underline{a})$, given for square matrices $\underline{a}$ by the following definition

$$D.\quad \phi_*(\underline{a}) = \begin{cases} (\phi(\underline{a}^{-1}))^{-1} & (|\underline{a}|\neq 0) \\ 0 & (|\underline{a}|=0) \end{cases},$$

and then extended to rectangular matrices by use of the relation

$$E.\quad \phi_*(\underline{a})=\phi_*(\underline{\bar{a}}'\underline{a}).$$

The upper and lower moduli, $|\underline{a}|^*$ and $|\underline{a}|_*$ are connected in this way, by which the continuity of $|\underline{a}|_*$ follows from that of $|\underline{a}|^*$. We ask if any function $\phi_*(\underline{a})$, thus defined, necessarily satisfies the four postulates P_*? It certainly satisfies P_* I,3 and 4; but the question is that if it satisfies P^*2 also, then are the functions $\phi^*(\underline{a})$ and $\phi_*(\underline{a})$ uniquely defined as $|\underline{a}|^*$ and $|\underline{a}|_*$ respectively? However, we cannot answer this question. In conclusion, for purposes of reference, we shall set out the fundamental results already established for $|\underline{a}|^*$ and $|\underline{a}|_*$.

THEOREM III. The upper and lower moduli $|\underline{a}|^*$ and $|\underline{a}|_*$ are non-negative real valued functions, continuous in the complex matrix argument $\underline{a}$, which satisfy the postulates P^* and P_* respectively, and are such that

$|\underline{a}|^* > 0$ $(\underline{a} \neq 0)$; $|0|^* = 0$;

$|\underline{a}|_* > 0$ $(|\bar{\underline{a}}'\underline{a}| \neq 0)$; $|\underline{a}|_* = 0$ $(|\bar{\underline{a}}'\underline{a}| = 0)$;

$|\underline{a}|^* = |\bar{\underline{a}}'|^*$; $|\underline{a}|_* = |\bar{\underline{a}}'|_*$ if $|\underline{a}|_* \neq 0$ and $|\bar{\underline{a}}'|_* \neq 0$;

$|\underline{a}|_* = (|(\bar{\underline{a}}'\underline{a})^{-1}|^*)^{-\frac{1}{2}}$; $|\underline{a}|_* = (|\underline{a}^{-1}|^*)^{-1}$ when $\underline{a}$ is square

$|\underline{k}|^* = |\underline{k}|_* = |k|$ for any scalar matrix $\underline{k}$ $(= k\underline{I}$ where k is any scalar$)$

$|\underline{X}|_* = |\underline{X}|^* = |\underline{X}|$, for any matrix $\underline{X}$, of length $|\underline{X}|$

$|\underline{u}|_* = |\underline{u}|^* = \underline{I}$, $|\bar{\underline{u}}'\underline{a}\underline{u}|^* = |\underline{a}|^*$ and $|\bar{\underline{u}}'\underline{a}\underline{u}|_* = |\underline{a}|_*$, for any unitary matrix $\underline{u}$, such that $\bar{\underline{u}}'\underline{u} = \underline{u}\bar{\underline{u}}' = \underline{I}$.

$|\underline{a}|_* \leq |\alpha| \leq |\underline{a}|^*$, for any matrix $\underline{a}$ with characteristic value α

$|a_{rs}| \leq |\underline{a}|^* \leq \sum_{ij} |a_{ij}|$, for any matrix $\underline{a}$ with elements a_{ij}.

$\left|\sum_{k} \bar{a}_{k_i} a_{k_j}\right|^{\frac{1}{2}} \leq |\underline{a}|^*$, for all i,j.

$|\underline{a}|_* \leq \sum_{k} |a_{kt}|^2 \leq |\underline{a}|^*$

3. Absolute trace and absolute determinant.

If $\underline{a}$ is a rectangular complex matrix, say of order mxn, then $\bar{\underline{a}}'\underline{a}$ is a square matrix, of order n, which is non-negative definite Hermitian, so that its characteristic values are all real and non-negative. The determinant of $\bar{\underline{a}}'\underline{a}$ is the sum of the squares of the moduli of the $\binom{n}{m}$ determinants on the columns of $\underline{a}$. If $m \geq n$, and a is of full rank n, then $\bar{\underline{a}}'\underline{a}$ is positive definite, and has positive determinant; otherwise the determinant of $\bar{\underline{a}}'\underline{a}$ id zero. The trace of $\bar{\underline{a}}'\underline{a}$ is the sum odf the squares of the moduli of the elements of $\underline{a}$. The absolute trace $|\underline{a}|_{(+)}$, and the absolute determinant $|\underline{a}|_{(\times)}$, of any complex matrix $\underline{a}$, are defined as the non-negative real numbers whose squares are equalt to the trace and the determinant, respectively, of the matrix $\bar{\underline{a}}'\underline{a}$. We note that the length of any vector $\underline{X}$ is given by

$$|\underline{X}| = |\underline{X}|^* = |\underline{X}|_* = |\underline{X}|_{(+)} = |\underline{X}|_{(\times)}.$$

Also there are the following relations:

THEOREM.I. If $\underline{a}$ is any complex matrix, and n is the number of its columns, then

$$n\,|\underline{a}|_* \leq |\underline{a}|_{(+)} \leq n\,|\underline{a}|^*,$$

$$(|\underline{a}|_*)^n \leq |\underline{a}|_{(\times)} \leq (|\underline{a}|^*)^n,$$

$$|\underline{a}|^* \leq |\underline{a}|_{(+)} \; ; \quad \text{and if } |\underline{a}|^* \leq I, \text{ then } |\underline{a}|_{(\times)} \leq |\underline{a}|_*.$$

By inspection of the coefficients in the characteristic equation of $\bar{\underline{a}}'\underline{a}$ we have that if $h_1,\ldots,h_n$ are the characteristic values of $\bar{\underline{a}}'\underline{a}$ in ascending order of magnitude, so that

$$0 \leq |\underline{a}|_*^2 = h_1 \leq \ldots \leq h_n = |\underline{a}|_*^2 ,$$

then

$$h_1 + \ldots + h_n = |\underline{a}|_{(4)},$$

$$h_1 \ldots h_n = |\underline{a}|_{(n)};$$

and the considered inequalities are equivalent to the following ones:

$$nh_1 \leq h_1 + \ldots + h_n \leq nh_n,$$

$$h_1^n \leq h_1 \ldots h_n \leq h_n^n ,$$

and

$$h_n \leq h_1 + \ldots + h_n ; \text{ and if } h_n \leq 1, \text{ then } h_1 \ldots h_n \leq h_1 .$$

THEOREM II. $$|\underline{a}|_{(4)}^2 = \sum_{\nu=1}^{n} |\underline{a}_{(\nu)}|^2 , \quad |\underline{a}|_{(n)}^2 \leq \prod_{\nu=1}^{n} |\underline{a}_{(\nu)}|^2 .$$

The first statement is immediate from definition. The second generalizes a theorem of Hadamard on the upper bound of the modulus of a determinant. It follows from the proposition, applied to $\bar{\underline{a}}'\underline{a}$, that the determinant of a positive definite Hermitian matrix is at most the product of its diagonal elements.

THEOREM III. $$|\underline{ab}|_{(4)} \leq |\underline{a}|_{(4)} |\underline{b}|_{(4)} , \quad |\underline{a}+\underline{b}|_{(4)} \leq |\underline{a}|_{(4)} + |\underline{b}|_{(4)} .$$

These results have been proved by Wedderburn. We may establish them as follows:

$$|\underline{ab}|_{(4)}^2 = \text{trace}(\bar{\underline{b}}'\bar{\underline{a}}'\underline{ab}) = \sum_k \bar{\underline{b}}_{(k)}' \bar{\underline{a}}' \underline{ab}_{(k)} \leq \sum_k |\underline{a}|_*^2 |\underline{b}_{(k)}|^2 = |\underline{a}|_*^2 \sum |\underline{b}_{(k)}|$$

then the first inequality follows, since $|\underline{a}|^{*} \leq |\underline{a}|_{(4)}$, by theorem I. Also, by two applications of Cauchy's inequality, we have the relations

$$
\begin{aligned}
|a+b|_{(4)}^{2} &= \operatorname{trace}((\bar{\underline{a}}'+\bar{\underline{b}}')(\underline{a}+\underline{b})) = \operatorname{trace}(\bar{\underline{a}}'\underline{a}+\bar{\underline{a}}'\underline{b}+\bar{\underline{b}}'\underline{a}+\bar{\underline{b}}'\underline{b}) \\
&= \sum_{\nu}(\bar{\underline{a}}_{(\nu)}'\underline{a}_{(\nu)}+\bar{\underline{a}}_{(\nu)}'\underline{b}_{(\nu)}+\bar{\underline{b}}_{(\nu)}'\underline{a}_{(\nu)}+\bar{\underline{b}}_{(\nu)}'\underline{b}_{(\nu)}) \\
&\leq \sum_{\nu}(\bar{\underline{a}}_{(\nu)}'\underline{a}_{(\nu)}+2(\bar{\underline{a}}_{(\nu)}'\underline{a}_{(\nu)})^{1/2}(\bar{\underline{b}}_{(\nu)}'\underline{b}_{(\nu)})^{1/2}+\bar{\underline{b}}_{(\nu)}'\underline{b}_{(\nu)}) \\
&= \sum_{\nu}|a_{(\nu)}|^{2}+2\sum_{\nu}|\underline{a}_{(\nu)}||\underline{b}_{(\nu)}|+\sum_{\nu}|\underline{b}_{(\nu)}|^{2} \\
&\leq \sum_{\nu}|a_{(\nu)}|^{2}+2\Big(\sum_{\nu}|\underline{a}_{(\nu)}|^{2}\Big)^{1/2}\Big(\sum_{\nu}|\underline{b}_{(\nu)}|^{2}\Big)^{1/2}+\sum_{\nu}|\underline{b}_{(\nu)}|^{2} \\
&= |\underline{a}|_{(4)}^{2}+2|\underline{a}|_{(4)}|\underline{b}|_{(4)}+|\underline{b}|_{(4)}^{2} \\
&= (|\underline{a}|_{(4)}+|\underline{b}|_{(4)})^{2},
\end{aligned}
$$

as required.

4. The condition $|\underline{a}|^* \leq k$.

The condition $|\underline{a}|^* < I$ arises as the general criterion for the convergence of certain iterative processes. Since $|\underline{a}|^* \leq |\underline{a}|_{(+)}$, it is a weaker condition than $|\underline{a}|_{(+)} < I$, which appears sometimes to have been introduced; but it is stronger than the condition $\rho(\underline{a}) < I$, where $\rho(a)$, called the characteristic radius of $\underline{a}$, denotes the maximum of the moduli of the characteristic values of $\underline{a}$; since we have $|\alpha| \leq |\underline{a}|^*$ for each characteristic value and hence $\rho(a) \leq |\underline{a}|^*$. In some numerical iterations there is a matrix or vector whose elements represent errors after a finite number n of steps, and it has the form $\underline{a}^n \underline{u}$, where $\underline{u}$ is some initial matrix or vector. Then the fitting general convergence condition is $\rho(a) < I$, since $\underline{a}^n \to 0$ $(n \to \infty)$, if and only if $\rho(\underline{a}) < I$, by a theorem of Oldenburger.[1] In other processes the error has the form $\underline{a}_m \ldots \underline{a}_1 \underline{u}$, where $\underline{a}_1, \underline{a}_2, \ldots$ is the sequence of matrix multipliers, where the nth is determined at the (n-I)th step of the process. In this case the iteration will converge, or equivalently $\underline{a}_m \ldots \underline{a}_1 \underline{u} \to 0$ $(n \to \infty)$ if, for example, we have $|\underline{a}_n| \leq k < I$ for all n, or at least for all n after a certain value. Thus the condition $|\underline{a}|^* < k$ has some fundamental importance in practical applications, and we shall express it in a form which is more convenient for computation, which demands nothing further than the elementary calculation of determinants, and thus saves us

from the generally difficult task of the calculation of characteristic values. It may be noted, however, that, in the calculation of the moduli, there is not the difficulty associated with the general determination of characteristic values; for there are special methods available for the calculation of characteristic values of maximum or minimum modulus, which in this case would apply.

An equivalent of the condition $|\underline{a}|^* < k$ is that for every vector $\underline{X}$ the vectors $\underline{aX}$ and $\underline{X}$ have their lengths in a ratio of almost k. This is identical with the condition of the Hermitian form $\bar{\underline{X}}'(k^2\underline{I}-\bar{\underline{a}}'\underline{a})\underline{X}$ being positive definite, for which, if $\mathbf{A}_{\prime\prime}^{(m)}$ denotes the leading mth order minor of any matrix $\mathbf{A}$, a necessary and sufficient condition may be written as

$$(k^2\underline{I}-\bar{\underline{a}}'\underline{a})_{\prime\prime}^{(m)} > 0 \quad (m=I,\ldots,n).$$

Now if $\underline{\mathbf{A}}_{\prime}^{(m}$ and $\mathbf{A}_{\prime}^{(m)}$ denote the submatrices formed of the leading m rows and the leading m columns respectively of $\underline{\mathbf{A}}$, and $\underline{\mathbf{A}}_{\prime\prime}^{(m\,m)}$ is the leading mxm submatrix, then we have

$$\begin{aligned}(k^2\underline{I}-\bar{\underline{a}}'\underline{a})_{\prime\prime}^{(m)} &= |(k^2\underline{I}-\bar{\underline{a}}'\underline{a})_{\prime\prime}^{(m\,m)}| \\ &= |k^2\underline{I}-\bar{\underline{a}}_{\prime}^{\prime(m}\underline{a}_{\prime}^{m)}| \\ &= \sum_{t=0}^{m}(-I)^t k^{2(m-t)}\,\mathrm{trace}(\bar{\underline{a}}_{\prime}^{\prime(m}\mathbf{a}_{\prime}^{m)})^{(t)} \\ &= \sum_{t=0}^{m}(-I)^t k^{2(m-t)}\,\mathrm{trace}[(\underline{a}_{\prime}^{\prime(m})^{(t)}(\underline{a}_{\prime}^{m)})^{(t)}] \\ &= \sum_{t=0}^{n}(-I)^t k^{2(n-t)}\,|(\underline{a}_{\prime}^{m)})^{(t)}|_{(t)}^2 \quad ,\end{aligned}$$

using the multiplication theorem for compound matrices, where $|(\underline{a}_{\prime}^{m)})^{(t)}|_{(t)}$ is in fact the sum of the squares of the moduli of the rth order determinants on the leading m

columns of $\underset{\sim}{a}$, and $(a_{,}^{(m-1)})^{(+)}=(a^{(+)})_{,}^{(m+)})$. Thus a necessary and sufficient condition that $|\underset{\sim}{a}|^{*}<k$, involving only the subdeterminants of $\underset{\sim}{a}$, is given as follows:

THEOREM. *If $\underset{\sim}{a}$ is any rectangular complex matrix with n columns, then a necessary and sufficient condition that $|\underset{\sim}{a}|^{*}<k$ is that*

$$k^{2m}-k^{2(m-1)}\left|\underset{\sim}{a}_{,}^{(m)}\right|_{(+)}^{2}+k^{2(m-2)}\left|(a^{(2)})_{,}^{(m-2)}\right|_{(+)}^{2}-\ldots+(-1)^{m}\left|\underset{\sim}{a}_{,,}^{(m)}\right|_{(+)}^{2}>0 \quad (m=1,\ldots,n)$$

COROLLARY. *$|\underset{\sim}{a}|^{*}<1$ if and only if*

$$\left|\underset{\sim}{a}_{,,}\right|_{(+)}^{2}<1$$

$$\left|\underset{\sim}{a}_{,}^{(2)}\right|_{(+)}^{2}-\left|\underset{\sim}{a}_{,,}^{(2)}\right|^{2}<1$$

$$\left|\underset{\sim}{a}_{,}^{(3)}\right|_{(+)}^{2}-\left|(\underset{\sim}{a}^{(2)})_{,}^{(3-2)}\right|^{2}+\left|\underset{\sim}{a}_{,,}^{(3)}\right|^{2}<1$$

$$\cdots\cdots \qquad \cdots\cdots \qquad \cdots\cdots$$

$$\left|\underset{\sim}{a}\right|_{(+)}-\left|(\underset{\sim}{a}^{(2)})_{,}^{(m-2)}\right|_{(+)}-\ldots+(-1)^{n}\left|(\,|\underset{\sim}{a}|\,)\right|_{(+)}^{2}<1$$

In some circumstances it may be more convenient to employ the conditions in this form, rather than evaluate $|\underset{\sim}{a}|^{*}$ directly; for example when $\underset{\sim}{a}$ is of a low order, and its elements are functions of one or more parameters ^{i}n which case it may be possible to maximise the functions of the elements entering into the present form of the conditions, and thus find the region of the parameter space in which $|\underset{\sim}{a}|^{*}<k$. Also it is possible that an application of one of the inequalities given in the theorem of the previous paragraph may prove decisive.

Suppose we have a system of equations arranged in the form

$$x_1 = f_1(x_1, \ldots x_n)$$
$$\ldots\ldots\ldots\ldots$$
$$x_n = f_n(x_1, \ldots, x_n) ,$$

and to be solved simultaneously for the unknowns $x_1, \ldots, x_n$. We shallw write $\underset{\sim}{x} = \underset{\sim}{f}\underset{\sim}{x}$ for these equations, so that $\underset{\sim}{f}$ represents a mapping of the complex n-vector space into itself, and the solutions are just the points of the space which are invariant under $\underset{\sim}{f}$. Denoting by $\partial\underset{\sim}{f}/\partial\underset{\sim}{x}$ the functional matrix $\partial(f_1, \ldots, f_n)/\partial(x_1, \ldots, x_n)$, with i,jth element $\partial f_i/\partial x_j$, we have the following:

If a is an invariant point of $\underset{\sim}{f}$, and $\underset{\sim}{f}$ and $\partial\underset{\sim}{f}/\partial\underset{\sim}{x}$ are continuous in a neighbourhood of a, and if $|\partial\underset{\sim}{f}/\partial\underset{\sim}{x}|^* < I$ at a, then there is a neighbourhood of a such that any approximation to a in this neighbourhood is refined indefinately by its successive iterated images under $\underset{\sim}{f}$.

Also we have:

If $\underset{\sim}{f}$ and $\partial\underset{\sim}{f}/\partial\underset{\sim}{x}$ are continuous in a region where the condition $|\partial\underset{\sim}{f}/\partial\underset{\sim}{x}|^* \leq k < I$ holds, then there is a number $d > 0$ such that for any point $\underset{\sim}{x}^\circ$ for which $|\underset{\sim}{f}\underset{\sim}{x}^\circ - \underset{\sim}{x}^\circ| < d$, an invariant point of $\underset{\sim}{f}$ is determined as the limit point of the iterated images under $\underset{\sim}{f}$ of $\underset{\sim}{x}^\circ$. If further, for $k < k' < I$, $|\underset{\sim}{F}| \leq k'$, when F is the matrix $\partial\underset{\sim}{f}/\partial\underset{\sim}{x}$ with its rows independently evaluated at points within a distance d of each other in the sphere with centre $\underset{\sim}{x}^\circ$ and radius

$a/(1-k')$, <u>then this sphere contains an invariant point of f which is the limit point of the iterated images under f of its centre.</u>

We may briefly illustrate with a pair of equations

$$x_1 = f_1(x_1, x_2), \quad x_2 = f_2(x_1, x_2),$$

and consider the condition $|\partial(f_1, f_2)/\partial(x_1, x_2)|^* < 1$. Our theorem in this case give criteria in the convenient form

$$|\partial f_1/\partial x_1|^2 + |\partial f_2/\partial x_1|^2 < 1,$$

$$|\partial f_1/\partial x_1|^2 + |\partial f_2/\partial x_1|^2 + |\partial f_1/\partial x_2|^2 + |\partial f_2/\partial x_2|^2 - |\partial f_1/\partial x_1 \, \partial f_2/\partial x_2 - \partial f_1/\partial x_2 \, \partial f_2/\partial x_1|^2 < 1.$$

REFERENCES

I. H. Hotelling, Some new methods in matrix calculation, Ann.Math.Stat.,I4(I943),I-34.

2. J.H.M. Wedderburn, Lectures on Matrices(New York I934), I25;I7

3. F. Rella, Uber den absoluten Betrag von Matrizen, International Congress of Mathematicians at Oslo (I936),II,29.

4. S.N. Afriat, Bounds for the characteristic values of matrix functions, Quart.J.Math.Oxford(2),2(I95I),8I-4

5. R.A. Frazer, W.J. Duncan and A.R. Collar, Elementary Matrices (Cambridge I938),I34.

6. J.H.M. Wedderburn, The absolute value of a product of two matrices, Bull.Amer.Math.Soc.,3I(I925),304-308.

7. A.H. Bowker, On the norm of a matrix,Ann.Math.Stats. I8(I947),285-288.

8. R. Oldenburger, Infinite powers of matrices and characteristic values, Duke Math.Journal,Vol.6(I940), 257.

9. P.Halmos, Finite Dimensional Vector Spaces (Princeton I948), 7I-4.